

Structural studies of eukaryotic and bacterial RecQ complexes

By

Kelly A. Manthei

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Biochemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2013

Date of final oral examination: 05/13/2013

The dissertation is approved by the following members of the Final Oral Committee:
James L. Keck, Professor, Biomolecular Chemistry
Michael M. Cox, Professor, Biochemistry
Catherine A. Fox, Professor, Biomolecular Chemistry
Melissa M. Harrison, Assistant Professor, Biomolecular Chemistry
Randal S. Tibbetts, Associate Professor, Human Oncology

Structural studies of eukaryotic and bacterial RecQ complexes

Kelly Antoinette Manthei (née Hoadley)

Under the supervision of Professor James L. Keck

University of Wisconsin-Madison

ABSTRACT

Cells have evolved many conserved mechanisms to repair continual genomic damage that is an ongoing threat. The importance of these processes is highlighted by multiple rare genetic diseases, in particular Bloom syndrome (BS) and Fanconi anemia (FA). Deleterious mutations in the gene encoding the RecQ-family DNA helicase BLM or any one of fifteen FA genes give rise to a similar disease that is manifested by a predisposition to cancer and genomic instability. Furthermore, these diseases are molecularly linked through protein-protein interactions. The RMI subcomplex, comprised of RMI1 and RMI2, interacts with BLM and Topo III α to form a complex referred to as the BLM dissolvasome, which acts to prevent DNA crossovers following homologous recombination-dependent double-strand break repair. The RMI subcomplex also interacts with FA protein FANCM, which is proposed to act by bringing the BLM dissolvasome to sites of DNA damage, specifically DNA interstrand crosslinks that are covalent linkages between DNA strands where a replication fork would stall.

In this work I describe the x-ray crystal structure of the RMI core complex, which is comprised of two oligonucleotide/oligosaccharide binding folds that interact in a similar manner as replication protein A. This structure directed mutational studies that assessed the importance of RMI subcomplex formation *in vivo*. Mutations that disrupt the RMI1/2 interface led to higher levels of crossover, or sister chromatid exchange (SCE), in cells. Next, I showed through both biochemical and crystallographic characterization that the RMI1/2 interface serves as the FANCM interaction site. When critical residues at this interface are mutated, levels of SCEs are also increased. I provide evidence that the interaction between the RMI subcomplex and FANCM is critical for a cell to maintain genomic integrity.

In addition to work on the RMI subcomplex, I also solved the first structure of a bacterial RecQ DNA helicase in complex with DNA. Both small- and large-scale rearrangements were observed between the apo and DNA-bound structure, and these movements are proposed to play a critical role in DNA unwinding. These studies will allow future characterization of DNA unwinding by both prokaryotic and eukaryotic RecQ DNA helicases.

CITATIONS

This dissertation contains the following primary research articles and a review article:

Hoadley KA, Xu D, Xue Y, Satyshur KA, Wang W, Keck JL (2010) “Structure and cellular roles of the RMI core complex from the Bloom syndrome dissolvasome”. *Structure* **18**, 1149-58.

Hoadley KA, Xue Y, Ling C, Takata M, Wang W, Keck JL (2012) “Defining the molecular interface that connects the Fanconi anemia protein FANCM to the Bloom syndrome dissolvasome”. *Proc Natl Acad Sci U S A* **12**, 4437-42.

Smith BC, Anderson, MA, Hoadley KA, Keck JL, Cleland WW, Denu JM. (2012) “Structural and Kinetic Isotope Effect Studies of Nicotinamidase (Pnc1) from *S. cerevisiae*”. *Biochemistry* **51**, 243-256.

Manthei, K.A. and Keck, J.L. (2013) “The BLM dissolvasome in DNA replication and repair”. *Cellular and molecular life sciences* (Epub ahead of print).

ACKNOWLEDGEMENTS

There are many who have supported me throughout this process and beyond. First, I'd like to thank my parents, Doug and Brenda Hoadley, who have supported and loved me without condition from the day I was born. I am so grateful for all the opportunities they have given me, and for setting such a good example for me to follow. I am so blessed to have them and their love in my life. I'm also thankful for my siblings, Mark and Erin, who have also provided support, friendship, and love throughout the years. My brother-in-law Ben has been a wonderful addition to the family, and I am thankful for the joy Asher and Jude have brought to our family. I hope we can all continue to support them in the same way our parents provided for us. I would also like to thank my newer family, Dennis, Barbara, Gretchen, Dawn, Vern, Erich and Lara, for welcoming me into their family and supporting David and me. We are both grateful to have supportive families and to be raised with similar values.

I am thankful for the many friends who have supported me throughout the years. My childhood friends were always so much fun and I am indebted to the LHS music department for providing the best family while in high school. To my friends from UIUC and Barcelona, who supported and strengthened my faith. I would specifically like to think Amy, Kim, Eli, Dia, Erin, and both Kendras. In Madison I have been so thankful to make so many new friends throughout graduate school and while I was skeptical of the other girl who joined the Keck lab at the same time, Aimee has turned out to be the best friend I could ever have asked for. I am so grateful for her

love and support and hope that we will continue to stay close even when our desks are not facing each other.

I owe gratitude to Dr. Scott Silverman, who gave me an incredible opportunity to begin research my sophomore year at Illinois. I also would like to thank Dr. Amber Flynn-Charlebois and the entire Silverman lab for their training and support as I began research. I am also thankful for Dr. Karin Musier-Forsyth, who took a chance and hired me as a technician while I was still overseas. She gave me the opportunity to get back into research and encouraged me to apply for graduate school within months of joining her lab.

I would like to thank my committee members for their encouragement, support, and ideas throughout my time in graduate school – Drs. Michael Cox, Catherine Fox, Randal Tibbetts, and Melissa Harrison. I would also like to remember and thank Dr. Paul Bertics for his support and advice.

I have thoroughly enjoyed my time and the company of the Keck Lab. Thank you to former members whom I overlapped with: Duo Lu, Nick George, Mike Killoran, Rob Shereda, Scott Lindner, and Aimee Marceau. Without their guidance I would have been lost as a new graduate student. Thanks also to the current lab: Basu Bhattacharyya, Sarah Wessel, Chrissy Petzold, and Angel Myers, who make spending time in lab fun and rewarding.

I am so thankful for the path that brought me to the Keck Lab, for MBTG which introduced me to Jim and caused me to consider joining the lab. It was the environment of the Keck Lab and the immediate realization that Jim would be an excellent mentor that caused me to join the lab. I would like to thank Jim for living up to those expectations. I am grateful for his open door policy, advice about research and life, excitement about my research (even when I had to kick him off the computer so I could work on it myself), terrible jokes, and incredible support.

Last, I would like to thank my husband David for his love and support. I am so blessed to have you in my life, and you continually amaze me with your encouragement and devotion to me. I am so glad to have you by my side and look forward to whatever, and wherever, comes next.

TABLE OF CONTENTS

ABSTRACT	i
CITATIONS	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	vii
LIST OF FIGURES	xii
LIST OF TABLES	xiv
CHAPTER 1. The BLM dissolvasome in DNA replication and repair	1
Abstract	2
Introduction.....	3
Bloom syndrome	3
BLM is a member of the RecQ family of DNA helicases	4
Biochemical properties of BLM	6
RecQ complex formation with Topo III	8
BLM interactions with RMI1 and RMI2	11
BLM and the BLM dissolvasome in double-strand break repair.....	13
The role of the BLM dissolvasome in replication restart	16
Links with Fanconi Anemia protein FANCM	18
BLM/FANCI interaction	21
BLM and FA proteins both act at the replication-dependent checkpoint	21
The BLM dissolvasome and FA proteins are both found associated with ultra-fine anaphase bridges.....	23
Concluding remarks	25
Acknowledgements.....	28
Figures.....	29
Table	35
References.....	36
CHAPTER 2. Structure and cellular roles of the RMI core complex from the Bloom syndrome dissolvasome	55
Summary	56
Introduction.....	57

Results.....	60
Defining the RMI core complex	60
Crystal structure of the RMI core complex.....	60
Comparison of the RMI core complex and the RPA trimerization core.....	61
The RMI core complex interface is distinct from RPA	63
Coimmunoprecipitation experiments demonstrate the importance of RMI core complex interface.....	64
Disrupting the RMI core complex interface increases SCE formation <i>in vivo</i>	65
Discussion	68
Comparison of the RMI core complex and the RPA trimerization core.....	68
Importance of RMI subcomplex interface in suppressing SCE levels	70
Concluding Remarks.....	72
Experimental Procedures	73
Plasmids, Protein Overexpression, and Purification.....	73
Limited Proteolysis Assay	75
Protein Crystallization and Structure Determination.....	75
Coimmunoprecipitation and SCE assays	76
Acknowledgements.....	77
Figures.....	78
Table	89
References.....	90
CHAPTER 3. Defining the molecular interface that connects the Fanconi anemia protein FANCM to the Bloom syndrome dissolvasome.....	94
Abstract.....	95
Introduction.....	96
Results.....	100
A role for RMI2 Lys121 in linking the dissolvasome to FANCM.....	100
X-ray crystal structure of FANCM MM2 bound to the RMI core complex.....	102
Assessing the importance of RMI/FANCM interactions <i>in vitro</i>	103
The RMI/FANCM interaction is important for suppressing cellu ^r SCEs	105
Discussion	106
Structure of the RMI/FANCM interface.....	106

Insights into the architecture of the BLM dissolvasome/FA core complex superstructure	107
Possible regulation of RMI/FANCM complex assembly	108
A model for complex coordination and SCE suppression.	109
Materials and Methods.....	111
Plasmids, Protein Overexpression, and Purification.....	111
FITC labeling of MM2 (F-MM2)	112
Size-exclusion Chromatography	113
Fluorescence polarization assays	113
RMI/MM2 complex crystallization and structure determination	114
Co-immunoprecipitation and SCE assays.....	114
Acknowledgements.....	116
Figures.....	117
Table	129
References.....	130
CHAPTER 4. Crystal structure of a DNA-bound bacterial RecQ helicase reveals substrate-dependent conformational changes	133
Summary	134
Introduction.....	135
Results.....	138
Interactions between CsRecQ and DNA.....	139
Reorientation of the ARL.....	140
Biochemical studies confirm the importance of certain residues and interactions.....	141
SAXS confirms movement of the WH domain	143
Discussion	145
Comparison between DNA-bound CsRecQ and human RecQ structures	145
The ARL is critical in coupling ATP hydrolysis to DNA unwinding	146
Crystal packing may offer hints at an oligomeric state.....	147
Future experiments.....	149
Experimental Procedures	150
Purification of RecQ proteins	150
CsRecQ/DNA complex crystallization and structure determination	150

ATPase assays.....	151
DNA Binding Assay	152
Helicase assays.....	152
Small angle X-ray scattering data collection	153
SAXS analysis and modeling.....	153
Acknowledgments.....	155
Figures.....	156
Tables.....	173
References.....	176
CHAPTER 5. Summary, Conclusions, and Future Directions.....	181
The importance of the RMI subcomplex	182
The RMI subcomplex interacts with FANCM at the RMI1/RMI2 interface.....	183
Future directions: a small-molecule screen to identify RMI/FANCM inhibitors.....	184
The structure of a bacterial RecQ in complex with DNA.....	184
APPENDIX 1. Structural and Kinetic Isotope Effect Studies of Nicotinamidase (Pnc1) from <i>S. cerevisiae</i>	186
Abstract.....	187
Introduction.....	188
Experimental Procedures	191
Expression and purification of Pnc1	191
Mutagenesis of Pnc1	191
Determination of kinetic parameters.....	191
Determination of K_i values	192
Protein crystallization and structure determination	193
^{15}N kinetic isotope effects.....	193
^{13}C kinetic isotope effects	195
Fraction of Reaction.....	197
Kinetic isotope effects.....	197
Dependence of activity on pH.....	198
Results.....	200
Product inhibition and irreversibility of reaction.....	200

Inhibition by nicotinamide analogs.....	201
Structure of nicotinaldehyde inhibited Pnc1	201
Kinetic parameters of Pnc1 active-site mutants.....	203
Kinetic parameters of nicotinamide analogs as alternate substrates	204
^{15}N and ^{13}C kinetic isotope effects with substrate analogs	205
^{15}N and ^{13}C kinetic isotope effects of Pnc1 mutants.....	205
pH dependence of kinetic parameters	206
Discussion	208
Nicotinamidase kinetic mechanism	208
Nicotinamide binding may proceed through a hydrophobic tunnel.....	208
Zinc is involved in both binding and activation of substrate	209
C167 is activated by Pnc1 active-site residues	211
Structure-activity relationships and rate-limiting steps of nicotinamide analogs.....	213
Nicotinamidase chemical mechanism.....	216
Conclusion	217
Supporting Discussion	219
Derivation explaining observed kinetic isotope effects for Pnc1 D51N.....	219
Acknowledgements.....	221
Figures.....	222
Tables	230
Schemes	235
References.....	237

LIST OF FIGURES

Figure 1.1. Properties of the RecQ helicases and related proteins.....	29
Figure 1.2. Pathways for the resolution of dHJ structures.....	31
Figure 1.3. Roles for the BLM dissolvasome in HR-dependent DSB repair.....	32
Figure 1.4. BLM dissolvasome and FA protein interactions described in this review	34
Figure 2.2. Structural features of the RMI core complex	80
Figure 2.3. RMI core complex interface.....	82
Figure 2.4. Electron density of the Asp141 interaction	84
Figure 2.5. Coimmunoprecipitations of the dissolvasome complex components with RMI2 variants.....	85
Figure 2.6. SCE frequencies observed with different RMI2 variants.....	86
Figure 2.7. Structure of the RMI2 Lys121Ala variant and surface proximity of Lys121	87
Figure 3.1. RMI2 Lys121Ala is defective in binding FANCM in vitro and in vivo	117
Figure 3.2. Size exclusion chromatographic analysis confirms that the RMI core complex interacts with GST-MM2	119
Figure 3.3. The RMI subcomplex acts as a bridge between FANCM and Top3 α	120
Figure 3.4. The structure of the RMI core complex bound to the MM2 peptide.....	121
Figure 3.5. The N-terminus of RMI2 orders to form β -sheet interactions with the MM2 peptide	123
Figure 3.6. Differences between the apo and MM2-bound RMI core complex	124
Figure 3.7. MM2 variants have defects in binding the RMI core complex	126

Figure 3.8. Co-immunoprecipitations using various chicken FANCM variants in HEK293 cells	127
Figure 3.9. FANCM variants with defects in binding increase SCEs	128
Figure 4.1. Sequence alignment between EcRecQ, CsRecQ, and BLM	156
Figure 4.2. Previous RecQ structures	158
Figure 4.3. The structure of CsRecQ bound to DNA	160
Figure 4.4. dsDNA binds to the WH domain similar to WRN	162
Figure 4.5. Details of the ssDNA binding interface.....	164
Figure 4.6. DNA connects the symmetry mates	166
Figure 4.7. ARL movement in response to DNA binding	168
Figure 4.8. EcRecQ and variants binding to F-3prOH	170
Figure 4.9. The WH reorients upon DNA-binding.....	171
Figure A1.1. Double-reciprocal inhibition plots.....	222
Figure A1.2. Structure of nicotinaldehyde covalently bound in the active site of yeast Pnc1 ..	223
Figure A1.3. Multiple-sequence alignment of the nicotinamidases for which structures have been solved.....	225
Figure A1.4. pH rate profiles	227
Figure A1.5. Proposed kinetic mechanism of the nicotinamidase reaction.....	228
Figure A1.6. Putative route of nicotinamide access to the Pnc1 active site	229

LIST OF TABLES

Table 1.1. A summary of studies examining cellular SCE	35
Table 2.1. X-ray data collection and structure determination statistics	89
Table 3.1. X-ray data collection and structure determination statistics	129
Table 4.1. X-ray data collection and structure determination statistics	173
Table 4.2. DNA-independent and -dependent ATPase activity	174
Table 4.3. Binding to DNA substrate F-3prOH	175
Table A1.1. Primers used in Pnc1 mutagenesis	230
Table A1.2. K_i values of Pnc1 competitive inhibition by nicotinamide analogs	231
Table A1.3. X-ray data collection and structure determination statistics	232
Table A1.4. Steady-state kinetic parameters of Pnc1 mutants	233
Table A1.5. Steady-state kinetic parameters and ^{15}N and ^{13}C KIE for the reaction catalyzed by Pnc1 with alternate substrates	234

CHAPTER 1

The BLM dissolvasome in DNA replication and repair

This chapter written by Kelly A. Manthei and has been published as a review article:

Manthei, K.A. and Keck, J.L. (2013) “The BLM dissolvasome in DNA replication and repair”.

Cellular and molecular life sciences (Epub ahead of print).

Abstract

RecQ DNA helicases are critical for proper maintenance of genomic stability and mutations in multiple human RecQ genes are linked with genetic disorders characterized by a predisposition to cancer. RecQ proteins are conserved from prokaryotes to humans and in all cases form higher-order complexes with other proteins to efficiently execute their cellular functions. The focus of this review is a conserved complex that is formed between RecQ helicases and type-I topoisomerases. In humans, this complex is referred to as the BLM dissolvasome or BTR complex, and is comprised of the RecQ helicase BLM, topoisomerase III α , and the RMI proteins. The BLM dissolvasome functions to resolve linked DNA intermediates without exchange of genetic material, which is critical in somatic cells. We will review the history of this complex and highlight its roles in DNA replication, recombination, and repair. Additionally, we will review recently established interactions between BLM dissolvasome and a second set of genome maintenance factors (the Fanconi anemia proteins) that appear to allow coordinated genome maintenance efforts between the two systems.

Introduction. Genomic instability is a shared phenotype in multiple hereditary diseases where DNA repair or replication-dependent checkpoint genes are mutated. In these disorders, malfunctions in the normal genome maintenance pathways of cells increase the chance that disease-causing changes will occur that allow for uncontrolled cell proliferation (1-3). Through the study of one such rare genetic disease, Bloom syndrome (BS), much has been learned about how cells ensure genomic fidelity. The gene mutated in BS encodes a RecQ-family DNA helicase, BLM, and the pronounced genomic instability of BS cells reflects the protein's multiple roles in DNA replication, recombination, and repair (2, 4, 5). In coordination with a large number of protein partners, BLM aids in cellular responses to replication stress and DNA damage. In this review we focus on the discovery of a key BLM-containing complex referred to as the "BLM dissolvasome" or "BTR complex" (6) that includes topoisomerase III α (Topo III α) and the RMI proteins, and on recent advances in understanding its functions. The coordinated action of the BLM helicase and Topo III α in this complex is critical for untangling intertwined DNA structures that can arise during DNA replication or repair. Recent insights into interactions between the BLM dissolvasome and FANCM will also be reviewed to provide an overview of the importance of coordinating the functions of these proteins in cells, particularly for the prevention of crossover formation in mitotic cells.

Bloom syndrome. BS is a rare autosomal recessive genetic disorder that was first described in 1954 by dermatologist David Bloom (7). BS manifestations include proportional dwarfism, sun sensitivity, increased susceptibility to infections and diabetes, and a predisposition to a broad array of cancers. Cancer is a frequent cause of death in BS patients, with the first diagnosis often

occurring early in life (mean age is in the mid-twenties) (4, 8, 9). Cells derived from BS patients display a remarkable increase in chromosomal damage as evidenced by DNA rearrangements, gaps, and breaks (8, 10). The major distinguishing genomic instability feature of somatic BS cells, which is used in diagnosis, is a ten-fold elevated frequency of sister-chromatid exchanges (SCEs) (4, 8, 11). One proposed mechanism that would lead to formation of these crossovers is the asymmetric resolution of Holliday junction (HJ) intermediates formed during homologous recombination (HR) reactions. The abnormal genomic instability behavior of BS cells provided an early clue that the gene mutated in the syndrome (*BLM*) encoded a protein that had important roles in suppressing chromosomal crossovers and maintaining genomic stability (11).

BLM is a member of the RecQ family of DNA helicases. In the mid-1990s, the *BLM* gene was mapped to chromosome 15 and positional cloning techniques identified its gene product (BLM) as a member of the RecQ family of DNA helicases (12, 13). Many of the mutations in *BLM* that give rise to BS are frameshift or nonsense mutations that lead to premature translation termination of BLM, resulting in a truncated protein. Since the nuclear localization signal (NLS) of BLM is encoded in its C-terminus, these truncation products are expected to fail to reach the nucleus (see the Bloom Syndrome registry, <http://med.cornell.edu/bsr/genetics/> and (14)) (Figure 1.1A). In addition, there are thirteen known BS-causing missense mutations, seven in the helicase domain and six in the “RecQ C-terminal” (RQC) domain (13-17). Through biochemical and structural studies, these mutations have been shown to confer loss of catalytic activity or are expected to hinder protein folding or stability (17-20).

The discovery of the *BLM* gene helped to explain the molecular basis for BS, since the RecQ helicase in *Escherichia coli* was already known to play a role in bacterial genomic integrity (21). At the time that BLM was discovered, ATP-dependent DNA helicase activity had been demonstrated for *E. coli* RecQ and roles for the protein in the RecF recombination pathway had been established (22). This background knowledge proved valuable for quickly ascertaining the functions of BLM and other eukaryotic RecQ DNA helicases. Bacteria and yeast typically code for a single RecQ protein, whereas humans have five: RecQ1, BLM, WRN, RecQ4, and RecQ5 (Figure 1.1). In addition to BS, two other recessive genetic diseases are caused by mutations in the *WRN* and *RecQ4* genes (Werner's syndrome (WS), (23) and Rothmund-Thomson syndrome (RTS), (24), respectively). Two more diseases, RAPADILINO syndrome and Baller-Gerold syndrome, both share some clinical characteristics with RTS and have also been linked to a subset of mutations of *RECQ4* (25, 26). As with BS, WS and RTS are characterized by increased genomic instability and cancer predisposition, although the breadth of cancers observed in WS and RTS is narrower than that of BS and the characteristic increase in SCEs of BS cells is not seen in cell lines derived from WS and RTS patients.

As is the case for all RecQ family members, BLM is a superfamily 2 (SF2) helicase that uses the energy derived from ATP hydrolysis to unwind DNA (13, 27). In addition to a conserved helicase domain that provides the motor function for DNA unwinding, BLM and most other identified RecQs also contain RQC and the "helicase and RNaseD C-terminal" (HRDC) domains (28) (Figure 1.1A). The RQC domain is comprised of a Zn^{2+} -binding scaffold and a winged-helix (WH) subdomain, and in some cases this domain has been shown to be important for

structure-specific binding of RecQ proteins to replication fork, HJ, or G-quartet DNA structures (20, 29-33). The HRDC domain is an independent globular domain that can confer further substrate specificity to RecQ helicases and can potentially enhance DNA unwinding processivity (34-37). Interestingly, although nonsense and frame-shift insertion/deletion mutations that cause BS span the entire coding-region of the *BLM* gene, missense mutations are restricted to the helicase and RQC domains, indicating that DNA binding and unwinding functions are critical for BLM cellular functions (14). In addition to these conserved domains, many eukaryotic RecQ helicases contain additional N- and C-terminal elements that encode for additional enzymatic functions (*e.g.* the exonuclease domain of WRN (Figure 1.1A)), provide binding sites for heterologous proteins, receive post-translational modifications, and/or facilitate proper subcellular localization (5, 27).

Biochemical properties of BLM. The BLM helicase is a 1417-residue ATP-dependent 3'-5' DNA helicase (4, 38). The 3'-5' directionality indicates that the preferred substrate for BLM unwinding contains a 3' single-stranded (ss) extension to which the helicase is thought to bind and translocate along; this directionality is conserved among RecQ helicase family members. Although BLM is capable of unwinding simple double-stranded (ds) DNA substrates, it preferentially acts on more elaborate DNA structures such as G-quartet, D-loop, and HJ DNAs (32, 39-43). When combined with the high level of chromosomal abnormalities observed in BS cell lines and the known roles for the prototypical RecQ in *E. coli*, these preferences suggested early on that the BLM helicase could have an important role in resolving unusual DNA structures that can arise in cells through processes such as homologous recombination.

BLM has been shown to form a large number of functionally important homooligomeric and heterooligomeric complexes. Full-length BLM can form homohexameric structures *in vitro* and an isolated domain comprised of the N-terminal 431 residues of BLM can self-associate, suggesting that this region aids in BLM oligomerization (44, 45). More recent results have suggested that BLM oligomers can dissociate into monomers upon ATP hydrolysis, and the helicase may function as a monomer during DNA unwinding (46). However, this does not rule out the possibility that higher-order oligomers are required to bind and/or unwind more complex DNA structures such as G-quartet and HJ DNA. Interestingly, the N-terminal region of another human RecQ, RecQ1, is required for both oligomerization and HJ unwinding (42). Furthermore, RecQ1 has been shown to form higher-order oligomers upon binding DNA, but these dissociate to monomers or dimers upon ATP-binding, suggesting that RecQ1 may unwind DNA as monomers or lower-order oligomers (47). The N-terminal exonuclease domain of WRN also appears to be responsible for hexamer formation, and this isolated exonuclease domain acts catalytically as a hexamer (48) (Figure 1.1A). Oligomerization of full-length WRN is stimulated by DNA binding in a similar manner to RecQ1 (49, 50). For BLM, constructs lacking the N-terminal oligomerization domain are able to unwind HJ DNA, and instead it appears that the HRDC domain is critical for HJ unwinding (37, 51) (Figure 1.1A). More research will be required in order to elucidate the roles of oligomerization for the BLM helicase for DNA binding and unwinding *in vivo*.

By itself, BLM is a poorly processive helicase and is only able to unwind duplexes of less than ~100 bp (52). However, the addition of the eukaryotic single-stranded DNA-binding protein

(Replication Protein A, RPA) strongly stimulates BLM DNA unwinding of longer duplexes, whereas the *E. coli* single-stranded DNA-binding protein (SSB) fails to stimulate BLM unwinding (52). Since BLM also interacts with RPA, it is thought that BLM/RPA complex formation is important for increasing the processivity of BLM DNA unwinding (52). BLM has also been shown to possess ssDNA annealing activity; however, RPA inhibits this activity. Since RPA or other DNA-binding proteins are bound to exposed ssDNA, the cellular relevance of this activity is unclear (53). Similar interactions and/or stimulation by SSBs have been observed in several other RecQ family members including *E. coli* RecQ/SSB, *Saccharomyces cerevisiae* Sgs1/RPA, human and *Caenorhabditis elegans* WRN/RPA, human RecQ1/RPA, and human RecQ5/RPA (54-71). In addition to its interactions with RPA, numerous studies have shown that BLM can interact with a large number of protein partners that have roles in DNA replication, recombination, and repair, including topoisomerase 3 α (Topo III α), RMI1, Topo II α , PICH, RAD51, EXO1, DNA2, FANCD1, RIF1, ATM, ATR, FEN1, DNA polymerase δ , MLH1, p53, WRN, and TRF2 (72-92). The focus of this review is the complex formed by BLM, Topo III α , RMI1, and RMI2, and therefore many of these other important interactions will not be discussed further (see reviews (4, 5)).

RecQ complex formation with Topo III. Physical and functional interactions with topoisomerases have been found to be critical for the genome maintenance activities of many RecQ proteins. Topoisomerases function to relieve the torsional stress that arises from DNA supercoiling by transiently breaking the DNA backbone in either one (type I) or both (type II) strands and then passing intact DNA through the opening before resealing the break (93).

Although RecQ proteins have been found to associate with both type-I and type-II topoisomerases, the partner that we focus on in this review is Topo III, a type-I topoisomerase. Linkages between RecQ proteins and Topo IIIs have been apparent since the discovery of the *S. cerevisiae* RecQ protein (Sgs1) as a suppressor of the slow-growth phenotype of cells lacking Topo III (94). The N-terminus of Sgs1 interacts with Topo III, and this interaction appears to be the sole function of the first 100 amino acids of Sgs1 (94-97). The discovery of this interaction in yeast prompted investigations in other organisms; for example, *E. coli* RecQ was shown to functionally interact with Topo III as well to stimulate DNA catenation/decatenation and the resolution of converging DNA replication forks (98-100). In humans, the N-terminus of BLM is responsible for interaction with Topo III α and both proteins localize to promyelocytic leukemia protein (PML) nuclear bodies (72, 101, 102). Furthermore, a study examining the effects of deleting the Topo III α -interaction domain from BLM (residues 1-133) revealed an arrangement in which interaction with BLM recruits Topo III α to PML bodies and showed that this interaction is important in suppressing SCEs (102). BLM appears to recruit Topo III α to DNA and to stimulate its decatenase activity (102-104). These functional RecQ/Topo III interactions appear to be conserved from bacteria to humans, and in all cases, the RecQ helicase stimulates the enzymatic activity of the topoisomerase (94, 98, 103, 104).

Although the cellular importance of RecQ/Topo III interactions had been established by the late 1990s, the precise substrates on which the BLM/Topo III α complex operated (especially with respect to the elevated SCEs observed in BS cells) remained unresolved for some time. An important breakthrough on this question was made by the Hickson group when they considered

the types of substrate that might require both BLM and Topo III α (105). Since there are no torsional constraints in simple HJ substrate that would mimic DNA *in vivo*, they sought out a constrained HJ-based substrate -- a double Holliday junction (dHJ). It was discovered that BLM and Topo III α worked together to resolve this substrate in a way that prevented exchange of flanking DNA, a process termed “dissolution” (105) (Figure 1.2A). This discovery also provided a satisfying explanation for the high levels of SCEs in BS cells. dHJ structures can arise in HR-dependent double-strand break (DSB) repair (106) and can be acted upon by either the BLM/Topo III α dissolution pathway or by HJ resolvases, which cleave symmetrically or asymmetrically to form a mixture of crossover and non-crossover products (107, 108) (Figure 1.2B). In BLM-deficient BS cells, resolvase-mediated cellular processes would be the dominant pathways for dealing with dHJ structures, which would lead to the exchange of large pieces of homologous chromosomes. In the process of dissolution, BLM is proposed to catalyze branch migration to merge the two HJs into a hemicatenated structure upon which Topo III α acts to create exclusively non-crossover products (6) (Figure 1.2A). This exclusivity is thought to be important to avoid a “loss of heterozygosity” (LOH) that can occur when the chromatids are separated into daughter cells, which might be tied to the increased cancer predisposition seen in BS patients (109, 110). While this mechanism provides an explanation for the increase in SCEs observed in BS cells, BLM and the BLM dissolvasome have multiple roles in cells beyond dHJ dissolution, and their roles in other pathways also prevent crossover formation (see **BLM and the BLM dissolvasome in double-strand break repair**).

BLM interactions with RMI1 and RMI2. Although the BLM/Topo III α connection provided an important link for understanding the origins of SCEs in BS cell lines, two additional proteins were subsequently found to be important for stabilizing the complex in HeLa cell nuclear extracts – RMI1 and RMI2 (RMI stands for *RecQ-mediated genome instability*) (Figure 1.1). Together, the BLM/Topo III α /RMI1/RMI2 complex is referred to as the “BLM dissolvasome” or “BTR complex.”

RMI1 was identified as a 625-residue protein that co-purified at near-stoichiometric levels with human BLM and Topo III α (73) (Figure 1.1). The N-terminus of RMI1 directly interacts with BLM and Topo III α and stimulates dHJ dissolution greater than ten-fold *in vitro* (111-113). This is accomplished by stimulation of Topo III α , since a variant of RMI1, K166A, which is able to bind BLM but not Topo III α , cannot stimulate dHJ dissolution (104, 113). RMI1 contains N- and C-terminal oligonucleotide/oligosaccharide binding (OB) domains, which led to speculation that it may bind DNA and stimulate Topo III α through recruitment effects (73, 114) (Figure 1.1). However, RMI1 has only very weak DNA binding affinity in its C-terminal OB domain, which is dispensable for dHJ stimulation *in vitro* (73, 111, 113, 115). Instead, RMI1 appears to be important to the stability of the BLM/Topo III α /RMI1 complex, since depletion of RMI1 significantly increases SCE levels and affects cellular protein levels of Topo III α and, to a lesser extent, BLM (73). Since RMI1 is able to stimulate Topo III α activity *in vitro*, it might also stabilize a conformation of Topo III α with increased catalytic activity (111, 112). In *S. cerevisiae*, a homolog of the human RMI1 (Rmi1) was identified as a smaller, 241-residue

protein that appears to be most similar to the N-terminus of human RMI1 in that it forms a complex with Sgs1 and Topo III and stimulates Topo III activity (116-118).

RMI2 was subsequently discovered as a binding partner to RMI1 in human cells, and the two together comprise the RMI subcomplex. RMI2 is a 147-residue protein comprised entirely of an OB fold that interacts with the C-terminal OB fold of RMI1 (115, 119) (Figure 1.1). The RMI subcomplex is formed through interactions between these OB folds in a manner similar to that of RPA70/RPA32, however, the RMI subcomplex lacks the ability to bind DNA and there is no conservation with the DNA-binding residues found in the RPA complex (115, 120, 121) (Figure 1.1). In HeLa cells, depletion of either RMI1 or RMI2 by siRNA significantly reduces the protein levels of the other, indicating their interdependence (115, 119). Furthermore, their knockdown leads to a strong decrease in Topo III α protein levels and a more modest decrease in BLM protein levels (115, 119). Therefore, it appears that RMI1, RMI2, Topo III α are more highly interdependent *in vivo* (perhaps forming a stable complex), whereas BLM may be a dissociable component. In addition, studies in hyper-recombinant chicken DT40 cells reveal that knockdown of RMI2 leads to an increase in SCE levels, and when both BLM and RMI2 are knocked down the levels are the same as BLM alone, indicating that RMI2 works with BLM to prevent crossovers (73, 115) (Table 1.1). Mutations that destabilize the RMI1/RMI2 interface also lead to increased levels of SCEs (121). In response to DNA damaging agents and replication inhibitors, RMI1 and RMI2 both co-localize with BLM to nuclear foci associated with DNA damage (115). Additionally, when either RMI1 or RMI2 is depleted, BLM localizes

to significantly fewer nuclear foci (73, 115). These studies confirm the importance of the RMI subcomplex and highlight the central importance of the BLM dissolvasome in DNA repair.

BLM and the BLM dissolvasome in double-strand break repair. Numerous studies have identified both pro- and anti-recombinogenic activities for BLM in DSB repair, even though the high levels of crossovers that form in its absence suggest a primarily anti-recombinogenic role. In fact, BLM may shuttle DSB repair intermediates towards a “pro-recombination” pathway in order to ensure they are resolved via dissolution, and therefore without crossover (9). Evidence supports roles for BLM in multiple steps in HR-dependent DSB repair, including 5' end resection, RAD51 filament and D-loop formation, and, as described earlier, in resolving dHJ structures (Figure 1.3). These roles are reviewed in the remainder of this section.

Multiple labs have shown that BLM is able to stimulate nuclease activity in 5' end resection (9, 77, 78, 122) (Figure 1.3). In humans, two pathways for end resection have been discovered, one involving the nuclease DNA2 and one involving exonuclease 1 (EXO1), and in both of these pathways BLM physically interacts with and stimulates the nuclease (77, 78). While both RecQ1 and WRN are able to simulate other EXO1 catalytic activities, only BLM has been shown to specifically stimulate the end resection activity of EXO1 and DNA2 *in vitro* (77, 78, 123, 124). However, in *Xenopus laevis* cell-free extracts, WRN was shown to function with DNA2 in end resection, yet more research is required before a conserved role for WRN can be established (125). The most extensive studies have been performed in yeast, where Sgs1 has been shown to stimulate Dna2 (126, 127). This is relevant *in vivo*, since when *exo1* and *sgs1* are simultaneously

deleted in yeast, thus disrupting both pathways for end resection, cells become sensitive toward a wide range of DNA damaging agents (122). By monitoring end resection in cells, it was further shown that Sgs1, Topo III α , and Rmi1 are all involved in the same resection pathway (127). In addition, the Sgs1/Dna2 pathway has been reconstituted *in vitro* and end resection was stimulated by the addition of Rmi1/Topo III α (128, 129). In this context, Rmi1/Topo III α are together able to stimulate the helicase activity of Sgs1 and are thought to aid Sgs1 DNA binding, however, this stimulation does not appear to require the catalytic activity of Topo III α (128, 129). Further research will be required to determine the importance of this stimulation *in vivo* and whether similar biochemical attributes are conserved in humans. Nonetheless, the role of BLM in 5' end resection is pro-recombinogenic since it aids in initiation of HR. This activity may serve to shuttle DSBs away from non-homologous end joining (NHEJ) pathways, which are more error prone and have been observed at higher levels in BS cells (130). NHEJ is active during the entire cell cycle, however, it cannot be initiated from the free 3' ssDNA end that results from end resection, and in this way BLM's role in end resection may be one of the factors that helps determine DSB repair pathway choice (131). Since BLM is cell cycle regulated, this would help ensure that the HR-dependent DSB repair is the primary pathway when an available sister chromatid exists to act as a template (132-134). Moreover, cells in S- and G2-phase appear to predominantly use HR to repair DSBs, and to complete this repair via dissolution to prevent exchange of genetic material (9).

In the second step of HR, the RAD51 recombinase forms a helical filament on the free 3' DNA end. A homology search for the DNA in the RAD51/ssDNA complex produces a D-loop

structure as a result of invasion of the ssDNA into a homologous sister chromatid or chromosome (135) (Figure 1.3). In this step, BLM interacts with RAD51 and is able to migrate and unwind D-loops, which disrupts nascent pairing in the first steps of HR (43, 136, 137). However, this activity appears to depend on the status of RAD51: BLM stimulates strand exchange of active ATP-bound RAD51 filaments, but dismantles inactive ADP-bound filaments (137-139). This may indicate that BLM surveys nascent D-loops to assure they are appropriate for HR. This activity could be related to a function of the prototypical *E. coli* RecQ in preventing illegitimate recombination between homologous but non-identical DNA sequences (140). Finally, some of the dismantled filaments may be repaired via single-strand annealing (SSA), which relies on flanking DNA repeat sequences commonly found in mammalian DNA (135).

Following D-loop formation and DNA synthesis, BLM has additional anti-recombinogenic rolls. First, in synthesis-dependent strand annealing (SDSA), the resulting D-loop product is dismantled and the broken chromosome is then able to re-anneal for repair in a non-crossover fashion (Figure 1.3). BLM may play a role in unwinding this D-loop (43, 136, 137). Second, in the HR-dependent pathway, the D-loop remains intact and second end capture recruits the second strand to form a dHJ substrate (Figure 1.3). This intermediate is the proposed substrate for dissolution via the BLM dissolvasome, as described above (6, 105) (Figure 1.2A). However, the dHJ substrate could also be resolved through cleavage by HJ resolvases, which leads to an equal distribution of non-crossover and crossover products (Figure 1.2B). Overall, it appears that BLM and potentially the entire BLM dissolvasome have multiple roles in HR, and while these

roles may be categorized as pro- and anti-recombinogenic, it is likely that BLM activity is focused on ensuring HR fidelity and the prevention of crossovers.

The role of the BLM dissolvasome in replication restart. Early observations with BS cell lines revealed a replication deficiency and hinted that BLM might help repair stalled or damaged replication forks. Cells from BS patients are slower to progress through S-phase and accumulate abnormal replication intermediates (141-143). Additionally, BLM levels are regulated in a cell cycle-dependent manner, accumulating in S-phase and persisting through G2/M before diminishing in G1 (132-134). Cells from BS patients are hypersensitive to replication stalling with hydroxyurea (HU) and BLM localizes to these stalled forks (82, 144).

Multiple studies point to pivotal roles for BLM in the repair of stalled DNA replication forks (145, 146). DNA replication processes can stall due to blockage of the replication fork by encounters with physical barriers, such as proteins that are tightly bound to the DNA template or impassable damage to the template. For example, replication through a nick in parental DNA creates a DSB when the replicative helicase proceeds through the nicked region. In these instances, the BLM-dependent DSB repair pathways described above would be triggered to drive replication restart (for review, see (147)). In addition, because BLM is a helicase that specifically acts upon DNA substrates that resemble replication and recombination substrates, a separate possible role for BLM is to produce DNA structures that are competent for replication restart. For example, a lesion on the leading strand that prevents progression of the replication fork could lead to replication fork regression and HJ formation (creating an intermediate referred

to as a “chicken foot” structure) whose formation could be driven by BLM branch migration (148-150). This structure can promote lesion bypass by supplying an intact template for the nascent leading strand (template switching) (145). Following lesion bypass, BLM may also act to reverse the regressed fork via branch migration so that replication can proceed past the lesion, or instead replication may restart through an HR-mediated process that would proceed through a dHJ structure that the BLM dissolvasome could act on (149).

More recently, single-molecule DNA fiber and molecular combing techniques have allowed detailed interrogation of the roles of BLM in mammalian DNA replication (151). These studies have shown that BLM is required for efficient recovery of replication forks blocked by aphidicolin or HU (152) and that BS cells exhibit reduced fork velocity and more incidences of fork pausing (153). Moreover, the BLM dissolvasome has been implicated in fork progression, as a recent study has shown that depletion of RMI1 leads to a reduction in replication fork rate and a failure to recover from replication fork arrest (154). Interestingly, these defects in replication fork rate can be suppressed when BLM is also depleted, which indicates that RMI1 and potentially the rest of the BLM dissolvasome act downstream of BLM (154). Finally, the protein RIF1 has been shown to interact with the BLM dissolvasome to promote replication fork restart, however, RIF1 does not play a role in SCE suppression (80) (Table 1.1). RIF1 can bind directly to replication fork and HJ DNA and it may coordinate fork regression with BLM (80). Therefore, the BLM dissolvasome may have roles in replication restart that are not related to its function in suppressing SCEs. In summary, a growing body of literature supports roles for the

BLM dissolvasome at replication forks to help maintain integrity through DSB repair and efficient replication restart.

Links with Fanconi Anemia protein FANCM. The discovery that the BLM dissolvasome was linked to proteins involved in the autosomal recessive disease Fanconi anemia (FA) made an important connection in genome maintenance that further solidified the importance of the BLM dissolvasome in DNA replication restart (155, 156). As is the case for BS cells, FA cell lines have significantly heightened levels of chromosomal instability, although they are exemplified by radial chromosomes and a hypersensitivity to agents that induce DNA interstrand crosslinks (ICLs). Mutations in at least fifteen genes (FANCA, B, C, D1 (BRCA2), D2, E, F, G, I, J, L, M, N, O, P) give rise to this highly heterogeneous disease, which includes symptoms of developmental abnormalities, progressive bone marrow failure, and a high occurrence of cancer (157-160). Prior to the discovery that key proteins in FA and BS directly interacted with one another, roles for the FA proteins in repair of DNA damage at ICL sites that have caused replication fork failure were known (161). In this response, a core complex of 8 FA proteins (FANCA, B, C, E, F, G, L, and M) recognizes the DNA damage and catalyzes monoubiquitination of the FANCD2/FANCI heterodimer. Ubiquitinated FANCD2/FANCI in turn recruits downstream FA proteins such as FANCI to facilitate removal of the ICL, which is followed by translesion synthesis, nucleotide excision repair, and HR to fully repair the DNA (160-162). Further details on the FA pathway can be found in a recent review by Kim and D'Andrea (160).

The initial finding that linked the BLM dissolvasome to FA was identification of an ICL-induced super-complex called BRAFT (BLM, Replication Protein A, FA, and Topo III α) that included components of both the BLM dissolvasome and FA proteins (155, 163). A subsequent study revealed that BLM and FA proteins co-localized in response to treatment with DNA crosslinkers or replication arrest and could be co-immunoprecipitated (164). Of the FA proteins, FANCM is the most conserved, with orthologs in archaeobacteria (Hef) and yeast (MPH1 in *S. cerevisiae*, and Fml1 in *Schizosaccharomyces pombe*) (165-168). FANCM is a large, 2048-residue subunit of the FA core complex that contains an N-terminal DEAH-box helicase domain and a degenerate C-terminal ERCC4-like endonuclease domain linked together by a highly dynamic linker element that has only short regions of conserved sequence (156, 165). FANCM is thought to act as a sensor to detect blocked replication forks in coordination with partner proteins FAAP24 and MHF (169-171) (Figure 1.4A). This complex is then thought to recruit other FA proteins in order to repair the DNA damage, and subsequently FANCM may act catalytically to remodel the replication fork in order for replication to proceed (158, 169, 170). FANCM-deficient cells are also characterized by a large increase in SCEs, which has also been described for other FA proteins such as FANCC and FANCI, though to a lesser degree than that observed with FANCM-depleted cells (172-175) (Table 1.1). This striking phenotypic parallel between FANCM- and BLM-deficient cells foreshadowed an interaction between FANCM and the BLM dissolvasome.

FANCM appears to have additional roles beyond recruitment of the FA core complex to ICL-damaged DNA (172, 176, 177). FANCM might also aid in general coordination of replication

restart, and only initiate the FA pathway under specific circumstances such as ICLs. While FANCM does not have demonstrated helicase activity, as a translocase FANCM can act on DNA substrates that mimic replication forks and HJs, and has been implicated in replication fork remodeling (165, 176, 178, 179). Molecular combing experiments have shown that the translocase activity of FANCM is important for the restart of stalled replication forks (180, 181). Moreover, these stalled replication forks have been shown to be prone to collapse, which generates DSBs that are subsequently repaired through HR (182). At these sites, the coordinated activity of the BLM dissolvasome and FANCM may be most evident.

A direct interaction between a conserved 34-residue motif in the middle of FANCM (called MM2, for FANCM Motif 2) and the C-terminal OB-fold of RMI1 was first demonstrated in 2009 (156) (Figure 1.4). A recent crystal structure of a proteolytically stable RMI1/RMI2 fragment bound to the MM2 peptide has revealed that the full interface is comprised of both the C-terminal OB-fold of RMI1 and the full-length RMI2 protein (183) (Figure 1.4B). In this complex, the side chains of conserved hydrophobic residues from FANCM MM2 dock into hydrophobic pockets on the RMI1/RMI2 surface, forming a stable tripartite complex (183). A FANCM mutation in two of the docking residues from MM2 (FANCM FF>AA) leads to an increase in SCEs upon treatment with mitomycin C (MMC), which induces ICLs (156) (Table 1.1). Moreover, mutation of a key residue on the RMI surface (Lys121 of RMI2) leads to an increase of SCEs in chicken DT40 cells that is similar to levels seen when FANCM or members of the BLM dissolvasome are depleted (73, 115, 121, 169, 172, 183) (Table 1.1 and Figure 1.4B). These data indicate that coordination between FANCM and the BLM dissolvasome is

critical in preventing crossover formation. Beyond their roles in SCE suppression, the BLM dissolvasome, FANCM, and FANCI (see below) all appear to be important in replication fork progression, remodeling, and integrity. Further research will be required to delineate their specific roles and interdependence.

BLM/FANCI interaction. In addition to the RMI/FANCM interaction, interactions between other FA- and BS-associated proteins have been reported. Notably, a functional and physical interaction has been described between BLM and FANCI, a helicase that acts downstream of the FA core complex (79) (Figure 1.4A). As with BS- and FANCM-deficient cells, SCEs are elevated in FANCI-deficient cells (although only ~2-3 fold for FANCI, see Table 1.1), and this elevation is synergistic with BLM (79, 175). FANCI is an ATP-dependent SF2 helicase that acts in the 5' to 3' direction and is also sensitive to replication stress, and therefore it may cooperate with BLM replication restart (79, 184) (for recent discussion on this interaction see (185)). In addition, both BLM and FANCI are able to unwind G-quartet DNA, and both may be important in replication to remove these secondary structure elements from DNA so that replication can proceed (39, 186, 187). A recent study in DT40 cells implied a genetic interaction between FANCI, BLM, and WRN that may allow coordination of replication through G-quartet DNA (188).

BLM and FA proteins both act at the replication-dependent checkpoint. In parallel with their activities in replication fork stability and restart, roles for BLM and multiple FA proteins have been established in the replication-dependent or S-phase cell cycle checkpoint. This

checkpoint triggers cell cycle arrest when replicative processes are blocked due to DNA damage, which leads to the activation of ATM (ataxia-telangiectasia mutated) and ATR (ATM and Rad3-related) kinases (189). ATM and ATR activation initiates a signaling cascade that stalls the cell cycle, providing the time needed to recover from replicative stress. While the two pathways are interdependent, ATR generally signals replication-dependent damage through CHK1 (checkpoint kinase 1) whereas ATM signals in response to DSBs via CHK2 (checkpoint kinase 2) (190, 191).

Within this checkpoint mechanism, BLM, FANCD2, and FANCI are targets of both the ATR/CHK1 and the ATM/CHK2 pathways within the pathway whereas FANCA and FANCE appear to be acted upon preferentially by the ATR/CHK1 pathway (81, 82, 192-199). BLM is phosphorylated at T99 and T122, and BS cells expressing an unphosphorylatable BLM variant are unable to recover following treatment with HU, and instead enter an extended arrest at the G2/M checkpoint (81, 82). Furthermore, this recovery is dependent on the helicase activity of BLM (152). Taken together, these data indicate a critical role for BLM at the replication-dependent checkpoint, and suggests that after replication has stalled, BLM is phosphorylated, which signals its localization to replication forks in order to aid in faithful restart. BLM may act without the rest of the BLM dissolvasome in this role, as phosphorylation on T99 by ATM or ATR causes BLM to dissociate from Topo III α , and presumably the RMI subcomplex as well (192).

FANCM and its binding partner FAAP24 have been implicated as upstream effectors in activation of the ATR pathway, which is independent from the rest of the FA core complex (179-

181). However, a more recent study has suggested that FANCM and FAAP24 have some independent functions, and that FAAP24 alone is responsible for checkpoint activation (200). In a similar fashion, FANCI appears to play a role in ATR activation (201). These insights have led to a model in which FANCM/FAAP24 (or FAAP24 alone) and FANCI act as sensors of replication stress in cells to initiate activation of ATR/CHK1 and the replication-dependent checkpoint. The activation of this checkpoint then leads to phosphorylation of other FA proteins and BLM to aid in DNA repair and replication restart (158, 181). Further research will be required to tease apart the roles of each at different stages and how they vary depending on the type of genetic insult.

The BLM dissolvasome and FA proteins are both found associated with ultra-fine anaphase bridges. DNA staining of sister chromatids as they separate during mitosis has revealed the presence of anaphase DNA bridge structures that link the two chromatids. More recently, a subclass of anaphase bridges, termed “ultra-fine anaphase bridges” (UFBs) have been discovered (202-204). These UFBs are more thread-like and cannot be visualized with conventional DNA dyes; instead they are localized with immuno-staining of proteins found on the UFBs. Several proteins, including members of the BLM dissolvasome and the PLK1-interacting checkpoint helicase (PICH) are localized to UFBs, and BLM may be specifically recruited by its interaction with PICH (75, 202, 203). Localization of Topo III α and RMI1 to UFBs is dependent on BLM, and therefore it appears that BLM recruits the rest of the BLM dissolvasome to UFBs (203). The majority of UFBs arise at centromeres, where it has been proposed that the BLM dissolvasome and PICH might cooperate to facilitate the association of

Topo II α , a type II topoisomerase, to drive centromere disjunction (75, 203, 205-207). To further support this hypothesis, BLM and Topo II α have been shown to interact both *in vitro* and *in vivo* (74, 208). The majority of these bridges appear to resolve as mitosis progresses without the formation of DSB intermediates, and may be normal structures that assist in centromeric cohesion (204).

A less common class of UFBs that are exclusively found at common fragile sites on chromosomes are proposed to result from incomplete DNA replication that is induced by replication stalling agents (209, 210). Normally under replication stress, additional replication initiation events will occur so that the DNA can be fully duplicated, but recent studies indicate that at fragile sites there is a lack of these extra initiation events, and so mitosis may begin before these regions have completed replication, resulting in a UFB (211, 212). Apart from their spatial differences, the other distinguishing feature of this class of UFBs is the localization of FA proteins FANCD2 and FANCI at their termini, which is consistent with their presence being related to replication stress (209, 210). FANCD2/I are proposed to be recruited to sites of incomplete replication in the late S- and G2-phases, which then become sites of UFBs that are coated by BLM and PICH in mitosis. FANCM has also been localized to UFBs in a BLM-dependent manner, though FANCM is observed on UFBs later in mitosis after BLM has already dissociated from UFBs (213). It has therefore been hypothesized that the BLM dissolvasome can recruit FANCM to UFBs in a hand-off mechanism, however, another possibility is that BLM alters the structure of the DNA in the bridge to make it accessible for FANCM later in mitosis (204, 213).

Most of these FA-dependent bridges disappear as mitosis progresses, although some persist and presumably the cell divides before replication is completed at these sites (209, 210). One hypothesized fate of UFBs is to resolve before replication is complete and form symmetrical DNA lesions in daughter cells that are marked as sites of DNA damage with 53BP1 (214). These sites become sequestered and protected by 53BP1 nuclear bodies until they can be repaired properly in the next S-phase (204, 214, 215). Alternatively, a more detrimental result of unresolved UFBs is chromosomal breakage, which leads to the formation of micronuclei. These products are observed at higher frequency when BLM or FA proteins are depleted in cells (10, 203, 209, 210, 213).

Concluding remarks. From the initial studies of the relatively simple *E. coli* RecQ protein to the more complex BLM dissolvasome and its many interactions, it has become clear that RecQ proteins and protein complexes have evolved central roles in maintaining genomic stability in all organisms. In all cases, these proteins do not act alone, but instead function as components in direct and indirect networks that integrate diverse enzymatic activities and cellular regulatory systems. In this review, we have focused on BLM and its interactions with Topo III α , RMI1, and RMI2, which comprise the BLM dissolvasome in humans. Overall, BLM appears to coordinate its functions with these and other proteins in order to maintain genetic stability in somatic cells. These functions include roles in replication restart, HR-dependent DSB repair, the replication-dependent checkpoint, and at UFBs. In many of these pathways, a role for BLM has only recently been described, and future studies will be required to understand the precise role for BLM and/or other human RecQ helicases *in vivo*.

We have also explored the role of the interaction between the BLM dissolvasome and FA proteins, specifically FANCM and FANCI. It appears that BLM and FA proteins work together in many pathways, and we have described one potential model for the localization of these proteins at ICLs (Figure 1.4A). The BLM dissolvasome and FA proteins coordinate their roles in dHJ dissolution, both respond to stalled replication forks, are involved in the replication-dependent checkpoint, and are found at UFBs. Another interesting collaboration may also exist in DSB repair, as FA proteins have been speculated to be involved in repair pathway choice, especially if the damage is an ICL. Studies in *C. elegans*, chicken DT40, and mammalian cells have provided evidence that the FA pathway helps to suppress NHEJ and promote repair via HR (216, 217). A recent study showing that DNA2 and FANCD2 can be copurified provides further evidence that the FA pathway is involved in end resection (218). However, the network of direct protein-protein interactions required to support this function remains to be defined.

With the stark effects of mutations in three of the five human RecQ genes, the central nature for RecQ proteins is well appreciated. Future research will continue to define the biochemical and cellular roles for RecQ proteins, leading to a better understanding of the molecular basis by which these critical enzymes function and the cellular consequences that result from their dysfunction. These studies may also pave the way for the development of new chemotherapeutics and allow for better treatment of patients with BS and FA. As an example, one recent study identified a small molecule that competitively inhibits BLM and induces an increase in SCEs (219). This compound will be useful to more specifically target BLM inactivation in further studies and as a potential chemotherapeutic. Other compounds may

similarly be discovered that inhibit specific proteins or interactions and could be used to sensitize cancer cells to DNA damaging agents in cancer treatment.

Acknowledgements

We apologize to all authors whose work we could not cite due to space limitations. Work in our lab was funded by a grant from the National Institutes of Health (GM068061) and KAM was supported in part by a National Institutes of Health training grant in Molecular Biosciences (GM07215).

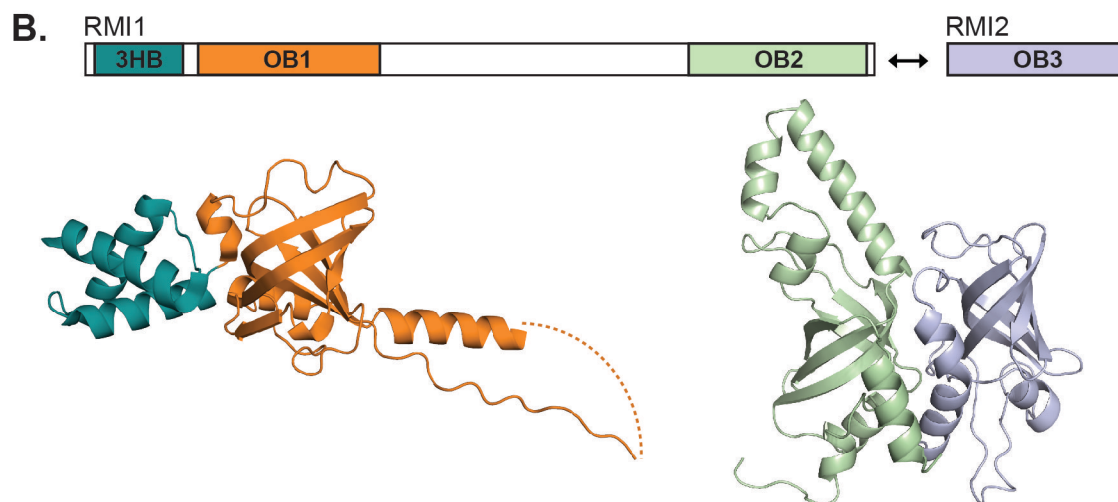
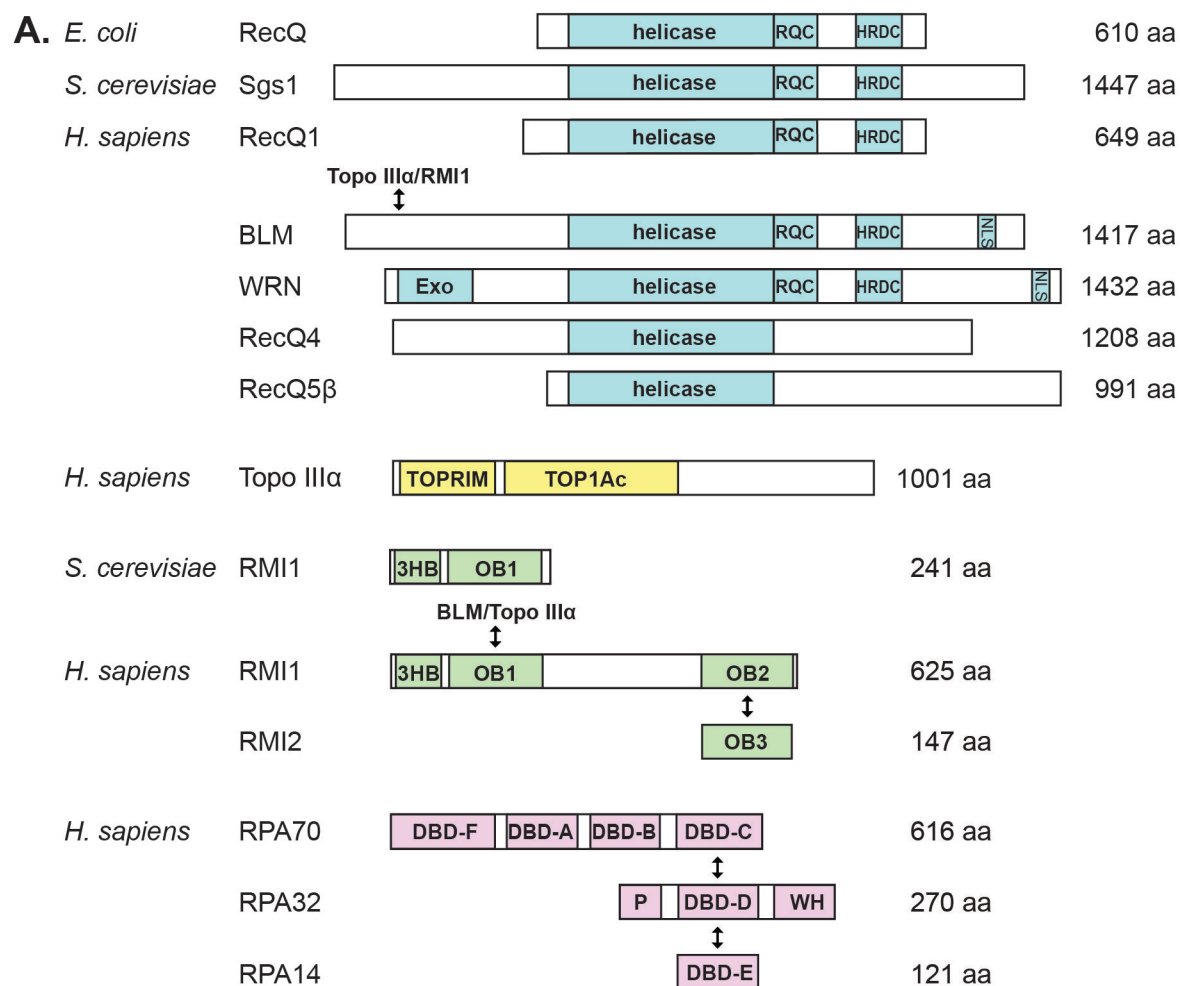
Figure 1.1.

Figure 1.1. Properties of the RecQ helicases and related proteins. (A) Domain architectures of RecQs, Topo III α , RMI1, RMI2 and RPA. Known interactions are depicted with arrows. RPA is shown because it shares sequence and structural similarities with the RMI subcomplex: RMI1 OB1 resembles RPA 70 DBD-F, RMI1 OB2 resembles RPA70 DBD-C, and RMI2 OB3 resembles RPA32 DBD-D. Abbreviations used: RQC = RecQ-conserved, HRDC = helicase and RNaseD C-terminal, NLS = nuclear localization signal, Exo = exonuclease, TOPRIM = topoisomerase-primase, TOP1Ac = topoisomerase IA conserved, 3HB = three helix bundle (also DUF1767), OB = OB-fold, DBD = DNA binding domain (also OB-folds), P = phosphorylated region, WH = winged helix, aa = amino acids. (B) Structures of RMI1 and RMI2. The structures are indicated by the colors in the domains above. On the left is a ribbon diagram of the crystal structure of the N-terminus of RMI1 (residues 2-204, excluding residues 109-119, and this region is connected with a dashed line) (120). On the right is a ribbon diagram of the crystal structure RMI1 OB2 (residues 475-625), and RMI2 (residues 17-147) (120, 121). The diagrams were rendered using PyMol (220).

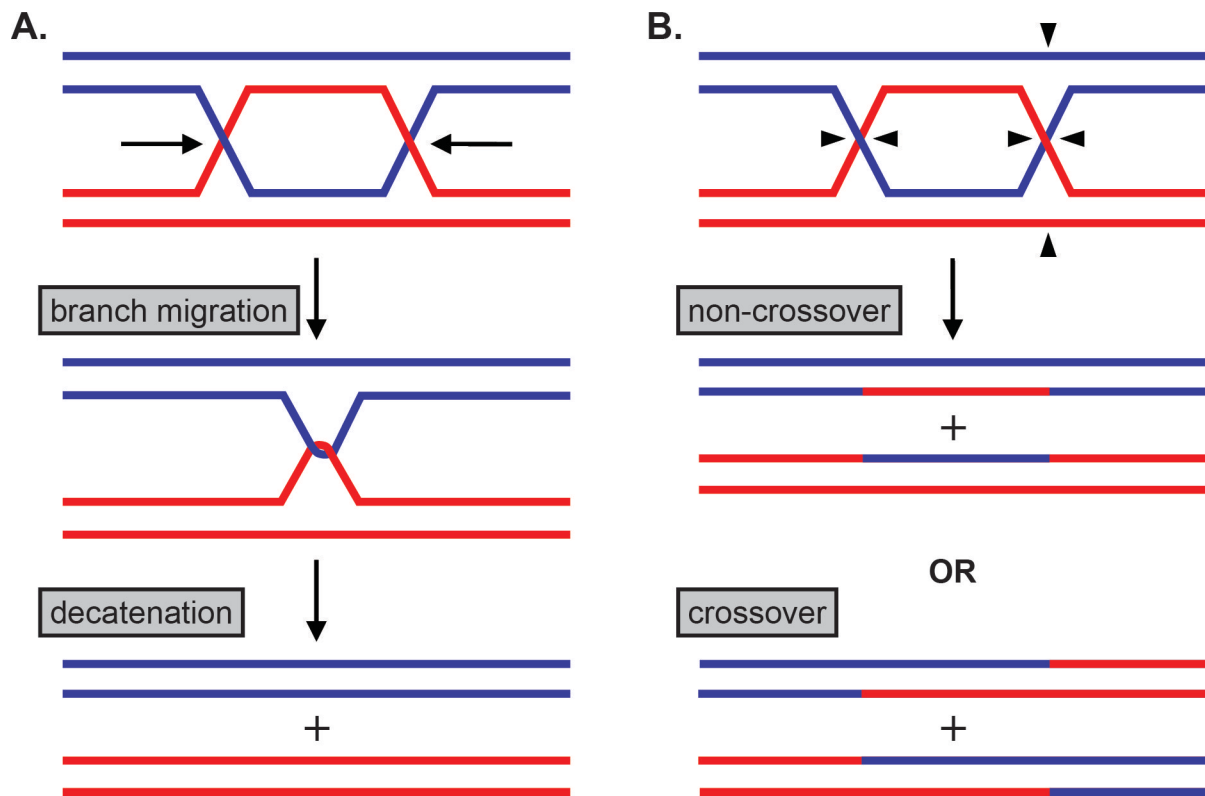
Figure 1.2.

Figure 1.2. Pathways for the resolution of dHJ structures. (A) Proposed pathway for dHJ dissolution. In this pathway BLM migrates the two HJs towards each other, then Topo III α decatenates the structure in the second step, producing exclusively non-crossover products. (B) Pathway for HJ resolvases. In this pathway each HJ is cleaved by an HJ resolvase and an equal mixture of non-crossover and crossover products result, depending on if the HJs are cleaved symmetrically or asymmetrically, as depicted by the arrows.

Figure 1.3.

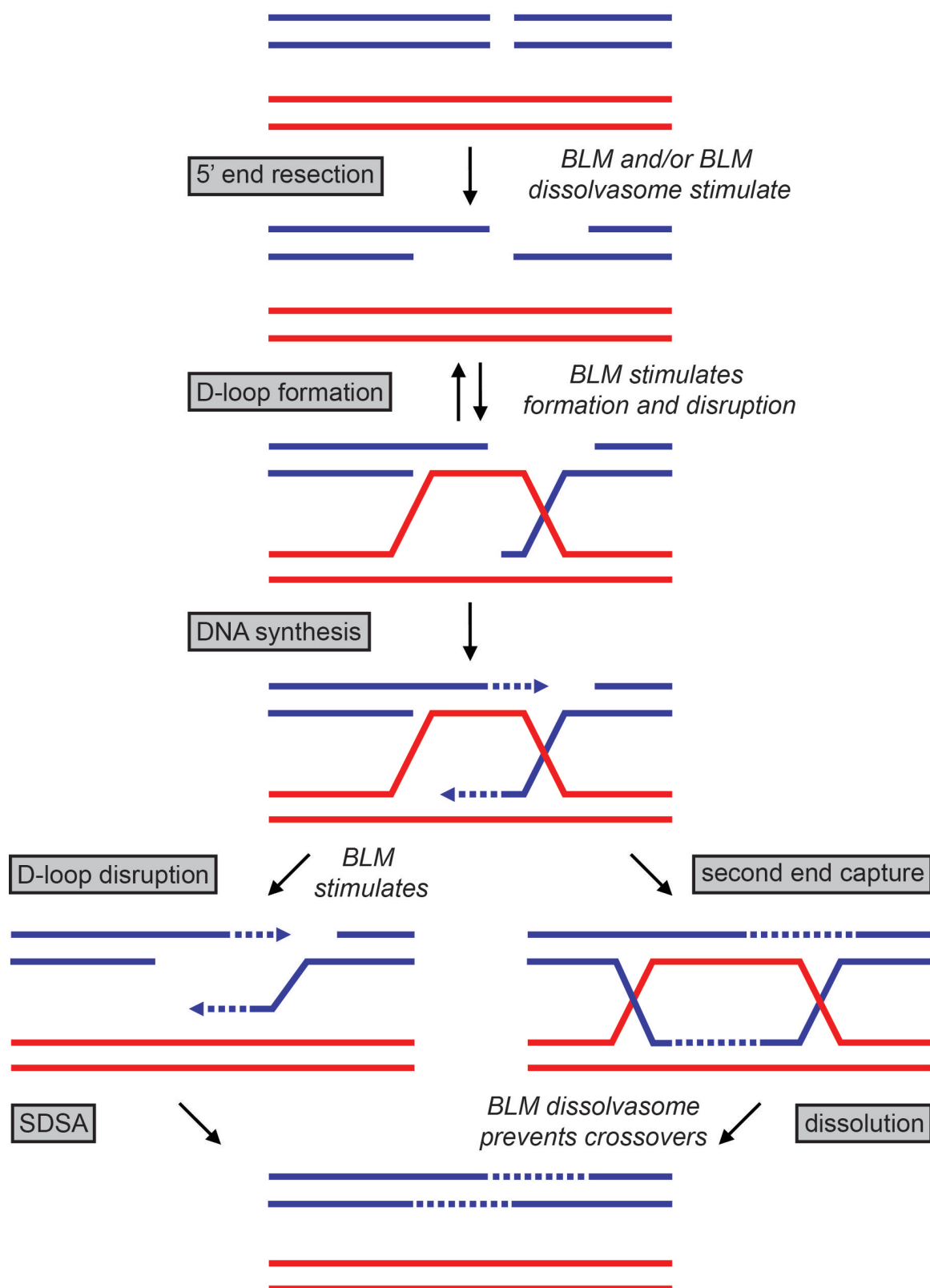


Figure 1.3. Roles for the BLM dissolvasome in HR-dependent DSB repair. Shown here is the pathway for DSB repair, showing the different steps where BLM and/or the BLM dissolvasome may act or stimulate. The pathways shown here are resolved to non-crossover products, either via synthesis-dependent strand annealing (SDSA) (left) or dHJ dissolution (right), however, the dHJ substrate may also be resolved via HJ resolvases as shown in Figure 1.2. New DNA synthesis is depicted with a dashed line.

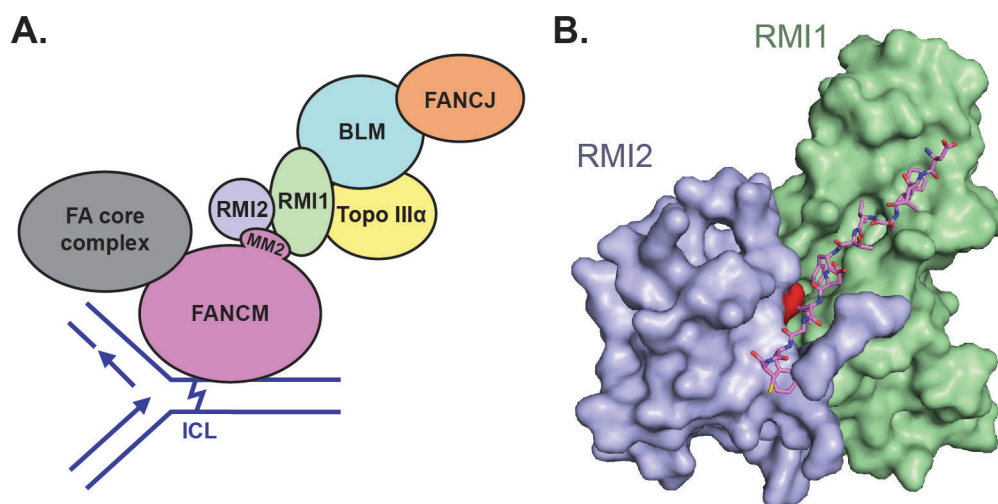
Figure 1.4.

Figure 1.4. BLM dissolvasome and FA protein interactions described in this review. (A) A potential model for the protein interactions between the BLM dissolvasome, FANCM, and FANCI. FANCM is shown at a site of DNA damage, depicted as an ICL, that may be encountered during replication and impede fork progression. In this model, FANCM then recruits the BLM dissolvasome through the MM2 interaction with the RMI subcomplex. FANCM is also part of the FA core complex and can localize the rest of the core complex to an ICL. FANCI may also be recruited to sites of DNA damage via its interaction with BLM. **(B)** The overall structure of the RMI core complex with bound MM2 peptide from FANCM. MM2 residues 1226 – 1237 are shown in purple, and RMI2 Lys121 is highlighted in red (183).

Table 1.1. A summary of studies examining cellular SCE

Protein disrupted or mutated	WT SCE	Mean SCE	Fold increase	Cell line	Significance
BLM (11)	6.9	89.0	12.9	BS patient cells	initial SCE report in BS cells
Topo III α (221)	-	-	-	Mouse embryonic	embryonic lethal
Topo III α (222)	-	-	-	DT40	knockout is lethal
BLM (223)	2.1	26.4	12.6	DT40	
RAD54	2.1	1.7	0.8	DT40	RAD54 is a DSB repair protein
BLM/RAD54	2.1	8.2	3.9	DT40	many SCEs come from DSB repair
BLM (73)	~0.2 ^a	~0.5 ^a	2-3	HeLa siRNA ^b	
RMI1	~0.2 ^a	~0.5 ^a	2-3	HeLa siRNA ^b	BLM and RMI1 -- same pathway
BLM (115)	2.1	19.0	9.0	DT40	
RMI2	2.1	11.4/9.5	5.4/4.5	DT40	two different experiments
BLM/RMI2	2.1	18.1	8.6	DT40	BLM and RMI2 -- same pathway
RMI2 Lys121Ala	2.1	8.2	3.9	DT40	disrupts RMI/FANCM interface
RMI2 (121)	1.85	11.3	6.1	DT40	
Lys121Ala	1.85	10.2	5.5	DT40	disrupts RMI/FANCM interface
Lys121Glu	1.85	10.9	5.9	DT40	also weakens RMI1/RMI2 (in Co-IPs)
Asp141Lys	1.85	15.3	8.3	DT40	disrupts RMI1/RMI2 interface
Lys121Glu/Asp141Lys	1.85	11.6	6.3	DT40	disrupts both interfaces
FANCC (173)	1.4	5.1	3.6	DT40	
FANCC (175)	1.8	4.1	2.3	DT40	
FANCCJ	1.8	5.2	2.9	DT40	
FANCM (172)	1.8	9.0	5.0	DT40	SCEs higher than other FA disruptions
FANCC	1.8	6.0	3.3	DT40	
FANCM/FANCC	1.8	14.9	8.3	DT40	SCEs are additive -- two pathways
FANCM D203A	1.8	8.9	4.9	DT40	FANCM catalytic activity required
BLM	1.8	14.0	7.8	DT40	
BLM/FANCM D203A	1.8	15.2	8.4	DT40	BLM and FANCM -- same pathway
FANCCJ (79)	0.31	0.59	1.9	EUFA30-F ^c	also no SCE increase in FA-C cells
FANCCJ	0.21	0.36	1.7	PSNF5 ^c /FANCCJ siRNA	
BLM	0.21	1.38	6.6	PSNG13 ^c	"WT" data is from above PSNF5 cells
BLM/FANCCJ	0.21	1.21	5.8	PSNG13 ^c /FANCCJ siRNA	BLM and FANCCJ -- same pathway
FANCM (156)	~6 ^a	~12.5 ^a	2-3	293 siRNA ^b with MMC	
FANCM FF>AA	~6 ^a	~12.5 ^a	2-3	293 siRNA ^b with MMC	disrupts RMI/FANCM interface (183)
FANCM (183)	2.0	17.7	8.9	DT40	
FANCM F1232A	2.0	14.2	7.1	DT40	disrupts RMI/FANCM interface (183)
RIF1 (80)	1.7	1.7	1.0	DT40	not involved in SCE suppression

^a mean SCE wasn't reported, inferred from data^b experiments using siRNA are known to not completely mimic the SCE levels observed in BS patients^c EUFA30-F = FA-J fibroblasts, PSNF5 = BLM^{+/+}, PSNG13 = BLM^{-/-}

References

1. Loeb LA (2001) A mutator phenotype in cancer. *Cancer Res* 61(8):3230-3239.
2. Ouyang KJ, Woo LL, & Ellis NA (2008) Homologous recombination and maintenance of genome integrity: cancer and aging through the prism of human RecQ helicases. *Mech Ageing Dev* 129(7-8):425-440.
3. Lengauer C, Kinzler KW, & Vogelstein B (1998) Genetic instabilities in human cancers. *Nature* 396(6712):643-649.
4. Bachrati CZ & Hickson ID (2003) RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem J* 374(Pt 3):577-606.
5. Bachrati CZ & Hickson ID (2008) RecQ helicases: guardian angels of the DNA replication fork. *Chromosoma* 117(3):219-233.
6. Mankouri HW & Hickson ID (2007) The RecQ helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific 'dissolvasome'? *Trends Biochem Sci* 32(12):538-546.
7. Bloom D (1954) Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs; probably a syndrome entity. *AMA Am J Dis Child* 88(6):754-758.
8. German J (1993) Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine (Baltimore)* 72(6):393-406.
9. Chu WK & Hickson ID (2009) RecQ helicases: multifunctional genome caretakers. *Nat Rev Cancer* 9(9):644-654.
10. Rosin MP & German J (1985) Evidence for chromosome instability *in vivo* in Bloom syndrome: increased numbers of micronuclei in exfoliated cells. *Hum Genet* 71(3):187-191.
11. Chaganti RS, Schonberg S, & German J (1974) A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc Natl Acad Sci U S A* 71(11):4508-4512.
12. McDaniel LD & Schultz RA (1992) Elevated sister chromatid exchange phenotype of Bloom syndrome cells is complemented by human chromosome 15. *Proc Natl Acad Sci U S A* 89(17):7968-7972.
13. Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, Ciocchi S, Proytcheva M, & German J (1995) The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 83(4):655-666.

14. German J, Sanz MM, Ciocci S, Ye TZ, & Ellis NA (2007) Syndrome-causing mutations of the BLM gene in persons in the Bloom's Syndrome Registry. *Hum Mutat* 28(8):743-753.
15. Foucault F, Vaury C, Barakat A, Thibout D, Planchon P, Jaulin C, Praz F, & Amor-Gueret M (1997) Characterization of a new BLM mutation associated with a topoisomerase II alpha defect in a patient with Bloom's syndrome. *Hum Mol Genet* 6(9):1427-1434.
16. Barakat A, Ababou M, Onclercq R, Dutertre S, Chadli E, Hda N, Benslimane A, & Amor-Gueret M (2000) Identification of a novel BLM missense mutation (2706T>C) in a Moroccan patient with Bloom's syndrome. *Hum Mutat* 15(6):584-585.
17. Guo RB, Rigolet P, Ren H, Zhang B, Zhang XD, Dou SX, Wang PY, Amor-Gueret M, & Xi XG (2007) Structural and functional analyses of disease-causing missense mutations in Bloom syndrome protein. *Nucleic Acids Res* 35(18):6297-6310.
18. Neff NF, Ellis NA, Ye TZ, Noonan J, Huang K, Sanz M, & Proytcheva M (1999) The DNA helicase activity of BLM is necessary for the correction of the genomic instability of Bloom syndrome cells. *Mol Biol Cell* 10(3):665-676.
19. Bahr A, De Graeve F, Kedinger C, & Chatton B (1998) Point mutations causing Bloom's syndrome abolish ATPase and DNA helicase activities of the BLM protein. *Oncogene* 17(20):2565-2571.
20. Bernstein DA, Zittel MC, & Keck JL (2003) High-resolution structure of the *E.coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910-4921.
21. Nakayama H, Nakayama K, Nakayama R, Irino N, Nakayama Y, & Hanawalt PC (1984) Isolation and genetic characterization of a thymineless death-resistant mutant of *Escherichia coli* K12: identification of a new mutation (recQ1) that blocks the RecF recombination pathway. *Molecular & general genetics : MGG* 195(3):474-480.
22. Nakayama K, Irino N, & Nakayama H (1985) The recQ gene of *Escherichia coli* K12: molecular cloning and isolation of insertion mutants. *Molecular & general genetics : MGG* 200(2):266-271.
23. Yu CE, Oshima J, Fu YH, Wijsman EM, Hisama F, Alisch R, Matthews S, Nakura J, Miki T, Ouais S, Martin GM, Mulligan J, & Schellenberg GD (1996) Positional cloning of the Werner's syndrome gene. *Science* 272(5259):258-262.
24. Kitao S, Shimamoto A, Goto M, Miller RW, Smithson WA, Lindor NM, & Furuichi Y (1999) Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. *Nat Genet* 22(1):82-84.

25. Siitonen HA, Kopra O, Kaariainen H, Haravuori H, Winter RM, Saamanen AM, Peltonen L, & Kestila M (2003) Molecular defect of RAPADILINO syndrome expands the phenotype spectrum of RECQL diseases. *Hum Mol Genet* 12(21):2837-2844.
26. Van Maldergem L, Siitonen HA, Jalkh N, Chouery E, De Roy M, Delague V, Muenke M, Jabs EW, Cai J, Wang LL, Plon SE, Fourneau C, Kestila M, Gillerot Y, Megarbane A, & Verloes A (2006) Revisiting the craniosynostosis-radial ray hypoplasia association: Baller-Gerold syndrome caused by mutations in the RECQL4 gene. *J Med Genet* 43(2):148-152.
27. Killoran MP & Keck JL (2006) Sit down, relax and unwind: structural insights into RecQ helicase mechanisms. *Nucleic Acids Res* 34(15):4098-4105.
28. Morozov V, Mushegian AR, Koonin EV, & Bork P (1997) A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem Sci* 22(11):417-418.
29. Pike AC, Shrestha B, Popuri V, Burgess-Brown N, Muzzolini L, Costantini S, Vindigni A, & Gileadi O (2009) Structure of the human RECQ1 helicase reveals a putative strand-separation pin. *Proc Natl Acad Sci U S A* 106(4):1039-1044.
30. Kitano K, Kim SY, & Hakoshima T (2010) Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN. *Structure* 18(2):177-187.
31. von Kobbe C, Thoma NH, Czyzewski BK, Pavletich NP, & Bohr VA (2003) Werner syndrome protein contains three structure-specific DNA binding domains. *J Biol Chem* 278(52):52997-53006.
32. Huber MD, Duquette ML, Shiels JC, & Maizels N (2006) A conserved G4 DNA binding domain in RecQ family helicases. *J Mol Biol* 358(4):1071-1080.
33. Vindigni A & Hickson ID (2009) RecQ helicases: multiple structures for multiple functions? *HFSP J* 3(3):153-164.
34. Bernstein DA & Keck JL (2005) Conferring substrate specificity to DNA helicases: role of the RecQ HRDC domain. *Structure* 13(8):1173-1182.
35. Kitano K, Yoshihara N, & Hakoshima T (2007) Crystal structure of the HRDC domain of human Werner syndrome protein, WRN. *J Biol Chem* 282(4):2717-2728.
36. Sato A, Mishima M, Nagai A, Kim SY, Ito Y, Hakoshima T, Jee JG, & Kitano K (2010) Solution structure of the HRDC domain of human Bloom syndrome protein BLM. *Journal of biochemistry* 148(4):517-525.

37. Wu L, Chan KL, Ralf C, Bernstein DA, Garcia PL, Bohr VA, Vindigni A, Janscak P, Keck JL, & Hickson ID (2005) The HRDC domain of BLM is required for the dissolution of double Holliday junctions. *Embo J* 24(14):2679-2687.
38. Karow JK, Chakraverty RK, & Hickson ID (1997) The Bloom's syndrome gene product is a 3'-5' DNA helicase. *J Biol Chem* 272(49):30611-30614.
39. Sun H, Karow JK, Hickson ID, & Maizels N (1998) The Bloom's syndrome helicase unwinds G4 DNA. *J Biol Chem* 273(42):27587-27592.
40. Huber MD, Lee DC, & Maizels N (2002) G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucleic Acids Res* 30(18):3954-3961.
41. Mohaghegh P, Karow JK, Brosh RM, Jr., Bohr VA, & Hickson ID (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res* 29(13):2843-2849.
42. Popuri V, Bachrati CZ, Muzzolini L, Mosedale G, Costantini S, Giacomini E, Hickson ID, & Vindigni A (2008) The Human RecQ helicases, BLM and RECQ1, display distinct DNA substrate specificities. *J Biol Chem* 283(26):17766-17776.
43. van Brabant AJ, Ye T, Sanz M, German IJ, Ellis NA, & Holloman WK (2000) Binding and melting of D-loops by the Bloom syndrome helicase. *Biochemistry* 39(47):14617-14625.
44. Karow JK, Newman RH, Freemont PS, & Hickson ID (1999) Oligomeric ring structure of the Bloom's syndrome helicase. *Curr Biol* 9(11):597-600.
45. Beresten SF, Stan R, van Brabant AJ, Ye T, Naureckiene S, & Ellis NA (1999) Purification of overexpressed hexahistidine-tagged BLM N431 as oligomeric complexes. *Protein Expr Purif* 17(2):239-248.
46. Xu YN, Bazeille N, Ding XY, Lu XM, Wang PY, Bugnard E, Grondin V, Dou SX, & Xi XG (2012) Multimeric BLM is dissociated upon ATP hydrolysis and functions as monomers in resolving DNA structures. *Nucleic Acids Res* 40(19):9802-9814.
47. Muzzolini L, Beuron F, Patwardhan A, Popuri V, Cui S, Niccolini B, Rappas M, Freemont PS, & Vindigni A (2007) Different quaternary structures of human RECQ1 are associated with its dual enzymatic activity. *PLoS biology* 5(2):e20.
48. Xue Y, Ratcliff GC, Wang H, Davis-Searles PR, Gray MD, Erie DA, & Redinbo MR (2002) A minimal exonuclease domain of WRN forms a hexamer on DNA and possesses both 3'- 5' exonuclease and 5'-protruding strand endonuclease activities. *Biochemistry* 41(9):2901-2912.

49. Compton SA, Tolun G, Kamath-Loeb AS, Loeb LA, & Griffith JD (2008) The Werner syndrome protein binds replication fork and holliday junction DNAs as an oligomer. *J Biol Chem* 283(36):24478-24483.
50. Huang S, Beresten S, Li B, Oshima J, Ellis NA, & Campisi J (2000) Characterization of the human and mouse WRN 3'-->5' exonuclease. *Nucleic Acids Res* 28(12):2396-2405.
51. Janscak P, Garcia PL, Hamburger F, Makuta Y, Shiraishi K, Imai Y, Ikeda H, & Bickle TA (2003) Characterization and mutational analysis of the RecQ core of the bloom syndrome protein. *J Mol Biol* 330(1):29-42.
52. Brosh RM, Jr., Li JL, Kenny MK, Karow JK, Cooper MP, Kureekattil RP, Hickson ID, & Bohr VA (2000) Replication protein A physically interacts with the Bloom's syndrome protein and stimulates its helicase activity. *J Biol Chem* 275(31):23500-23508.
53. Cheok CF, Wu L, Garcia PL, Janscak P, & Hickson ID (2005) The Bloom's syndrome helicase promotes the annealing of complementary single-stranded DNA. *Nucleic Acids Res* 33(12):3932-3941.
54. Umezu K & Nakayama H (1993) RecQ DNA helicase of *Escherichia coli*. Characterization of the helix-unwinding activity with emphasis on the effect of single-stranded DNA-binding protein. *J Mol Biol* 230(4):1145-1150.
55. Harmon FG & Kowalczykowski SC (1998) RecQ helicase, in concert with RecA and SSB proteins, initiates and disrupts DNA recombination. *Genes Dev* 12(8):1134-1144.
56. Harmon FG & Kowalczykowski SC (2001) Biochemical characterization of the DNA helicase activity of the *Escherichia coli* RecQ helicase. *J Biol Chem* 276(1):232-243.
57. Shereda RD, Bernstein DA, & Keck JL (2007) A central role for SSB in *Escherichia coli* RecQ DNA helicase function. *J Biol Chem* 282(26):19247-19258.
58. Lecointe F, Serena C, Velten M, Costes A, McGovern S, Meile JC, Errington J, Ehrlich SD, Noirot P, & Polard P (2007) Anticipating chromosomal replication fork arrest: SSB targets repair DNA helicases to active forks. *Embo J* 26(19):4239-4251.
59. Shereda RD, Reiter NJ, Butcher SE, & Keck JL (2009) Identification of the SSB binding site on *E. coli* RecQ reveals a conserved surface for binding SSB's C terminus. *J Mol Biol* 386(3):612-625.
60. Cejka P & Kowalczykowski SC (2010) The full-length *Saccharomyces cerevisiae* Sgs1 protein is a vigorous DNA helicase that preferentially unwinds Holliday junctions. *J Biol Chem* 285(11):8290-8301.

61. Cobb JA, Bjergbaek L, Shimada K, Frei C, & Gasser SM (2003) DNA polymerase stabilization at stalled replication forks requires Mec1 and the RecQ helicase Sgs1. *Embo J* 22(16):4325-4336.
62. Hegnauer AM, Hustedt N, Shimada K, Pike BL, Vogel M, Amsler P, Rubin SM, van Leeuwen F, Guenole A, van Attikum H, Thoma NH, & Gasser SM (2012) An N-terminal acidic region of Sgs1 interacts with Rpa70 and recruits Rad53 kinase to stalled forks. *Embo J* 31(18):3768-3783.
63. Garcia PL, Bradley G, Hayes CJ, Krintel S, Soultanas P, & Janscak P (2004) RPA alleviates the inhibitory effect of vinylphosphonate internucleotide linkages on DNA unwinding by BLM and WRN helicases. *Nucleic Acids Res* 32(12):3771-3778.
64. Doherty KM, Sommers JA, Gray MD, Lee JW, von Kobbe C, Thoma NH, Kureekattil RP, Kenny MK, & Brosh RM, Jr. (2005) Physical and functional mapping of the replication protein A interaction domain of the Werner and Bloom syndrome helicases. *J Biol Chem* 280(33):29494-29505.
65. Ahn B, Lee JW, Jung H, Beck G, & Bohr VA (2009) Mechanism of Werner DNA helicase: POT1 and RPA stimulates WRN to unwind beyond gaps in the translocating strand. *PLoS One* 4(3):e4673.
66. Sowd G, Wang H, Pretto D, Chazin WJ, & Opresko PL (2009) Replication protein A stimulates the Werner syndrome protein branch migration activity. *J Biol Chem* 284(50):34682-34691.
67. Machwe A, Lozada E, Wold MS, Li GM, & Orren DK (2011) Molecular cooperation between the Werner syndrome protein and replication protein A in relation to replication fork blockage. *J Biol Chem* 286(5):3497-3508.
68. Hyun M, Park S, Kim E, Kim DH, Lee SJ, Koo HS, Seo YS, & Ahn B (2012) Physical and functional interactions of *Caenorhabditis elegans* WRN-1 helicase with RPA-1. *Biochemistry* 51(7):1336-1345.
69. Cui S, Arosio D, Doherty KM, Brosh RM, Jr., Falaschi A, & Vindigni A (2004) Analysis of the unwinding activity of the dimeric RECQ1 helicase in the presence of human replication protein A. *Nucleic Acids Res* 32(7):2158-2170.
70. Cui S, Klima R, Ochem A, Arosio D, Falaschi A, & Vindigni A (2003) Characterization of the DNA-unwinding activity of human RECQ1, a helicase specifically stimulated by human replication protein A. *J Biol Chem* 278(3):1424-1432.
71. Garcia PL, Liu Y, Jiricny J, West SC, & Janscak P (2004) Human RECQ5beta, a protein with DNA helicase and strand-annealing activities in a single polypeptide. *Embo J* 23(14):2882-2891.

72. Wu L, Davies SL, North PS, Goulaouic H, Riou JF, Turley H, Gatter KC, & Hickson ID (2000) The Bloom's syndrome gene product interacts with topoisomerase III. *J Biol Chem* 275(13):9636-9644.
73. Yin J, Sobeck A, Xu C, Meetei AR, Hoatlin M, Li L, & Wang W (2005) BLAP75, an essential component of Bloom's syndrome protein complexes that maintain genome integrity. *Embo J* 24(7):1465-1476.
74. Bhattacharyya S, Keirse J, Russell B, Kavecansky J, Lillard-Wetherell K, Tahmaseb K, Turchi JJ, & Groden J (2009) Telomerase-associated protein 1, HSP90, and topoisomerase IIalpha associate directly with the BLM helicase in immortalized cells using ALT and modulate its helicase activity using telomeric DNA substrates. *J Biol Chem* 284(22):14966-14977.
75. Ke Y, Huh JW, Warrington R, Li B, Wu N, Leng M, Zhang J, Ball HL, & Yu H (2011) PICH and BLM limit histone association with anaphase centromeric DNA threads and promote their resolution. *Embo J* 30(16):3309-3321.
76. Wu L, Davies SL, Levitt NC, & Hickson ID (2001) Potential role for the BLM helicase in recombinational repair via a conserved interaction with RAD51. *J Biol Chem* 276(22):19375-19381.
77. Nimonkar AV, Ozsoy AZ, Genschel J, Modrich P, & Kowalczykowski SC (2008) Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. *Proc Natl Acad Sci U S A* 105(44):16906-16911.
78. Nimonkar AV, Genschel J, Kinoshita E, Polaczek P, Campbell JL, Wyman C, Modrich P, & Kowalczykowski SC (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev* 25(4):350-362.
79. Suhasini AN, Rawtani NA, Wu Y, Sommers JA, Sharma S, Mosedale G, North PS, Cantor SB, Hickson ID, & Brosh RM, Jr. (2011) Interaction between the helicases genetically linked to Fanconi anemia group J and Bloom's syndrome. *Embo J* 30(4):692-705.
80. Xu D, Muniandy P, Leo E, Yin J, Thangavel S, Shen X, Li M, Agama K, Guo R, Fox D, 3rd, Meetei AR, Wilson L, Nguyen H, Weng NP, Brill SJ, Li L, Vindigni A, Pommier Y, Seidman M, & Wang W (2010) Rif1 provides a new DNA-binding interface for the Bloom syndrome complex to maintain normal replication. *Embo J* 29(18):3140-3155.
81. Beamish H, Kedar P, Kaneko H, Chen P, Fukao T, Peng C, Beresten S, Gueven N, Purdie D, Lees-Miller S, Ellis N, Kondo N, & Lavin MF (2002) Functional link between BLM defective in Bloom's syndrome and the ataxia-telangiectasia-mutated protein, ATM. *J Biol Chem* 277(34):30515-30523.

82. Davies SL, North PS, Dart A, Lakin ND, & Hickson ID (2004) Phosphorylation of the Bloom's syndrome helicase and its role in recovery from S-phase arrest. *Mol Cell Biol* 24(3):1279-1291.
83. Sharma S, Sommers JA, Wu L, Bohr VA, Hickson ID, & Brosh RM, Jr. (2004) Stimulation of flap endonuclease-1 by the Bloom's syndrome protein. *J Biol Chem* 279(11):9847-9856.
84. Wang W & Bambara RA (2005) Human Bloom protein stimulates flap endonuclease 1 activity by resolving DNA secondary structure. *J Biol Chem* 280(7):5391-5399.
85. Selak N, Bachrati CZ, Shevelev I, Dietschy T, van Loon B, Jacob A, Hubscher U, Hoheisel JD, Hickson ID, & Stagljar I (2008) The Bloom's syndrome helicase (BLM) interacts physically and functionally with p12, the smallest subunit of human DNA polymerase delta. *Nucleic Acids Res* 36(16):5166-5179.
86. Langland G, Kordich J, Creaney J, Goss KH, Lillard-Wetherell K, Bebenek K, Kunkel TA, & Groden J (2001) The Bloom's syndrome protein (BLM) interacts with MLH1 but is not required for DNA mismatch repair. *J Biol Chem* 276(32):30031-30035.
87. Pedrazzi G, Perrera C, Blaser H, Kuster P, Marra G, Davies SL, Ryu GH, Freire R, Hickson ID, Jiricny J, & Stagljar I (2001) Direct association of Bloom's syndrome gene product with the human mismatch repair protein MLH1. *Nucleic Acids Res* 29(21):4378-4386.
88. Wang XW, Tseng A, Ellis NA, Spillare EA, Linke SP, Robles AI, Seker H, Yang Q, Hu P, Beresten S, Bemmels NA, Garfield S, & Harris CC (2001) Functional interaction of p53 and BLM DNA helicase in apoptosis. *J Biol Chem* 276(35):32948-32955.
89. Garkavtsev IV, Kley N, Grigorian IA, & Gudkov AV (2001) The Bloom syndrome protein interacts and cooperates with p53 in regulation of transcription and cell growth control. *Oncogene* 20(57):8276-8280.
90. von Kobbe C, Karmakar P, Dawut L, Opresko P, Zeng X, Brosh RM, Jr., Hickson ID, & Bohr VA (2002) Colocalization, physical, and functional interaction between Werner and Bloom syndrome proteins. *J Biol Chem* 277(24):22035-22044.
91. Stavropoulos DJ, Bradshaw PS, Li X, Pasic I, Truong K, Ikura M, Ungrin M, & Meyn MS (2002) The Bloom syndrome helicase BLM interacts with TRF2 in ALT cells and promotes telomeric DNA synthesis. *Hum Mol Genet* 11(25):3135-3144.
92. Opresko PL, von Kobbe C, Laine JP, Harrigan J, Hickson ID, & Bohr VA (2002) Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom syndrome helicases. *J Biol Chem* 277(43):41110-41119.
93. Wang JC (1996) DNA topoisomerases. *Annu Rev Biochem* 65:635-692.

94. Gangloff S, McDonald JP, Bendixen C, Arthur L, & Rothstein R (1994) The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol Cell Biol* 14(12):8391-8398.
95. Bennett RJ, Noirot-Gros MF, & Wang JC (2000) Interaction between yeast sgs1 helicase and DNA topoisomerase III. *J Biol Chem* 275(35):26898-26905.
96. Fricke WM, Kaliraman V, & Brill SJ (2001) Mapping the DNA topoisomerase III binding domain of the Sgs1 DNA helicase. *J Biol Chem* 276(12):8848-8855.
97. Bennett RJ & Wang JC (2001) Association of yeast DNA topoisomerase III and Sgs1 DNA helicase: studies of fusion proteins. *Proc Natl Acad Sci U S A* 98(20):11108-11113.
98. Harmon FG, DiGate RJ, & Kowalczykowski SC (1999) RecQ helicase and topoisomerase III comprise a novel DNA strand passage function: a conserved mechanism for control of DNA recombination. *Mol Cell* 3(5):611-620.
99. Suski C & Marians KJ (2008) Resolution of converging replication forks by RecQ and topoisomerase III. *Mol Cell* 30(6):779-789.
100. Harmon FG, Brockman JP, & Kowalczykowski SC (2003) RecQ helicase stimulates both DNA catenation and changes in DNA topology by topoisomerase III. *J Biol Chem* 278(43):42668-42678.
101. Johnson FB, Lombard DB, Neff NF, Mastrangelo MA, Dewolf W, Ellis NA, Marciniak RA, Yin Y, Jaenisch R, & Guarente L (2000) Association of the Bloom syndrome protein with topoisomerase IIIalpha in somatic and meiotic cells. *Cancer Res* 60(5):1162-1167.
102. Hu P, Beresten SF, van Brabant AJ, Ye TZ, Pandolfi PP, Johnson FB, Guarente L, & Ellis NA (2001) Evidence for BLM and Topoisomerase IIIalpha interaction in genomic stability. *Hum Mol Genet* 10(12):1287-1298.
103. Wu L & Hickson ID (2002) The Bloom's syndrome helicase stimulates the activity of human topoisomerase IIIalpha. *Nucleic Acids Res* 30(22):4823-4829.
104. Yang J, Bachrati CZ, Ou J, Hickson ID, & Brown GW (2010) Human topoisomerase IIIalpha is a single-stranded DNA decatenase that is stimulated by BLM and RMI1. *J Biol Chem* 285(28):21426-21436.
105. Wu L & Hickson ID (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* 426(6968):870-874.
106. Szostak JW, Orr-Weaver TL, Rothstein RJ, & Stahl FW (1983) The double-strand-break repair model for recombination. *Cell* 33(1):25-35.

107. Ip SC, Rass U, Blanco MG, Flynn HR, Skehel JM, & West SC (2008) Identification of Holliday junction resolvases from humans and yeast. *Nature* 456(7220):357-361.
108. Wechsler T, Newman S, & West SC (2011) Aberrant chromosome morphology in human cells defective for Holliday junction resolution. *Nature* 471(7340):642-646.
109. Luo G, Santoro IM, McDaniel LD, Nishijima I, Mills M, Youssoufian H, Vogel H, Schultz RA, & Bradley A (2000) Cancer predisposition caused by elevated mitotic recombination in Bloom mice. *Nat Genet* 26(4):424-429.
110. LaRocque JR, Stark JM, Oh J, Bojilova E, Yusa K, Horie K, Takeda J, & Jasin M (2011) Interhomolog recombination and loss of heterozygosity in wild-type and Bloom syndrome helicase (BLM)-deficient mammalian cells. *Proc Natl Acad Sci U S A* 108(29):11971-11976.
111. Wu L, Bachrati CZ, Ou J, Xu C, Yin J, Chang M, Wang W, Li L, Brown GW, & Hickson ID (2006) BLAP75/RMI1 promotes the BLM-dependent dissolution of homologous recombination intermediates. *Proc Natl Acad Sci U S A* 103(11):4068-4073.
112. Raynard S, Bussen W, & Sung P (2006) A double Holliday junction dissolvosome comprising BLM, topoisomerase IIIalpha, and BLAP75. *J Biol Chem* 281(20):13861-13864.
113. Raynard S, Zhao W, Bussen W, Lu L, Ding YY, Busygina V, Meetei AR, & Sung P (2008) Functional role of BLAP75 in BLM-topoisomerase IIIalpha-dependent holliday junction processing. *J Biol Chem* 283(23):15701-15708.
114. Murzin AG (1993) Ob(Oligonucleotide Oligosaccharide Binding)-Fold - Common Structural and Functional Solution for Nonhomologous Sequences. *EMBO J* 12(3):861-867.
115. Xu D, Guo R, Sobeck A, Bachrati CZ, Yang J, Enomoto T, Brown GW, Hoatlin ME, Hickson ID, & Wang W (2008) RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. *Genes Dev* 22(20):2843-2855.
116. Chang M, Bellaoui M, Zhang C, Desai R, Morozov P, Delgado-Cruzata L, Rothstein R, Freyer GA, Boone C, & Brown GW (2005) RMI1/NCE4, a suppressor of genome instability, encodes a member of the RecQ helicase/Topo III complex. *Embo J* 24(11):2024-2033.
117. Mullen JR, Nallaseth FS, Lan YQ, Slagle CE, & Brill SJ (2005) Yeast Rmi1/Nce4 controls genome stability as a subunit of the Sgs1-Top3 complex. *Mol Cell Biol* 25(11):4476-4487.
118. Chen CF & Brill SJ (2007) Binding and activation of DNA topoisomerase III by the Rmi1 subunit. *J Biol Chem* 282(39):28971-28979.

119. Singh TR, Ali AM, Busygina V, Raynard S, Fan Q, Du CH, Andreassen PR, Sung P, & Meetei AR (2008) BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvasome. *Genes Dev* 22(20):2856-2868.
120. Wang F, Yang Y, Singh TR, Busygina V, Guo R, Wan K, Wang W, Sung P, Meetei AR, & Lei M (2010) Crystal structures of RMI1 and RMI2, two OB-fold regulatory subunits of the BLM complex. *Structure* 18(9):1159-1170.
121. Hoadley KA, Xu D, Xue Y, Satyshur KA, Wang W, & Keck JL (2010) Structure and cellular roles of the RMI core complex from the Bloom syndrome dissolvasome. *Structure* 18(9):1149-1158.
122. Gravel S, Chapman JR, Magill C, & Jackson SP (2008) DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes Dev* 22(20):2767-2772.
123. Doherty KM, Sharma S, Uzdilla LA, Wilson TM, Cui S, Vindigni A, & Brosh RM, Jr. (2005) RECQ1 helicase interacts with human mismatch repair factors that regulate genetic recombination. *J Biol Chem* 280(30):28085-28094.
124. Sharma S, Sommers JA, Driscoll HC, Uzdilla L, Wilson TM, & Brosh RM, Jr. (2003) The exonucleolytic and endonucleolytic cleavage activities of human exonuclease 1 are stimulated by an interaction with the carboxyl-terminal region of the Werner syndrome protein. *J Biol Chem* 278(26):23487-23496.
125. Liao S, Toczylowski T, & Yan H (2008) Identification of the *Xenopus* DNA2 protein as a major nuclease for the 5'->3' strand-specific processing of DNA ends. *Nucleic Acids Res* 36(19):6091-6100.
126. Mimitou EP & Symington LS (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* 455(7214):770-774.
127. Zhu Z, Chung WH, Shim EY, Lee SE, & Ira G (2008) Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell* 134(6):981-994.
128. Cejka P, Cannavo E, Polaczek P, Masuda-Sasa T, Pokharel S, Campbell JL, & Kowalczykowski SC (2010) DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature* 467(7311):112-116.
129. Niu H, Chung WH, Zhu Z, Kwon Y, Zhao W, Chi P, Prakash R, Seong C, Liu D, Lu L, Ira G, & Sung P (2010) Mechanism of the ATP-dependent DNA end-resection machinery from *Saccharomyces cerevisiae*. *Nature* 467(7311):108-111.
130. Gaymes TJ, North PS, Brady N, Hickson ID, Mufti GJ, & Rassool FV (2002) Increased error-prone non homologous DNA end-joining--a proposed mechanism of chromosomal instability in Bloom's syndrome. *Oncogene* 21(16):2525-2533.

131. Symington LS & Gautier J (2011) Double-strand break end resection and repair pathway choice. *Annual review of genetics* 45:247-271.
132. Dutertre S, Ababou M, Onclercq R, Delic J, Chatton B, Jaulin C, & Amor-Gueret M (2000) Cell cycle regulation of the endogenous wild type Bloom's syndrome DNA helicase. *Oncogene* 19(23):2731-2738.
133. Sanz MM, Proytcheva M, Ellis NA, Holloman WK, & German J (2000) BLM, the Bloom's syndrome protein, varies during the cell cycle in its amount, distribution, and co-localization with other nuclear proteins. *Cytogenet Cell Genet* 91(1-4):217-223.
134. Bischof O, Kim SH, Irving J, Beresten S, Ellis NA, & Campisi J (2001) Regulation and localization of the Bloom syndrome protein in response to DNA damage. *J Cell Biol* 153(2):367-380.
135. San Filippo J, Sung P, & Klein H (2008) Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* 77:229-257.
136. Bachrati CZ, Borts RH, & Hickson ID (2006) Mobile D-loops are a preferred substrate for the Bloom's syndrome helicase. *Nucleic Acids Res* 34(8):2269-2279.
137. Bugreev DV, Yu X, Egelman EH, & Mazin AV (2007) Novel pro- and anti-recombination activities of the Bloom's syndrome helicase. *Genes Dev* 21(23):3085-3094.
138. Bugreev DV, Mazina OM, & Mazin AV (2009) Bloom syndrome helicase stimulates RAD51 DNA strand exchange activity through a novel mechanism. *J Biol Chem* 284(39):26349-26359.
139. Kikuchi K, Abdel-Aziz HI, Taniguchi Y, Yamazoe M, Takeda S, & Hirota K (2009) Bloom DNA helicase facilitates homologous recombination between diverged homologous sequences. *J Biol Chem* 284(39):26360-26367.
140. Hanada K, Ukita T, Kohno Y, Saito K, Kato J, & Ikeda H (1997) RecQ DNA helicase is a suppressor of illegitimate recombination in *Escherichia coli*. *Proc Natl Acad Sci U S A* 94(8):3860-3865.
141. Hand R & German J (1975) A retarded rate of DNA chain growth in Bloom's syndrome. *Proc Natl Acad Sci U S A* 72(2):758-762.
142. Ockey CH & Saffhill R (1986) Delayed DNA maturation, a possible cause of the elevated sister-chromatid exchange in Bloom's syndrome. *Carcinogenesis* 7(1):53-57.
143. Lonn U, Lonn S, Nylen U, Winblad G, & German J (1990) An abnormal profile of DNA replication intermediates in Bloom's syndrome. *Cancer Res* 50(11):3141-3145.

144. Sengupta S, Linke SP, Pedoux R, Yang Q, Farnsworth J, Garfield SH, Valerie K, Shay JW, Ellis NA, Wasylyk B, & Harris CC (2003) BLM helicase-dependent transport of p53 to sites of stalled DNA replication forks modulates homologous recombination. *Embo J* 22(5):1210-1222.
145. Wu L (2007) Role of the BLM helicase in replication fork management. *DNA Repair (Amst)* 6(7):936-944.
146. Petermann E & Helleday T (2010) Pathways of mammalian replication fork restart. *Nat Rev Mol Cell Biol* 11(10):683-687.
147. Jones RM & Petermann E (2012) Replication fork dynamics and the DNA damage response. *Biochem J* 443(1):13-26.
148. Karow JK, Constantinou A, Li JL, West SC, & Hickson ID (2000) The Bloom's syndrome gene product promotes branch migration of holliday junctions. *Proc Natl Acad Sci U S A* 97(12):6504-6508.
149. Ralf C, Hickson ID, & Wu L (2006) The Bloom's syndrome helicase can promote the regression of a model replication fork. *J Biol Chem* 281(32):22839-22846.
150. Machwe A, Xiao L, Groden J, & Orren DK (2006) The Werner and Bloom syndrome proteins catalyze regression of a model replication fork. *Biochemistry* 45(47):13939-13946.
151. Tuduri S, Tourriere H, & Pasero P (2010) Defining replication origin efficiency using DNA fiber assays. *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 18(1):91-102.
152. Davies SL, North PS, & Hickson ID (2007) Role for BLM in replication-fork restart and suppression of origin firing after replicative stress. *Nat Struct Mol Biol* 14(7):677-679.
153. Rao VA, Conti C, Guirouilh-Barbat J, Nakamura A, Miao ZH, Davies SL, Sacca B, Hickson ID, Bensimon A, & Pommier Y (2007) Endogenous gamma-H2AX-ATM-Chk2 checkpoint activation in Bloom's syndrome helicase deficient cells is related to DNA replication arrested forks. *Molecular cancer research : MCR* 5(7):713-724.
154. Yang J, O'Donnell L, Durocher D, & Brown GW (2012) RMI1 promotes DNA replication fork progression and recovery from replication fork stress. *Mol Cell Biol* 32(15):3054-3064.
155. Meetei AR, Sechi S, Wallisch M, Yang D, Young MK, Joenje H, Hoatlin ME, & Wang W (2003) A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome. *Mol Cell Biol* 23(10):3417-3426.

156. Deans AJ & West SC (2009) FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. *Mol Cell* 36(6):943-953.
157. Moldovan GL & D'Andrea AD (2009) How the Fanconi anemia pathway guards the genome. *Annu Rev Genet* 43:223-249.
158. Kee Y & D'Andrea AD (2010) Expanded roles of the Fanconi anemia pathway in preserving genomic stability. *Genes Dev* 24(16):1680-1694.
159. Niedernhofer LJ, Lalai AS, & Hoeijmakers JH (2005) Fanconi anemia (cross)linked to DNA repair. *Cell* 123(7):1191-1198.
160. Kim H & D'Andrea AD (2012) Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev* 26(13):1393-1408.
161. Nakanishi K, Cavallo F, Perrouault L, Giovannangeli C, Moynahan ME, Barchi M, Brunet E, & Jasin M (2011) Homology-directed Fanconi anemia pathway cross-link repair is dependent on DNA replication. *Nat Struct Mol Biol* 18(4):500-503.
162. Knipscheer P, Raschle M, Smogorzewska A, Enou M, Ho TV, Scharer OD, Elledge SJ, & Walter JC (2009) The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* 326(5960):1698-1701.
163. Wang W (2007) Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet* 8(10):735-748.
164. Pichierri P, Franchitto A, & Rosselli F (2004) BLM and the FANC proteins collaborate in a common pathway in response to stalled replication forks. *Embo J* 23(15):3154-3163.
165. Meetei AR, Medhurst AL, Ling C, Xue Y, Singh TR, Bier P, Steltenpool J, Stone S, Dokal I, Mathew CG, Hoatlin M, Joenje H, de Winter JP, & Wang W (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet* 37(9):958-963.
166. Komori K, Fujikane R, Shinagawa H, & Ishino Y (2002) Novel endonuclease in Archaea cleaving DNA with various branched structure. *Genes & genetic systems* 77(4):227-241.
167. Scheller J, Schurer A, Rudolph C, Hettwer S, & Kramer W (2000) MPH1, a yeast gene encoding a DEAH protein, plays a role in protection of the genome from spontaneous and chemically induced damage. *Genetics* 155(3):1069-1081.
168. Sun W, Nandi S, Osman F, Ahn JS, Jakovleska J, Lorenz A, & Whitby MC (2008) The FANCM ortholog Fml1 promotes recombination at stalled replication forks and limits crossing over during DNA double-strand break repair. *Mol Cell* 32(1):118-128.

169. Yan Z, Delannoy M, Ling C, Dae D, Osman F, Muniandy PA, Shen X, Oostra AB, Du H, Steltenpool J, Lin T, Schuster B, Decaillet C, Stasiak A, Stasiak AZ, Stone S, Hoatlin ME, Schindler D, Woodcock CL, Joenje H, Sen R, de Winter JP, Li L, Seidman MM, Whitby MC, Myung K, Constantinou A, & Wang W (2010) A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. *Mol Cell* 37(6):865-878.
170. Singh TR, Saro D, Ali AM, Zheng XF, Du CH, Killen MW, Sachpatzidis A, Wahengbam K, Pierce AJ, Xiong Y, Sung P, & Meetei AR (2010) MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Mol Cell* 37(6):879-886.
171. Ciccia A, Ling C, Coulthard R, Yan Z, Xue Y, Meetei AR, Laghmani el H, Joenje H, McDonald N, de Winter JP, Wang W, & West SC (2007) Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. *Mol Cell* 25(3):331-343.
172. Rosado IV, Niedzwiedz W, Alpi AF, & Patel KJ (2009) The Walker B motif in avian FANCM is required to limit sister chromatid exchanges but is dispensable for DNA crosslink repair. *Nucleic Acids Res* 37(13):4360-4370.
173. Niedzwiedz W, Mosedale G, Johnson M, Ong CY, Pace P, & Patel KJ (2004) The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol Cell* 15(4):607-620.
174. Hirano S, Yamamoto K, Ishiai M, Yamazoe M, Seki M, Matsushita N, Ohzeki M, Yamashita YM, Arakawa H, Buerstedde JM, Enomoto T, Takeda S, Thompson LH, & Takata M (2005) Functional relationships of FANCC to homologous recombination, translesion synthesis, and BLM. *Embo J* 24(2):418-427.
175. Bridge WL, Vandenberg CJ, Franklin RJ, & Hiom K (2005) The BRIP1 helicase functions independently of BRCA1 in the Fanconi anemia pathway for DNA crosslink repair. *Nat Genet* 37(9):953-957.
176. Xue Y, Li Y, Guo R, Ling C, & Wang W (2008) FANCM of the Fanconi anemia core complex is required for both monoubiquitination and DNA repair. *Hum Mol Genet* 17(11):1641-1652.
177. Singh TR, Bakker ST, Agarwal S, Jansen M, Grassman E, Godthelp BC, Ali AM, Du CH, Rooimans MA, Fan Q, Wahengbam K, Steltenpool J, Andreassen PR, Williams DA, Joenje H, de Winter JP, & Meetei AR (2009) Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. *Blood* 114(1):174-180.

178. Gari K, Decaillet C, Delannoy M, Wu L, & Constantinou A (2008) Remodeling of DNA replication structures by the branch point translocase FANCM. *Proc Natl Acad Sci U S A* 105(42):16107-16112.
179. Collis SJ, Ciccio A, Deans AJ, Horejsi Z, Martin JS, Maslen SL, Skehel JM, Elledge SJ, West SC, & Boulton SJ (2008) FANCM and FAAP24 function in ATR-mediated checkpoint signaling independently of the Fanconi anemia core complex. *Mol Cell* 32(3):313-324.
180. Luke-Glaser S, Luke B, Grossi S, & Constantinou A (2010) FANCM regulates DNA chain elongation and is stabilized by S-phase checkpoint signalling. *Embo J* 29(4):795-805.
181. Schwab RA, Blackford AN, & Niedzwiedz W (2010) ATR activation and replication fork restart are defective in FANCM-deficient cells. *Embo J* 29(4):806-818.
182. Blackford AN, Schwab RA, Nieminszczy J, Deans AJ, West SC, & Niedzwiedz W (2012) The DNA translocase activity of FANCM protects stalled replication forks. *Hum Mol Genet* 21(9):2005-2016.
183. Hoadley KA, Xue Y, Ling C, Takata M, Wang W, & Keck JL (2012) Defining the molecular interface that connects the Fanconi anemia protein FANCM to the Bloom syndrome dissolvosome. *Proc Natl Acad Sci U S A* 109(12):4437-4442.
184. Cantor S, Drapkin R, Zhang F, Lin Y, Han J, Pamidi S, & Livingston DM (2004) The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proc Natl Acad Sci U S A* 101(8):2357-2362.
185. Suhasini AN & Brosh RM, Jr. (2012) Fanconi anemia and Bloom's syndrome crosstalk through FANCF-BLM helicase interaction. *Trends Genet* 28(1):7-13.
186. Wu Y, Shin-ya K, & Brosh RM, Jr. (2008) FANCF helicase defective in Fanconi anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol Cell Biol* 28(12):4116-4128.
187. London TB, Barber LJ, Mosedale G, Kelly GP, Balasubramanian S, Hickson ID, Boulton SJ, & Hiom K (2008) FANCF is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J Biol Chem* 283(52):36132-36139.
188. Sarkies P, Murat P, Phillips LG, Patel KJ, Balasubramanian S, & Sale JE (2012) FANCF coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. *Nucleic Acids Res* 40(4):1485-1498.
189. Abraham RT (2001) Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev* 15(17):2177-2196.

190. Syljuasen RG, Sorensen CS, Hansen LT, Fugger K, Lundin C, Johansson F, Helleday T, Sehested M, Lukas J, & Bartek J (2005) Inhibition of human Chk1 causes increased initiation of DNA replication, phosphorylation of ATR targets, and DNA breakage. *Mol Cell Biol* 25(9):3553-3562.
191. O'Driscoll M & Jeggo PA (2006) The role of double-strand break repair - insights from human genetics. *Nat Rev Genet* 7(1):45-54.
192. Rao VA, Fan AM, Meng L, Doe CF, North PS, Hickson ID, & Pommier Y (2005) Phosphorylation of BLM, dissociation from topoisomerase IIIalpha, and colocalization with gamma-H2AX after topoisomerase I-induced replication damage. *Mol Cell Biol* 25(20):8925-8937.
193. Ababou M, Dutertre S, Lecluse Y, Onclercq R, Chatton B, & Amor-Gueret M (2000) ATM-dependent phosphorylation and accumulation of endogenous BLM protein in response to ionizing radiation. *Oncogene* 19(52):5955-5963.
194. Taniguchi T, Garcia-Higuera I, Xu B, Andreassen PR, Gregory RC, Kim ST, Lane WS, Kastan MB, & D'Andrea AD (2002) Convergence of the Fanconi anemia and ataxia telangiectasia signaling pathways. *Cell* 109(4):459-472.
195. Pichierri P & Rosselli F (2004) The DNA crosslink-induced S-phase checkpoint depends on ATR-CHK1 and ATR-NBS1-FANCD2 pathways. *Embo J* 23(5):1178-1187.
196. Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER, 3rd, Hurov KE, Luo J, Ballif BA, Gygi SP, Hofmann K, D'Andrea AD, & Elledge SJ (2007) Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. *Cell* 129(2):289-301.
197. Ishiai M, Kitao H, Smogorzewska A, Tomida J, Kinomura A, Uchida E, Saberi A, Kinoshita E, Kinoshita-Kikuta E, Koike T, Tashiro S, Elledge SJ, & Takata M (2008) FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nat Struct Mol Biol* 15(11):1138-1146.
198. Collins NB, Wilson JB, Bush T, Thomashevski A, Roberts KJ, Jones NJ, & Kupfer GM (2009) ATR-dependent phosphorylation of FANCA on serine 1449 after DNA damage is important for FA pathway function. *Blood* 113(10):2181-2190.
199. Wang X, Kennedy RD, Ray K, Stuckert P, Ellenberger T, & D'Andrea AD (2007) Chk1-mediated phosphorylation of FANCE is required for the Fanconi anemia/BRCA pathway. *Mol Cell Biol* 27(8):3098-3108.
200. Wang Y, Leung JW, Jiang Y, Lowery MG, Do H, Vasquez KM, Chen J, Wang W, & Li L (2013) FANCM and FAAP24 Maintain Genome Stability via Cooperative as Well as Unique Functions. *Mol Cell* 49(5):997-1009.

201. Gong Z, Kim JE, Leung CC, Glover JN, & Chen J (2010) BACH1/FANCI acts with TopBP1 and participates early in DNA replication checkpoint control. *Mol Cell* 37(3):438-446.
202. Baumann C, Korner R, Hofmann K, & Nigg EA (2007) PICH, a centromere-associated SNF2 family ATPase, is regulated by Plk1 and required for the spindle checkpoint. *Cell* 128(1):101-114.
203. Chan KL, North PS, & Hickson ID (2007) BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *Embo J* 26(14):3397-3409.
204. Chan KL & Hickson ID (2011) New insights into the formation and resolution of ultrafine anaphase bridges. *Seminars in cell & developmental biology* 22(8):906-912.
205. Wang LH, Schwarzbach T, Speicher MR, & Nigg EA (2008) Persistence of DNA threads in human anaphase cells suggests late completion of sister chromatid decatenation. *Chromosoma* 117(2):123-135.
206. Rouzeau S, Cordelieres FP, Buhagiar-Labarchede G, Hurbain I, Onclercq-Delic R, Gemble S, Magnaghi-Jaulin L, Jaulin C, & Amor-Gueret M (2012) Bloom's syndrome and PICH helicases cooperate with topoisomerase IIalpha in centromere disjunction before anaphase. *PloS one* 7(4):e33905.
207. Spence JM, Phua HH, Mills W, Carpenter AJ, Porter AC, & Farr CJ (2007) Depletion of topoisomerase IIalpha leads to shortening of the metaphase interkinetochore distance and abnormal persistence of PICH-coated anaphase threads. *J Cell Sci* 120(Pt 22):3952-3964.
208. Russell B, Bhattacharyya S, Keirsey J, Sandy A, Grierson P, Perchiniak E, Kavecansky J, Acharya S, & Groden J (2011) Chromosome breakage is regulated by the interaction of the BLM helicase and topoisomerase IIalpha. *Cancer Res* 71(2):561-571.
209. Chan KL, Palmai-Pallag T, Ying S, & Hickson ID (2009) Replication stress induces sister-chromatid bridging at fragile site loci in mitosis. *Nat Cell Biol* 11(6):753-760.
210. Naim V & Rosselli F (2009) The FANCD pathway and BLM collaborate during mitosis to prevent micro-nucleation and chromosome abnormalities. *Nat Cell Biol* 11(6):761-768.
211. Letessier A, Millot GA, Koundrioukoff S, Lachages AM, Vogt N, Hansen RS, Malfoy B, Brison O, & Debatisse M (2011) Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature* 470(7332):120-123.
212. Ozeri-Galai E, Lebofsky R, Rahat A, Bester AC, Bensimon A, & Kerem B (2011) Failure of origin activation in response to fork stalling leads to chromosomal instability at fragile sites. *Mol Cell* 43(1):122-131.

213. Vinciguerra P, Godinho SA, Parmar K, Pellman D, & D'Andrea AD (2010) Cytokinesis failure occurs in Fanconi anemia pathway-deficient murine and human bone marrow hematopoietic cells. *J Clin Invest* 120(11):3834-3842.
214. Lukas C, Savic V, Bekker-Jensen S, Doil C, Neumann B, Pedersen RS, Grofte M, Chan KL, Hickson ID, Bartek J, & Lukas J (2011) 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat Cell Biol* 13(3):243-253.
215. Harrigan JA, Belotserkovskaya R, Coates J, Dimitrova DS, Polo SE, Bradshaw CR, Fraser P, & Jackson SP (2011) Replication stress induces 53BP1-containing OPT domains in G1 cells. *J Cell Biol* 193(1):97-108.
216. Adamo A, Collis SJ, Adelman CA, Silva N, Horejsi Z, Ward JD, Martinez-Perez E, Boulton SJ, & La Volpe A (2010) Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia. *Mol Cell* 39(1):25-35.
217. Pace P, Mosedale G, Hodskinson MR, Rosado IV, Sivasubramaniam M, & Patel KJ (2010) Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science* 329(5988):219-223.
218. Karanja KK, Cox SW, Duxin JP, Stewart SA, & Campbell JL (2012) DNA2 and EXO1 in replication-coupled, homology-directed repair and in the interplay between HDR and the FA/BRCA network. *Cell Cycle* 11(21):3983-3996.
219. Nguyen GH, Dexheimer TS, Rosenthal AS, Chu WK, Singh DK, Mosedale G, Bachrati CZ, Schultz L, Sakurai M, Savitsky P, Abu M, McHugh PJ, Bohr VA, Harris CC, Jadhav A, Gileadi O, Maloney DJ, Simeonov A, & Hickson ID (2013) A Small Molecule Inhibitor of the BLM Helicase Modulates Chromosome Stability in Human Cells. *Chem Biol* 20(1):55-62.
220. Delano WL (2002) *The PyMol Molecular Graphics System*. (DeLano Scientific, San Carlos, CA).
221. Li W & Wang JC (1998) Mammalian DNA topoisomerase IIIalpha is essential in early embryogenesis. *Proc Natl Acad Sci U S A* 95(3):1010-1013.
222. Seki M, Nakagawa T, Seki T, Kato G, Tada S, Takahashi Y, Yoshimura A, Kobayashi T, Aoki A, Otsuki M, Habermann FA, Tanabe H, Ishii Y, & Enomoto T (2006) Bloom helicase and DNA topoisomerase IIIalpha are involved in the dissolution of sister chromatids. *Mol Cell Biol* 26(16):6299-6307.
223. Wang W, Seki M, Narita Y, Sonoda E, Takeda S, Yamada K, Masuko T, Katada T, & Enomoto T (2000) Possible association of BLM in decreasing DNA double strand breaks during DNA replication. *Embo J* 19(13):3428-3435.

CHAPTER 2

Structure and cellular roles of the RMI core complex from the Bloom syndrome dissolvasome

This work has been published:

Hoadley KA, Xu D, Xue Y, Satyshur KA, Wang W, Keck JL (2010) “Structure and cellular roles of the RMI core complex from the Bloom syndrome dissolvasome”. *Structure* **18**, 1149-58.

Kelly A. Manthei (nee Hoadley) performed protein purification, limited proteolysis experiment, structure determinations, and circular dichroism experiments. Kenneth A. Satyshur assisted in structure determination. Dongyi Xu, Yutong Xue, and Weidong Wang designed and executed the coimmunoprecipitation and sister chromatid exchange assays.

Summary

BLM, the protein product of the gene mutated in Bloom syndrome, is one of five human RecQ helicases. It functions to separate double Holliday junction DNA without genetic exchange as a component of the “dissolvasome”, which also includes topoisomerase III α and the RMI (RecQ-mediated genome instability) subcomplex (RMI1 and RMI2). We describe the crystal structure of the RMI core complex, comprising RMI2 and the C-terminal OB domain of RMI1. The overall RMI core structure strongly resembles two-thirds of the trimerization core of the eukaryotic single-strand DNA-binding protein, Replication Protein A. Immunoprecipitation experiments with RMI2 variants confirm key interactions that stabilize the RMI core interface. Disruption of this interface leads to a dramatic increase in cellular sister chromatid exchange events similar to that seen in BLM-deficient cells. The RMI core interface is therefore crucial for BLM dissolvasome assembly and may have additional cellular roles as a docking hub for other proteins.

Introduction

Bloom syndrome (BS) is a rare autosomal recessive genetic disorder characterized by proportional dwarfism, sun sensitivity, an increased susceptibility to infections and diabetes, and a high occurrence of most types of cancer (1, 2). BS patients are often first diagnosed with cancer in their mid-twenties, which is the typical cause of death. Cells derived from persons with BS show a remarkable increase in chromosomal rearrangements, gaps, and breaks. The hallmark feature of BS cells, which is used in diagnosis, is a ten-fold elevated frequency of sister chromatid exchanges (SCEs). The gene mutated in BS, *BLM*, encodes one of five human RecQ DNA helicases. Two other RecQ-linked genetic diseases are caused by defects in *WRN* and *RecQ4*, causing Werner's syndrome and Rothmund-Thomson syndrome, respectively. All three of these genetic disorders are characterized by increased genomic instability and cancer predisposition, highlighting the importance of RecQ proteins in genome maintenance (2).

RecQ DNA helicases are ubiquitously conserved enzymes that catalyze the ATP-dependent unwinding of DNA with a preference for DNA structures resembling replication and recombination intermediates such as replication forks, Holliday junctions, D-loops, and G-quadruplexes (3). For the BLM protein, these activities drive multiple cellular activities. BLM is important in homologous recombination (HR)-dependent DNA repair, with roles in branch migration (4-6), RAD51 filament disruption (7), and in restarting stalled DNA replication processes (8). For several of these reactions, BLM is thought to function as a component of a larger protein complex termed the "dissolvasome", in which it partners with topoisomerase III α

(Top3 α) and the RMI (*RecQ-mediated genome instability*) subcomplex, comprised of RMI1 and RMI2 (9-13).

The dissolvasome specifically acts on double Holliday junction intermediates in HR, creating non-crossover products in a process termed “dissolution” (12, 14). This is of particular interest both for the normal function of the dissolvasome in HR and for BS pathology: since only non-crossover products are formed, a defect in this process could explain the increase in SCEs in Bloom syndrome. Dissolution activity is proposed to occur through branch migration by BLM, which creates a hemicatenated structure that can then be resolved by Top3 α (14). RMI1 stimulates this process (12) and the more recently discovered RMI2 is thought to provide stability to the complex (9, 10).

Sequence analysis of RMI1 and RMI2 indicates that the proteins have three oligonucleotide/oligosaccharide binding (OB) folds: two in RMI1 (OB1 and OB2) and one in RMI2 (OB3) (Figure 2.2A). RMI1 OB1 interacts with BLM and Top3 α , whereas OB2 interacts with OB3 from RMI2 (9). Interestingly, the OB2 and OB3 domains both share sequence homology with domains found within Replication Protein A (RPA), the heterotrimeric eukaryotic single-stranded DNA binding protein (9). RPA contains six OB folds, which are also known as DNA binding domains (DBDs) because of their importance in ssDNA binding. A subset of these domains in RPA is responsible for the assembly of RPA heterotrimers, with one OB domain from each RPA subunit (RPA70 DBD-C, RPA32 DBD-D, and RPA14) contributing to a core trimerization structure. RPA trimerization is primarily driven by formation of a parallel

bundle of α -helices from the OB domains (15, 16). RMI OB2 and OB3 domains share significant sequence similarity with two of the three trimerization domains from RPA (RPA70 DBD-C and RPA32 DBD-D, respectively), but whether the RMI complex interface was similar to RPA was not known.

Here we describe the high-resolution X-ray crystal structure of the human RMI core complex, composed of RMI1 OB2 and RMI2. Similarly to the arrangement observed in RPA, the RMI core complex associates via OB folds that are flanked by C-terminal α -helices that form a helical bundle. This overall structure resembles two-thirds of the RPA trimerization core (RPA70 DBD-C/RPA32 DBD-D), but the specific interactions at the interface are very different. The RPA interface mostly consists of hydrophobic interactions between C-terminal helices, whereas the RMI interface is much larger and includes electrostatic interactions as well as an extensive hydrophobic interface. We have identified residues critical to the RMI interface and show that disruption of the RMI subcomplex leads to an overall increase in SCEs *in vivo*. These findings point to a primary role for the RMI core complex in stabilizing the BLM dissolvasome and also hint at possible functions for the core as a hub for docking additional cellular factors.

Results

Defining the RMI core complex. We initially set out to investigate the structure of full-length human RMI1 in complex with RMI2 (the RMI subcomplex). However, isolated RMI1 and RMI1 in complex with RMI2 were prone to proteolytic degradation, making either a poor crystallographic target. Limited proteolysis was used to define a stable complex that would be more appropriate for crystallographic trials. The RMI subcomplex was purified and subjected to limited proteolytic treatment with subtilisin, trypsin, or α -chymotrypsin and then analyzed by SDS-PAGE. The stability of RMI2 was remarkable, as it was almost completely resistant to all three proteases, while RMI1 was rapidly degraded by each of the proteases. Trypsin treatment produced a fragment migrating near the 16.5-kDa marker, which appeared after 5 minutes and was stable throughout the experiment (Figure 2.1). Mass spectrometric analysis determined that this fragment included the C-terminus of RMI1 (residues 473-625), which comprises the C-terminal RMI1 OB-fold (OB2) (Figure 2.2A). This protease-resistant complex consisting of RMI1 OB2 and RMI2 defined a minimal RMI core complex. The RMI core complex was co-expressed and purified for subsequent structural and biochemical experiments. The complex eluted with 1:1 stoichiometry in size-exclusion chromatographic analysis (Figure 2.1) and trypsin proteolysis yielded no degradation, consistent with formation of a stable minimal heterodimeric RMI core.

Crystal structure of the RMI core complex. The RMI core complex was crystallized and its structure was determined to 1.55-Å resolution using single wavelength anomalous dispersion to phase a selenomethionine-substituted variant of the complex (Table 2.1). The refined structure

contains residues 475-625 of RMI1 and residues 17-147 of RMI2. Electron density defining the N-terminus of RMI2 was missing, which could indicate that this region of the domain is structurally dynamic. The same N-terminal segment is not well conserved among RMI2 homologs (9). The RMI core complex structure was refined with good bond geometry and crystallographic statistics (Table 2.1).

Within the RMI core complex, both the RMI1 OB2 domain and RMI2 fold to form OB domains that are flanked by additional helical elements (Figure 2.2B). The canonical OB fold is defined by a five-stranded β -barrel structure that is often supported by an α -helix present in the region between the third and fourth strands (17). The RMI1 OB2 domain has each of these elements along with two additional helices between strands 3 and 4 and a C-terminal helix. These helices are structurally similar to elements that are present in RPA, as is described below. RMI2 lacks the canonical OB-fold helix but clearly forms the core β -barrel OB-fold structure with an N-terminal helix that packs against the β -barrel RMI2 core. As is the case with the RMI1 OB2 domain, RMI2 also has an additional C-terminal helix that is similar to a helix found in RPA.

Comparison of the RMI core complex and the RPA trimerization core. Structural comparison of the RMI core complex to other OB-domain proteins revealed a striking similarity with the core trimerization domain of RPA, the major eukaryotic single-strand DNA-binding protein. The RPA trimerization core interface is formed among OB folds present in RPA70, RPA32, and RPA14, with C-terminal α -helices that flank the core OB folds. These helices form a parallel trihelical bundle that provides a large portion of the interface (16) (Figure 2.2C). In a

similar fashion, the interface between RMI1 and RMI2 is stabilized by contacts between the two C-terminal α -helices, which align in parallel. A least-squares alignment (18) of RPA70 DBD-C with the RMI1 OB2 domain correspondingly aligns RMI2 with RPA32 DBD-D, with the two C-terminal helices aligning analogously (Figure 2.2D). The RPA70 DBD-C and RMI1 OB2 domains align with a root mean square deviation (rmsd) of 3.2 Å for 122 common C_α atoms, which includes structurally similar helical motifs positioned in the loop between β -strands 3 and 4 (L_{34} loop) of both proteins. The RMI1 and RPA70 L_{34} loops are distinct, however, as the helices in RMI1 L_{34} extend further from the core OB domain than those from RPA70. A second major difference between these two domains is that the loop connecting β -strands 2 and 3 (L_{23} loop) in RPA70 DBD-C contains a zinc ribbon motif whereas the corresponding loop in RMI1 lacks this motif. This RPA70 zinc-binding region is thought to contribute to DNA binding in RPA (16, 19, 20).

Least-squares alignment of RMI2 with RPA32 DBD-D, which is the most similar protein to RMI2 in the Protein Data Bank, show that the two proteins align with a rmsd of 2.2 Å for 112 common C_α atoms (Figure 2.2E). Interestingly, RMI2 also shares significant structural similarity with *S. pombe* Stn1N (2.5 Å for 115 common C_α atoms) and Ten1 (2.3 Å for 102 common C_α atoms) which, along with Cdc13, form a heterotrimeric RPA-like complex called CST involved in telomere maintenance (Gao et al., 2007; Sun et al., 2009). The CST complex is conserved in humans as Stn1, Ten1, and Ctc1, and is thought to assemble through an RPA-like clustering of OB domains (Miyake et al., 2009; Surovtseva et al., 2009).

There are two key functional differences between the RMI and RPA complexes that appear to be reflected in their respective core structures. The first is that RPA binds DNA whereas previous studies indicate that RMI does not (9). Accordingly, the aromatic residues used for DNA-binding in RPA are not conserved in the structure of the RMI core complex (Figure 2.2E). The second major difference is that the RMI core complex exists as a heterodimer, but the RPA core is a heterotrimer (Figure 2.2B and 2.2C). Whether a third protein also docks against the RMI core complex to complete a heterotrimer is not yet known.

The RMI core complex interface is distinct from RPA. The RMI interface is large, burying 33 residues of RMI1 OB2 and 36 residues of RMI2, which together corresponds to a buried surface area of 2610 Å² (computed using PISA (21)). Surface electrostatic representation of the structure reveals the complexity of the interface, as hydrophobic, basic, and acidic surfaces on RMI1 and RMI2 are all involved in complex formation (Figure 2.3). Lys121 of RMI2 has been previously posited to play a role at the interface (9, 10), and indeed it forms apparent interactions with carbonyl groups from Ile514 and Ile586 of RMI1 in the core complex structure (Figure 2.3B). A second important ionic interaction is formed between the side chains of RMI2 Asp141 and RMI1 Arg622 (Figure 2.4), both of which are in the C-terminal helices of the proteins. Other contributing residues from this helix in RMI2 are Glu140, which interacts with RMI1 Lys625, and residues His131 and Glu138, which interact with Lys511 of RMI1. Additionally, Asn128, which is adjacent to the RMI2 C-terminal α -helix interacts with the carbonyl oxygen of Val611 of RMI1. There are two other potentially important interfacing residues away from the C-terminus: the carbonyl oxygen of RMI2 Ser20 interacts with the Tyr540 side chain of RMI1,

and the main chain amide of RMI2 Arg93 interacts with the RMI1 Asp587 side chain. Finally, two highly conserved RMI2 residues that have been previously mutated and shown to disrupt complex formation (Lys24 and Trp59) were located in the center of the RMI core complex interface (Figure 2.3C) (10). Trp59 forms a hydrophobic network in the interface, stacking between Phe513 from RMI1 and the hydrophobic portion of the Lys24 side chain from RMI2 at the center of the interface. Notably, the residues that comprise the RMI core complex interface are evolutionarily well conserved (Figure 2.3C and 2.3D), indicating that the overall arrangement of subunits with the core observed in the structure is likely to be representative of most RMI complexes.

Although the overall arrangement of RMI and RPA OB domains is similar, the size and chemical composition of the interfaces are distinct. In contrast to the large RMI interface (2610 \AA^2), the corresponding RPA interface (between RPA70 DBD-C and RPA 32 DBD-D) is much smaller (900 \AA^2), with fewer amino acids from each OB fold contributing to the interaction. In addition, the RMI interface is composed of many electrostatic and hydrophobic interactions, whereas the RPA interface is dominated by hydrophobic contacts (16). Furthermore, the residues that make up the RPA trimerization core are primarily from the C-terminal α -helices that extend from the RPA OB domains, whereas those composing the RMI core complex are from both the OB domains and the C-terminal helices.

Coimmunoprecipitation experiments demonstrate the importance of RMI core complex interface. Coimmunoprecipitation experiments were used to test the importance of the

identified interface of the RMI core complex in assembly of the BLM dissolvasome. Based on the structural insights described above, expression constructs were produced that encode FLAG-epitope labeled single-site variants of RMI2 that are predicted to destabilize the complex. These RMI2 variants were then transiently transfected into HeLa cells and the integrity of the dissolvasome was measured by immunoprecipitating the RMI2 variants and probing for the presence of other dissolvasome members (BLM, Top3 α , and RMI1) by Western analysis.

Initial experiments tested RMI2 variants in which Lys121, His131, Glu138, or Asp141 were individually changed to alanine. In agreement with earlier studies, BLM, Top3 α , and RMI1 coimmunoprecipitate with wild-type FLAG-RMI2 (9) (Figure 2.5A). Interestingly, when the RMI2 alanine variants were used in the assay, only the Asp141Ala variant showed a major defect in coimmunoprecipitation with BLM, Top3 α and RMI1. Given the large interface in the RMI core complex structure, we reasoned that alanine-substitution variants may not be sufficient to disrupt the interface in all cases. To test this idea, charge reversal RMI2 variants Lys121Glu, Asp141Lys, and a variant that included both changes were tested in the coimmunoprecipitation assay. With these more dramatic changes, coimmunoprecipitation was either greatly diminished (Lys121Glu RMI2 variant) or entirely eliminated (Asp141Lys and double-reversal variants) (Figure 2.5A). These results thus support a structural model in which Lys121 and Asp141 are key residues at the interface between RMI1 and RMI2.

Disrupting the RMI core complex interface increases SCE formation *in vivo*. The consequences of destabilizing the RMI core complex interface on dissolvasome function *in vivo*

were tested next. Disruption of the *RMI2* gene in chicken DT40 cells leads to a ~5-fold increase in SCE formation that can be suppressed by expressing human RMI2 protein in the mutant cells (9). As was observed in HeLa cells, experiments in DT40 cells verified that immunoprecipitation of the human RMI2 protein coprecipitated a member of the dissolvasome (Top3 α) whereas RMI core complex-destabilizing RMI2 variants failed to do so (Figure 2.5B). These controls allow the use of the DT40 system to analyze the chromosomal effects of the RMI2 interface variants.

Assays were next performed to test the effects of RMI2 variants on SCE levels *in vivo*. As was observed previously (9), the mean number of SCE per cell increased ~6-fold in the *rmi2*^{-/-} cells compared to that of the wild type DT40 cells, but this difference was suppressed by expression of the FLAG-tagged human RMI2 protein (Figure 2.6). Consistent with retention of the RMI subcomplex, the His131Ala RMI2 variant showed no significant increase in the mean number of SCEs/cell when compared to wild type RMI2. Similarly, the Glu138Ala RMI2 variant, which also retained complex integrity, showed only a slight increase in mean SCEs/cell (less than 2-fold). In contrast, drastic changes in the mean number of SCEs/cell were observed with the Lys121 and Asp141 RMI2 variants. The Asp141Ala and Asp141Lys variants had 2.8- and 8.6-fold increases, respectively. The Lys121Ala, Lys121Glu, and Lys121Glu/Asp141Lys mutations all had 5- to 6-fold increases in mean SCEs/cell relative to wild type RMI2-complemented experiments. These values are very similar to the SCE levels of *rmi2*^{-/-} cells. With one exception (Lys121Ala), these data parallel the coimmunoprecipitation results; RMI2 variants that failed to assemble with the dissolvasome led to increases in cellular SCE events. Folding defects

or aberrant RMI core complex formation by the Lys121Ala RMI2 variant were ruled out as possible explanations for its unusual behavior. The circular dichroic spectrum of the isolated Lys121Ala RMI2 variant showed that the protein was folded, and an X-ray crystal structure of the variant in complex with the RMI1 OB2 domain was nearly indistinguishable from that of the wild type RMI core complex except for the Lys-to-Ala alteration (Figure 2.7). While the possibility still exists that there may be differences in thermodynamic stability between the two proteins, the effects of the Lys121Ala RMI2 variant appear to indicate a role for the Lys121 residue beyond RMI complex stabilization.

Discussion

RMI1 and RMI2 define a subcomplex that is an integral component of the BLM dissolvasome (9, 10, 13). The RMI subcomplex is stabilized by binding of the C-terminal OB fold of RMI1 (OB2) to RMI2, which itself comprises an OB fold (OB3) (Figure 2.2A) (9, 10). Here we report the 1.55 Å-resolution X-ray crystal structure of the RMI core complex, revealing a large interface that mediates RMI1/RMI2 heterodimer formation. This interface has been verified using coimmunoprecipitation experiments in HeLa and DT40 cells. Furthermore, disruption of the RMI interface was shown to be linked to a profound increase of SCEs in DT40 cells. Our data support a model in which the RMI core complex interface is critical for stabilizing the BLM dissolvasome, because RMI2 destabilizing mutations at the RMI interface lead to increased SCE levels similar to those observed in cells lacking RMI2 altogether. Additional possible roles for the RMI core complex domains in higher order protein complex assembly are also supported as described below.

Comparison of the RMI core complex and the RPA trimerization core. It has been suggested previously that the RMI OB domains would resemble those from RPA (9). Consistent with this hypothesis, the overall structures of the RMI OB domains are strikingly similar to those of the RPA OB domains, and the two complexes assemble using analogous surfaces (Figure 2.2). Interestingly, in spite of this overall similarity, the RMI core complex and the RPA trimerization core interfaces also display significant differences. First, while the RPA interface is primarily hydrophobic in nature, the RMI interface consists of a network of hydrogen bonds and salt bridges in addition to hydrophobic interactions. Second, while nearly all of the interactions

within the RPA interface are mediated by residues in the α -helices C-terminal to the OB domains, significant contacts between the RMI core subunits are contributed both by the OB core domains and the C-terminal helices. This difference extends the buried surface area in the RMI core complex relative to that of RPA. These distinctions could contribute to the particular functions of the two complexes.

One important functional difference between RPA and RMI is that the RMI subcomplex lacks DNA binding activity. Previous work has shown that isolated human and yeast RMI1 bind DNA weakly *in vitro* (12, 22-24), but that the RMI subcomplex is unable to bind DNA (9). Accordingly, the RMI core complex lacks analogous DNA-binding residues found in RPA, as the surface-exposed aromatic residues that might stack with ssDNA bases are buried in the RMI core complex interface. Exposure of these interface residues in the isolated RMI1 could explain its previously observed weak DNA binding (23).

Perhaps the most noticeable difference between the RMI and RPA core structures is the lack of a third partner in the RMI subcomplex. RPA forms a heterotrimeric complex that, in addition to the RPA70 and RPA32, includes RPA14 (Figure 2.2C). In contrast, the RMI core complex lacks a third partner to complete a possible trimeric structure. Since the RMI subcomplex further associates with other proteins, it is tantalizing to speculate that another protein could dock onto the RMI core structure in a similar manner to that of RPA14 binding RPA70/RPA32. Possibilities include BLM or Top3 α , which comprise the remaining dissolvasome members. However these proteins have been shown to associate with an N-terminal OB domain (OB1)

fragment of RMI1 and in the absence of RMI2 (23). A second possibility is that the RMI1 OB1 domain or some other OB-fold-containing protein might bind to the RMI core complex; such an interaction could potentially regulate dissolvasome complex formation or activation. A final, non-exclusive possibility is that other proteins could form complexes with the dissolvasome by docking onto the “open” site on the RMI core complex. These could include FANCM, for example, which has recently been shown to be able to bind to the RMI1 OB2 domain (25). Future experiments will be required to determine whether the heterodimeric RMI core complex described here is able to associate with an additional protein partner (or partners) in an RPA-like assembly.

Importance of RMI subcomplex interface in suppressing SCE levels. Based on the RMI core complex structure, RMI2 variants that have lost the ability to assemble with RMI1 and the rest of the dissolvasome were shown to lead to an increase in SCE levels *in vivo*. RMI2 variants Asp141Ala, Lys121Glu, Asp141Lys, and Lys121Glu/Asp141Lys either fail to assemble with the dissolvasome or do so with diminished stability (Figure 2.5) and lead to 2.8- to 8.6-fold increase in mean SCE events/cell compared to wild-type RMI2 (Figure 2.6). In contrast, RMI2 variants His131Ala and Glu138Ala retain the ability to assemble with the dissolvasome and have more modest increases in mean SCE events/cell of 1.2- and 1.8-fold, respectively. These data thus define a relationship between destabilization of the RMI subcomplex and an increase in cellular SCE events, and indicate a role for the RMI proteins in stabilization and/or activation of the dissolvasome within cells.

The Lys121Ala RMI2 variant presented an intriguing exception to the parallel between RMI subcomplex destabilization and elevated SCE levels. This variant retained a normal ability to assemble into the dissolvasome (Figure 2.5) but had mean SCE/cell levels that were nearly the same as the *rmi2^{-/-}* control (Figure 2.6). Indeed, even the Lys121Glu charge reversal variant retained a weakened ability to assemble within the dissolvasome. Similar observations for RMI2 Lys121Ala variants have been noted in the past (9, 10). What could account for this behavior? One possibility is that the Lys121 RMI2 variants are misfolded or that they form aberrant complexes with RMI1 that have an altered interface. However solution studies and a crystal structure of the Lys121Ala RMI2 in complex with RMI1 OB2 domain rule out this explanation, at least *in vitro* (Figure 2.7). Another possibility is that the observed difference is caused by a change in the rates of association or dissociation of the RMI core complex that cannot be detected by the coimmunoprecipitation assay. It is possible that the dissolvasome could be regulated by this key interaction through the dynamics of association instead of a constant equilibrium state. A final hypothesis is that Lys121 from RMI2 is important for functions beyond RMI complex stabilization. Notably, RMI2 Lys121 is positioned on the periphery of the RMI1/RMI2 interface and its interaction with two carbonyl groups in RMI1 appears to contort the Lys121 side chain away from the surrounding solvent (Figure 2.7). It is possible that in the context of the full dissolvasome Lys121 assumes an alternate conformation in which its charged group is presented on the surface of the RMI subcomplex to foster interactions with other cellular binding partners. Such a possibility could explain the dispensability of RMI2 Lys121 for dissolvasome formation but its clear importance in cellular activity. Furthermore, the Lys121 surface is on the opposite side of the RMI core complex from the putative RPA-like interaction

site, which could indicate that the complex has two protein interaction sites and could possibly act as a hub for docking multiple protein partners.

Concluding Remarks. The structure of the RMI core complex and analysis of RMI protein variants has revealed the importance of RMI subcomplex integrity in genome stability. It is interesting that not only are RMI and RPA both complexes that interact via OB folds, but both complexes also mediate BLM function. RPA physically interacts with BLM and stimulates its ability to unwind DNA (26). Furthermore, there are other OB-fold-containing complexes with roles in genome maintenance such as the CST complex and TPP-POT1, both of which act at telomeres. The CST complex has been proposed to form an RPA-like complex specifically at telomeres (Miyake et al., 2009; Surovtseva et al., 2009), and a recent *S. pombe* Stn1-Ten1 structure has also identified a similarity to the RPA interface (27). This and other studies point to a growing number of genome maintenance roles for OB-fold-containing proteins that extend beyond the ssDNA-binding activities for which the domain is named.

Experimental Procedures

Plasmids, Protein Overexpression, and Purification

Purification of the RMI subcomplex. A plasmid pCDFDuet-1 (Novagen) encoding RMI1 and RMI2 (9) was transformed into Rosetta 2 (DE3) *E. coli* transformed with pLysS (Novagen). Cells were grown at 30°C in Luria Broth supplemented with 25 µg/ml streptomycin and 50 µg/ml chloramphenicol. At an OD₆₀₀ of ~0.6 cells were induced to overexpress protein with 0.3 mM isopropyl β-D-thiogalactopyranoside at ~25°C for 3 hours. Cells were harvested with centrifugation, frozen at -80°C, and resuspended at 4°C in PBS (10 mM phosphate, pH 7.4, 138 mM NaCl and 2.7 mM KCl) with 5% glycerol, 1 mM phenylmethanesulfonylfluoride (PMSF), and protease-inhibitor tablets (Roche). All subsequent purification steps were performed at 4°C. Cells were lysed by sonication and centrifuged for 30 min at 15,000 rpm. Ammonium sulfate (35% w/v) was added to the soluble lysate and the resulting pellet was resuspended in PBS with 5% glycerol, then incubated with Ni-NTA resin (Qiagen) for 45 min. The resin was poured into a column and washed twice with PBS containing 10 mM imidazole and 500 mM NaCl. The RMI subcomplex eluted in 300 mM imidazole in TBS buffer (50 mM Tris-HCl, pH 7.5, 5% glycerol, 150 mM NaCl). The eluate was diluted 1:5 in Q buffer (25 mM Tris-HCl, pH 7.5, 5% glycerol) and loaded onto a HiPrep Q FF column (Amersham). The bound protein was eluted with a linear (0.1 M to 0.5 M) NaCl gradient. The peak fractions were used for limited proteolysis.

Purification of the RMI core complex. Open-reading frames encoding RMI1 OB2 and RMI2 (or an RMI2 variant) were separately subcloned from the above plasmid into a modified

pGEX4T-1 plasmid (Amersham) that expresses an amino-terminal glutathione S-transferase (GST) tagged RMI1 OB2 and amino-terminal 6x-histidine tagged RMI2. This plasmid was transformed into Rosetta 2 (DE3) *E. coli* cells transformed with pLysS (Novagen) and grown at 30°C in LB medium supplemented with 100 µg/ml ampicillin and 50 µg/ml chloramphenicol. At an OD₆₀₀ of ~0.6 cells were induced to overexpress protein with 0.5 mM isopropyl β-D-thiogalactopyranoside at 30°C for 3 hours. Cells were harvested with centrifugation, frozen at -80°C, and resuspended at 4°C in lysis buffer (50 mM Tris-HCl pH 7.5, 20% sucrose, and 600 mM KCl) supplemented with 1 mM PMSF and protease-inhibitor tablets (Roche). All subsequent purification steps were performed at 4°C. Cells were lysed by sonication and centrifuged for 30 min at 15,000 rpm. The soluble lysate was incubated with Ni-NTA resin (Qiagen) for 30 min, then poured into a column and washed twice with wash buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM imidazole, 100 mM dextrose, and 10% glycerol). The complex was eluted with 300 mM imidazole in TBS buffer, then incubated with Glutathione Sepharose 4B resin (Amersham) for 40 min. The resin was then poured into a column, washed twice with GS4B buffer (25 mM Tris-HCl pH 7.5, 0.5 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 10% glycerol) with 500 mM KCl, and eluted with GS4B buffer supplemented with 150 mM KCl and 20 mM reduced glutathione. The eluate was incubated with thrombin overnight to remove both affinity tags, leaving Gly-Ser-His peptides on the amino terminus of both proteins. The eluate was dialyzed against SPFF buffer (20 mM MES pH 6.0, 50 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 10 % glycerol) and then loaded on a HiPrep SP FF column (Amersham). The bound protein was eluted with a linear (0.05 M to 0.50 M) NaCl gradient. Fractions containing the complex were pooled, concentrated to <2 ml, then

further purified on a Sephacryl S-100 column in SPFF buffer. Selenomethionine-incorporated protein was expressed as previously described (28) and purified identically to native protein. Analytical size-exclusion chromatography was performed on with a Sephacryl S-200 column in SPFF buffer.

Limited Proteolysis Assay. The purified RMI subcomplex was incubated at a 150:1 molar ratio (complex:protease) with subtilisin, trypsin, or α -chymotrypsin in 25 mM Tris-HCl pH 8, 8 mM DTT, 1 mM EDTA, 150 mM NaCl, and 10% glycerol. Proteolysis was terminated after 1, 5, 15, and 30 minutes by adding 2X sample buffer (12.5 mM Tris-HCl pH 6.8, 20% glycerol, 5% w/v SDS, 10% β -mercaptoethanol, bromophenol blue) and freezing on dry ice. The products were analyzed by 15% SDS-PAGE and mass spectrometry at the UW Biotechnology Center.

Protein Crystallization and Structure Determination. Wild type RMI core complex or a RMI core complex formed with Lys121Ala RMI2 was dialyzed against water. Selenomethionine-incorporated RMI core complex was dialyzed against 10 mM HEPES pH 7.1, 1 mM DTT. Core complex samples (5 mg/ml) were mixed with mother liquor (10 mM Hepes pH 7.1, 0.05 M NaSCN, 20% ethylene glycol, 2 mM DTT, and 13% polyethylene glycol 3350) at a 1:1 (vol) ratio, and benzamidine (1%) was added to the solution. Crystals were formed by hanging drop vapor diffusion. Crystals were transferred to a cryoprotectant solution (10mM Hepes pH 7.1, 0.05 M NaSCN, 30% ethylene glycol, 2mM DTT, 15% polyethylene glycol 3350, 1% benzamidine) and flash-frozen in liquid nitrogen. The highest quality crystals were obtained by microseeding with native crystals.

The structure of the selenomethionine-incorporated RMI core complex was determined by single-wavelength anomalous dispersion (SAD) phasing. Data were indexed and scaled using HKL2000 (29). Selenium positions were identified using SHELX within Crank (30) and an initial structure was built using ARP/wARP (31). The structure was improved by rounds of manual fitting using Coot (32) and refinement against the native data set using REFMAC5 (33). The structure of the RMI core complex that includes RMI2 Lys121Ala variant was determined by molecular replacement using Phaser (34). The two proteins in the RMI core complex were positioned as individual search models so that any differences in packing would be revealed. Refinement was limited to positioning the RMI core complex subunits. Coordinate and structure factor files have been deposited in the Protein Data Bank (PDB ID 3MXN).

Coimmunoprecipitation and SCE assays. FLAG-tagged RMI2 variants were transiently expressed in HeLa cells or stably expressed in chicken DT40 *rmi2*^{-/-} cells. Nuclear extract was prepared as described previously (35) for coimmunoprecipitation. Coimmunoprecipitation experiments were performed using the anti-FLAG M2-agarose (Sigma) as described (35). Polyclonal antibodies for Western blotting against BLM, Topo 3 α , RMI1 or RMI2 were described elsewhere (9, 13, 36, 37). The different lanes were normalized by comparison of FLAG-RMI2 levels. The SCE assay was performed as described previously (35).

Acknowledgements

We thank Advanced Photon Source staff (LS-CAT beamline) for assistance with data collection, and Dr. David Schlessinger, James Berger, David Fox, and members of the Keck Lab for critical reading of this manuscript. This work was funded by a grant from the NIH (GM068061, JLK) and by the Intramural Research Program of the National Institute on Aging (Z01 AG000657-08, WW). KAH was supported in part by an NIH training grant in Molecular Biosciences (GM07215).

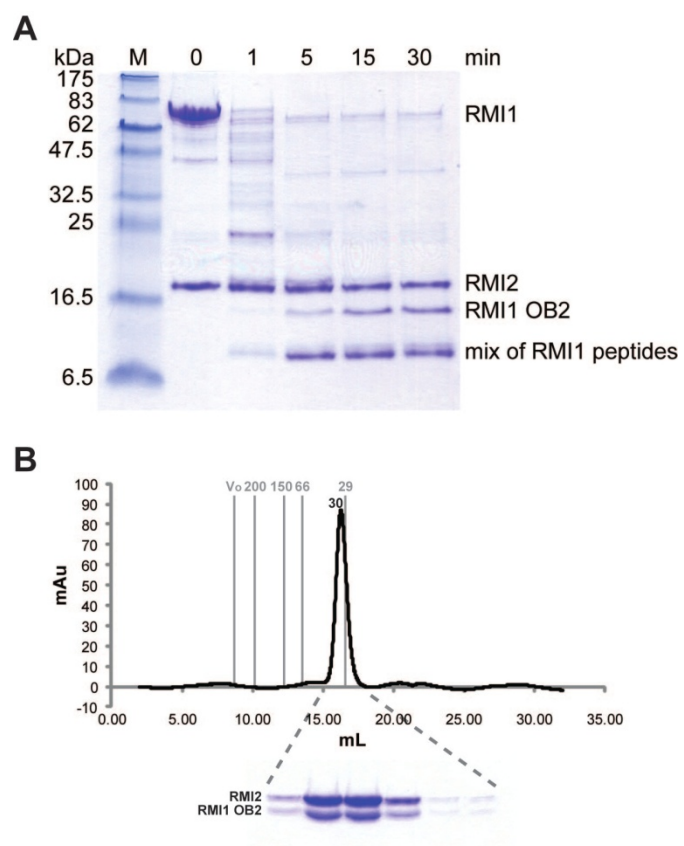
Figure 2.1

Figure 2.1. Characterizing the RMI core complex. **(A)** Limited proteolysis of the RMI subcomplex. Full-length RMI subcomplex was incubated with trypsin (1:150 trypsin:RMI subcomplex, molar ratio) for indicated times and samples were resolved by 15% SDS-PAGE. M denotes pre-stained protein markers with their molecular masses (kDa) shown on the left side of the gel. The RMI1 breakdown product identified in the study (RMI1 OB2) is labeled on the right side of the gel along with full-length RMI1 and RMI2. **(B)** Size exclusion chromatographic analysis of the RMI core complex. The chromatograph from the size exclusion column is shown in black with grey lines indicating the elution volumes of molecular weight standards resolved under the same conditions. The molecular masses (kDa) of the standards are indicated above the lines with the predicted molecular weight above the RMI core complex elution. Dashed lines indicate individual fractions that were resolved on an SDS-PAGE gel to show the purified RMI core complex. Size exclusion analysis predicts a molecular weight of 30.4 kDa for the RMI core complex, indicating 1:1 complex formation.

Figure 2.2.

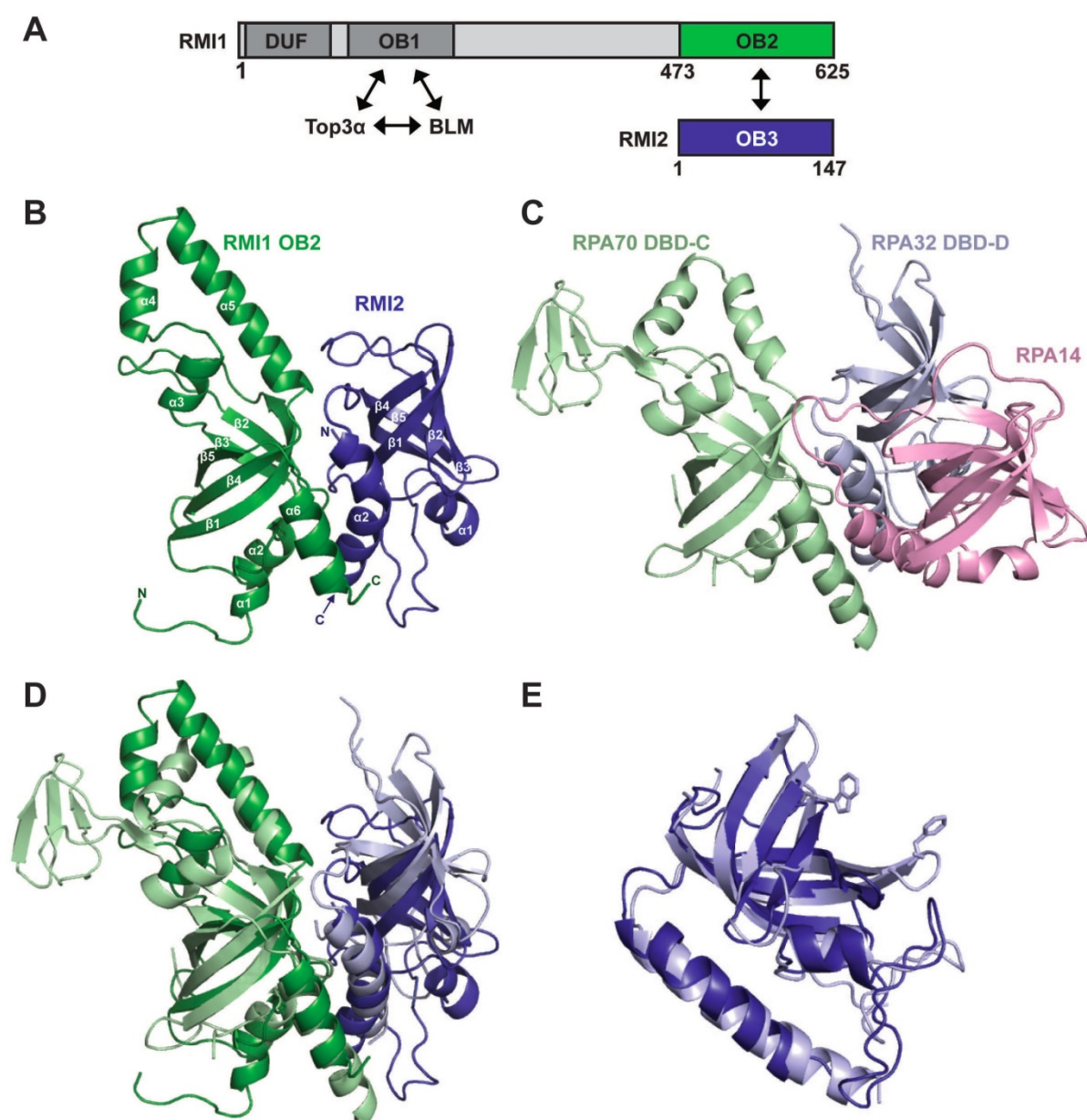


Figure 2.2. Structural features of the RMI core complex. **(A)** Schematic diagram of the domain structures of human RMI1 and RMI2. Interacting OB domains comprising the RMI core complex are shown in green (RMI1 OB2) and blue (RMI2 OB3). Additional RMI1 domains (domain of unknown function, DUF, and an OB domain that interacts with Top3 α and BLM (OB1)) are indicated. **(B)** Ribbon diagram of the crystal structure of the RMI core complex. RMI1 OB2 (green, residues 475-625), and RMI2 (blue, residues 17-147) are rendered using PyMol (38). **(C)** Ribbon diagram of the crystal structure of the RPA trimerization core (16) shown with RPA70 in the same orientation as RMI1 OB2 in (B). RPA70 DBD-C (pale green), RPA32 DBD-D (pale blue), and RPA14 (pink) are shown. **(D)** Superposition of RMI1 OB2 with RPA70 DBD-C. The color scheme and orientation are presented as in (B) and (C). **(E)** Superposition of RMI2 with RPA32 DBD-D (39). DNA-stacking side chains are shown for RPA32 DBD-D. There are no analogous aromatic residues in RMI2.

Figure 2.3.

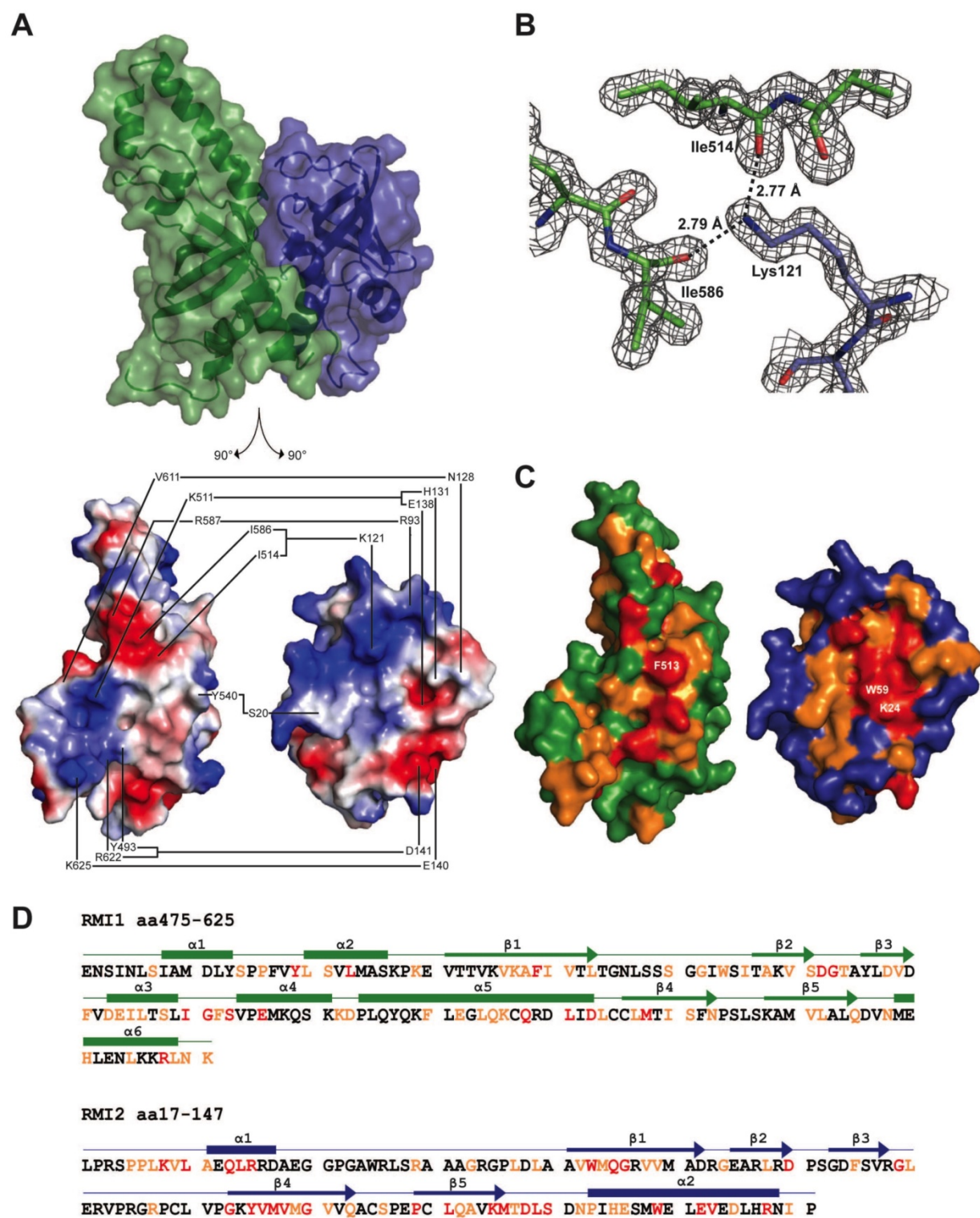


Figure 2.3. RMI core complex interface. (A) Top: structure of the RMI core complex showing semi-transparent surface and ribbon representations in the same orientation and color scheme as in 1B. Bottom: split view of the RMI core complex in which both proteins are separated and rotated 90° to show interacting surfaces. The interface between RMI1 and RMI2 is shown with connected lines indicating residues forming hydrogen bonds or ionic interactions. The interface is depicted using surface electrostatics, showing electropositive (blue) and electronegative (red) surface potential. (B) $2F_o - F_c$ electron density map (1.5σ) showing interaction of RMI1 Ile514 and Ile586 with RMI2 Lys121. (C) The interface between RMI1 and RMI2 is shown in the same orientation as 2A, bottom. RMI1 and RMI2 are colored to indicate residues that are invariant (red), highly conserved (orange), or poorly conserved (green for RMI1 and blue for RMI2) among identified RMI1 and RMI2 proteins from human, bovine, mouse, zebrafish, and arabidopsis sequences. Three of the invariant residues that form a hydrophobic interaction network are labeled: RMI1 Phe513, RMI2 Lys24, and RMI2 Trp59. (D) The sequence for the RMI core complex is shown, with secondary structural features labeled above as boxes (α -helices) and arrows (β -strands). Red and orange lettering indicates invariant and highly conserved residues, respectively.

Figure 2.4.

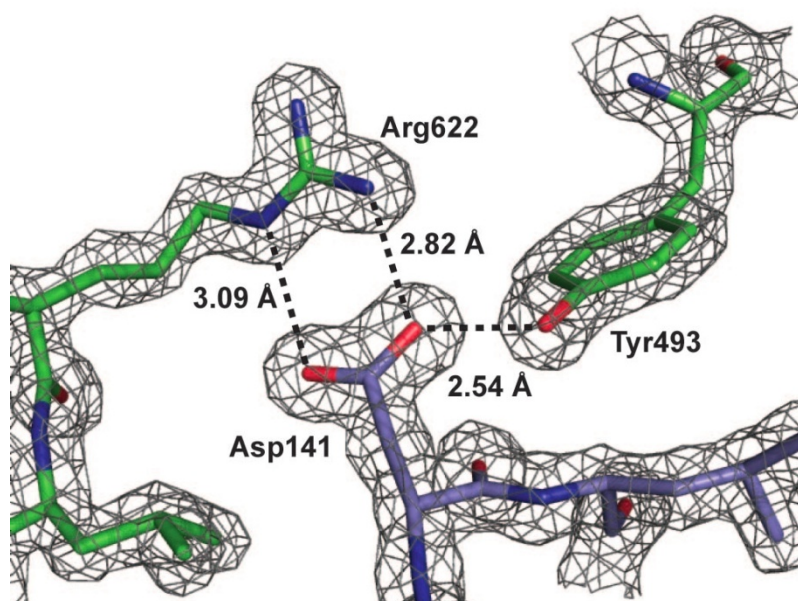


Figure 2.4. Electron density of the Asp141 interaction. $2F_o - F_c$ electron density map (1.5σ) showing interaction of RMI1 Tyr493 and Arg622 with RMI2 Asp141.

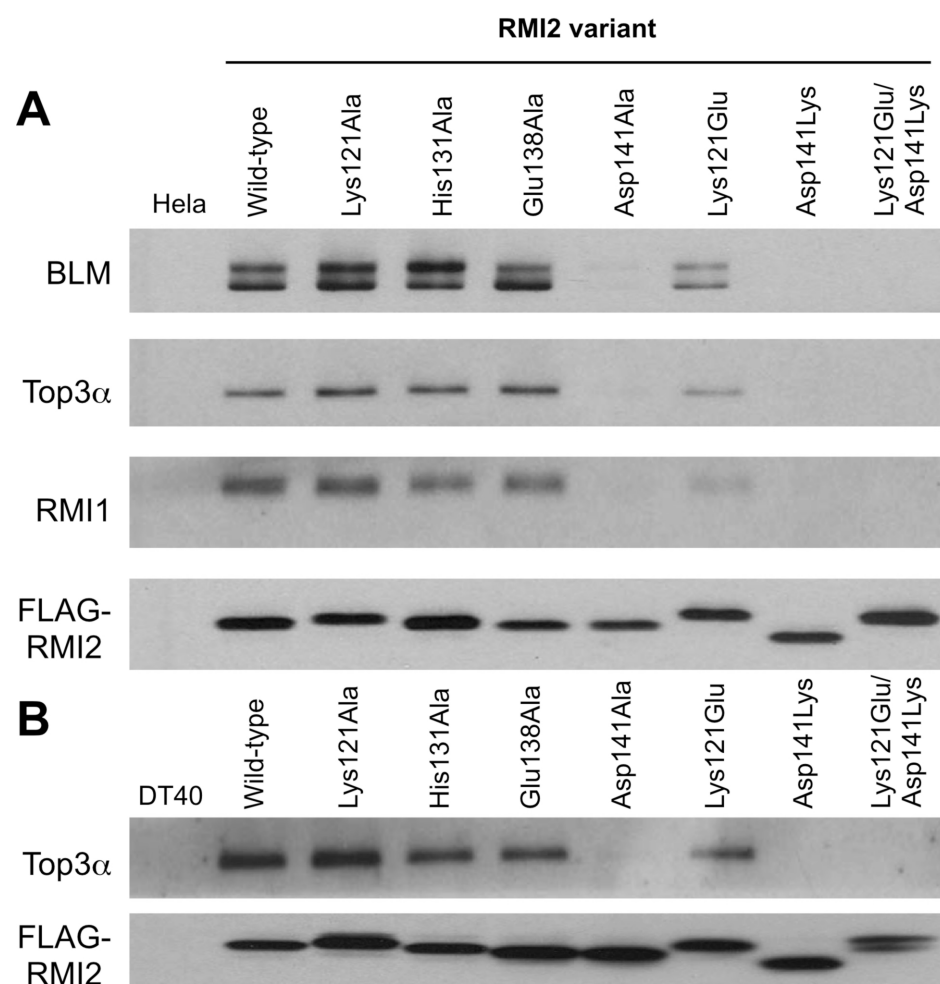
Figure 2.5.

Figure 2.5. Coimmunoprecipitations of the dissolvasome complex components with RMI2 variants. **(A)** Western analysis of dissolvasome complex components (BLM, Top3 α , RMI1, and FLAG-RMI2) from immunoprecipitates of FLAG-tagged RMI2 transiently expressed in HeLa cells. **(B)** Western analysis of dissolvasome complex components (Top3 α and FLAG-RMI2) from immunoprecipitates of FLAG-tagged RMI2 stably expressed in chicken DT40 cells. RMI1 and BLM were not detected in DT40 cells because the antibodies raised against human proteins fail to recognize the chicken homologs.

Figure 2.6.

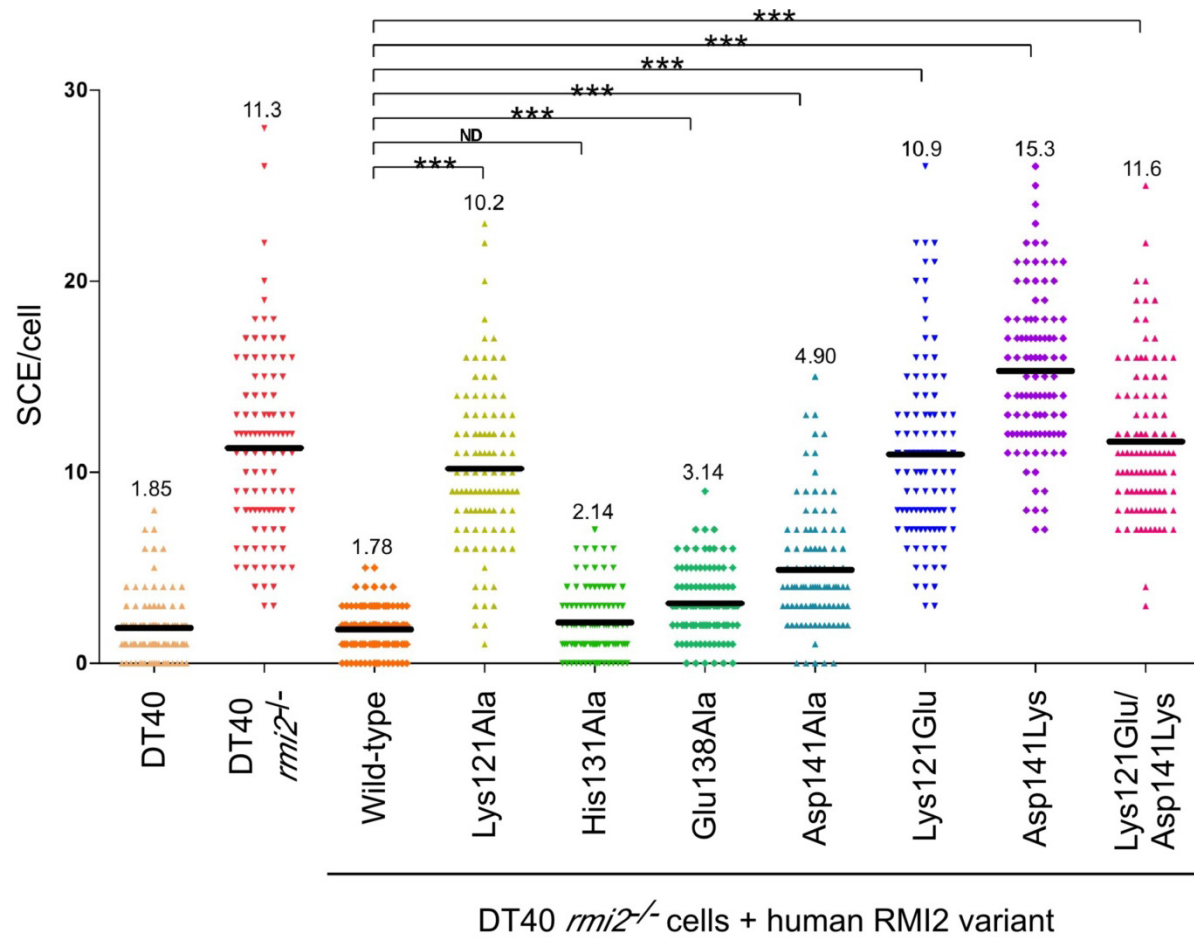


Figure 2.6. SCE frequencies observed with different RMI2 variants. The number of SCEs per cell in chicken DT40 cells are shown as a histogram in which each point in a column is the number of SCE events in a single cell. The mean level of SCEs/cell is displayed as a dark black line and numerically at the top of the column. Statistical significance is indicated with *** ($p < 0.001$) and ND ($p > 0.05$). “DT40” denotes wild type DT40 cells, “*rmi2*^{-/-}” denotes DT40 cells lacking RMI2, “wild-type” denotes *rmi2*^{-/-} cells complemented with human RMI2 and the remainder denote the human RMI2 variants that have been substituted.

Figure 2.7.

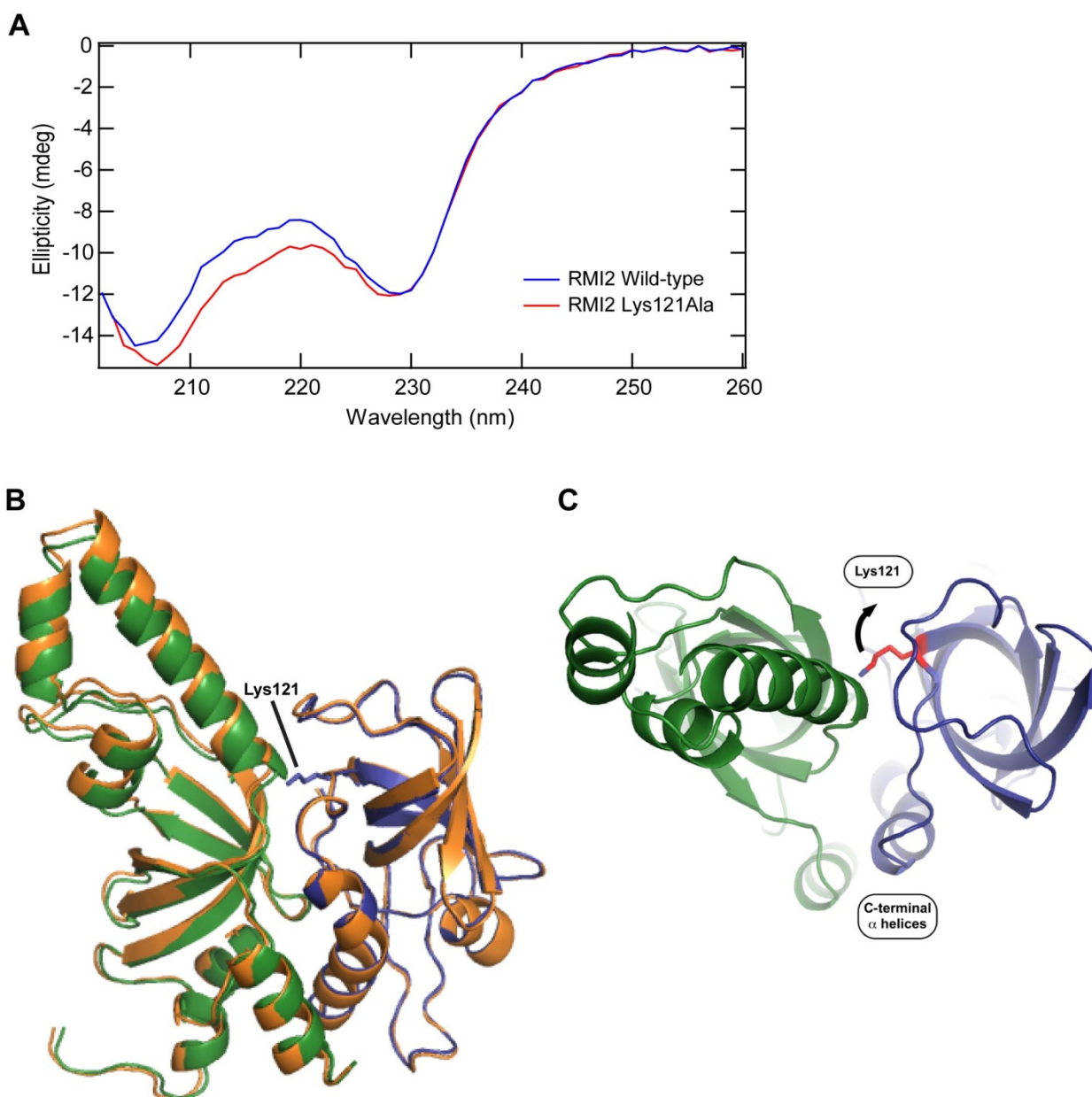


Figure 2.7. Structure of the RMI2 Lys121Ala variant and surface proximity of Lys121. (A)

Circular dichroic (CD) spectra showing recombinant wild type RMI2 (blue) and RMI2 Lys121Ala (red). CD was performed at 4° C with proteins at 300 µg/ml in 10 mM Tris-HCl, pH 9, 100 mM NaCl, 3% glycerol. Thermal denaturation experiments resulted in protein aggregation, which prevented determination of melting temperatures for RMI2 or RMI2 Lys121Ala. **(B)** Overlay of the wild type RMI core complex (blue and green) with the X-ray crystal structure of the core complex with the RMI2 Lys121Ala mutation. RMI2 from both structures was superimposed to observe any changes in the RMI1 OB domain relative position. The overall rmsd including all C_α atoms from both RMI1 and RMI2 was 0.56 Å. Lys121 from the wild type RMI structure is shown to highlight its location near the surface of the complex. **(C)** Close-up image showing the proximity of RMI2 Lys121 (colored red and blue) to the surface of the RMI core complex. The arrow indicates that alternative Lys121 rotamers could extend the side chain into the solvent.

Table 2.1. X-ray data collection and structure determination statistics

	Native	SeMet
Data Collection		
Wavelength, Å	0.97856	0.97933
Resolution Range (high resolution bin), Å	30-1.55 (1.58-1.55)	50-2.14 (2.18-2.14)
Space Group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit Cell (a, b, c (Å))	42.10, 42.22, 152.24	41.92, 42.03, 152.87
(α , β , γ (°))	90, 90, 90	90, 90, 90
Completeness, %	97.0 (83.0)	99.9 (99.6)
Total/Unique Reflections	423,921/39,281	213,642/15,791
Redundancy	10.8 (5.3)	13.5 (12.6)
$\langle I/\sigma I \rangle$	25.7 (2.4)	21.5 (20.2)
R_{sym}^{\dagger} , %	11.5 (37.8)	10.0 (20.8)
Phasing		
Figure of Merit	0.756	
Refinement		
Resolution, Å	30-1.55	
$R_{\text{work}}/R_{\text{free}}^{\ddagger}$, %	19.5/23.1	
Rms deviations		
Bonds, Å	0.008	
Angles, °	1.20	
Ramachandran statistics, %		
Most favored	93.8	
Allowed	5.8	
Generously allowed	0.4	
Disallowed	0.0	
# atoms		
protein	2269	
solvent	259	
$\langle B \text{ factor} \rangle$, Å ²		
protein	17.4	
solvent	28.9	

[†] $R_{\text{sym}} = \frac{\sum_j |I_j - \langle I \rangle|}{\sum_j I_j}$, where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

[‡] $R_{\text{work}}/R_{\text{free}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$, where the working and free R factors are calculated by using the working and free reflection sets, respectively. The free R reflections (5% of the total) were held aside throughout refinement.

References

1. German J (1993) Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine (Baltimore)* 72(6):393-406.
2. Bachrati CZ & Hickson ID (2003) RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem J* 374(Pt 3):577-606.
3. Ouyang KJ, Woo LL, & Ellis NA (2008) Homologous recombination and maintenance of genome integrity: cancer and aging through the prism of human RecQ helicases. *Mech Ageing Dev* 129(7-8):425-440.
4. Bachrati CZ, Borts RH, & Hickson ID (2006) Mobile D-loops are a preferred substrate for the Bloom's syndrome helicase. *Nucleic Acids Res* 34(8):2269-2279.
5. Karow JK, Constantinou A, Li JL, West SC, & Hickson ID (2000) The Bloom's syndrome gene product promotes branch migration of holliday junctions. *Proc Natl Acad Sci U S A* 97(12):6504-6508.
6. van Brabant AJ, Ye T, Sanz M, German IJ, Ellis NA, & Holloman WK (2000) Binding and melting of D-loops by the Bloom syndrome helicase. *Biochemistry* 39(47):14617-14625.
7. Bugreev DV, Yu X, Egelman EH, & Mazin AV (2007) Novel pro- and anti-recombination activities of the Bloom's syndrome helicase. *Genes Dev* 21(23):3085-3094.
8. Ralf C, Hickson ID, & Wu L (2006) The Bloom's syndrome helicase can promote the regression of a model replication fork. *J Biol Chem* 281(32):22839-22846.
9. Xu D, Guo R, Sobeck A, Bachrati CZ, Yang J, Enomoto T, Brown GW, Hoatlin ME, Hickson ID, & Wang W (2008) RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. *Genes Dev* 22(20):2843-2855.
10. Singh TR, Ali AM, Busygina V, Raynard S, Fan Q, Du CH, Andreassen PR, Sung P, & Meetei AR (2008) BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvasome. *Genes Dev* 22(20):2856-2868.
11. Raynard S, Bussen W, & Sung P (2006) A double Holliday junction dissolvasome comprising BLM, topoisomerase IIIalpha, and BLAP75. *J Biol Chem* 281(20):13861-13864.
12. Wu L, Bachrati CZ, Ou J, Xu C, Yin J, Chang M, Wang W, Li L, Brown GW, & Hickson ID (2006) BLAP75/RMI1 promotes the BLM-dependent dissolution of homologous recombination intermediates. *Proc Natl Acad Sci U S A* 103(11):4068-4073.

13. Yin J, Sobeck A, Xu C, Meetei AR, Hoatlin M, Li L, & Wang W (2005) BLAP75, an essential component of Bloom's syndrome protein complexes that maintain genome integrity. *Embo J* 24(7):1465-1476.
14. Wu L & Hickson ID (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* 426(6968):870-874.
15. Bochkarev A & Bochkareva E (2004) From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol* 14(1):36-42.
16. Bochkareva E, Korolev S, Lees-Miller SP, & Bochkarev A (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *Embo J* 21(7):1855-1863.
17. Murzin AG (1993) Ob(Oligonucleotide Oligosaccharide Binding)-Fold - Common Structural and Functional Solution for Nonhomologous Sequences. *Embo Journal* 12(3):861-867.
18. Holm L, Kaariainen S, Rosenstrom P, & Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* 24(23):2780-2781.
19. Lao Y, Lee CG, & Wold MS (1999) Replication protein A interactions with DNA. 2. Characterization of double-stranded DNA-binding/helix-destabilization activities and the role of the zinc-finger domain in DNA interactions. *Biochemistry* 38(13):3974-3984.
20. Pestryakov PE, Weissbart K, Schlott B, Khodyreva SN, Kremmer E, Grosse F, Lavrik OI, & Nasheuer HP (2003) Human replication protein A. The C-terminal RPA70 and the central RPA32 domains are involved in the interactions with the 3'-end of a primer-template DNA. *J Biol Chem* 278(19):17515-17524.
21. Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774-797.
22. Mullen JR, Nallasetth FS, Lan YQ, Slagle CE, & Brill SJ (2005) Yeast Rmi1/Nce4 controls genome stability as a subunit of the Sgs1-Top3 complex. *Mol Cell Biol* 25(11):4476-4487.
23. Raynard S, Zhao W, Bussen W, Lu L, Ding YY, Busygina V, Meetei AR, & Sung P (2008) Functional role of BLAP75 in BLM-topoisomerase IIIalpha-dependent holliday junction processing. *J Biol Chem* 283(23):15701-15708.
24. Chen CF & Brill SJ (2007) Binding and activation of DNA topoisomerase III by the Rmi1 subunit. *J Biol Chem* 282(39):28971-28979.
25. Deans AJ & West SC (2009) FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. *Mol Cell* 36(6):943-953.

26. Brosh RM, Jr., Li JL, Kenny MK, Karow JK, Cooper MP, Kureekattil RP, Hickson ID, & Bohr VA (2000) Replication protein A physically interacts with the Bloom's syndrome protein and stimulates its helicase activity. *J Biol Chem* 275(31):23500-23508.
27. Sun J, Yu EY, Yang Y, Confer LA, Sun SH, Wan K, Lue NF, & Lei M (2009) Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes Dev* 23(24):2900-2914.
28. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, & Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* 229(1):105-124.
29. Otwinowski Z & Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology*, eds Carter CW, Jr. & Sweet RM (Academic Press, New York), Vol 276, pp 307-326.
30. Ness SR, de Graaff RA, Abrahams JP, & Pannu NS (2004) CRANK: new methods for automated macromolecular crystal structure solution. *Structure* 12(10):1753-1761.
31. Perrakis A, Morris R, & Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6(5):458-463.
32. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
33. Winn MD, Isupov MN, & Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 1):122-133.
34. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.
35. Guo R, Xu D, & Wang W (2009) Identification and analysis of new proteins involved in the DNA damage response network of Fanconi anemia and Bloom syndrome. *Methods*.
36. Meetei AR, Sechi S, Wallisch M, Yang D, Young MK, Joenje H, Hoatlin ME, & Wang W (2003) A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome. *Mol Cell Biol* 23(10):3417-3426.
37. Wu L, Davies SL, North PS, Goulaouic H, Riou JF, Turley H, Gatter KC, & Hickson ID (2000) The Bloom's syndrome gene product interacts with topoisomerase III. *J Biol Chem* 275(13):9636-9644.
38. Delano WL (2002) *The PyMol Molecular Graphics System*. (DeLano Scientific, San Carlos, CA).

39. Deng X, Habel JE, Kabaleeswaran V, Snell EH, Wold MS, & Borgstahl GE (2007) Structure of the full-length human RPA14/32 complex gives insights into the mechanism of DNA binding and complex formation. *J Mol Biol* 374(4):865-876.

CHAPTER 3

Defining the molecular interface that connects the Fanconi anemia protein FANCM to the Bloom syndrome dissolvasome

This work has been published:

Hoadley KA, Xue Y, Ling C, Takata M, Wang W, Keck JL (2012) “Defining the molecular interface that connects the Fanconi anemia protein FANCM to the Bloom syndrome dissolvasome”. *Proc Natl Acad Sci U S A* **12**, 4437-42.

Kelly A. Manthei (nee Hoadley) performed protein and peptide purifications, structure determination and fluorescence polarization experiments. Yutong Xue, Chen Ling, Minoru Takata, and Weidong Wang designed and executed the co-immunoprecipitation and sister chromatid exchange assays.

Abstract

The RMI subcomplex (RMI1/RMI2) functions with the BLM helicase and topoisomerase III α in a complex called the “dissolvasome”, which separates double-Holliday junction DNA structures that can arise during DNA repair. This activity suppresses potentially harmful sister chromatid exchange (SCE) events in wildtype cells but not in cells derived from Bloom Syndrome patients with inactivating *BLM* mutations. The RMI subcomplex also associates with FANCM, a component of the Fanconi Anemia (FA) core complex that is important for repair of stalled DNA replication forks. The RMI/FANCM interface appears to help coordinate dissolvasome and FA core complex activities but its precise role remains poorly understood. Here, we define the structure of the RMI/FANCM interface and investigate its roles in coordinating cellular DNA repair activities. The X-ray crystal structure of the RMI core complex bound to a well-conserved peptide from FANCM shows that FANCM binds to both RMI proteins through a hydrophobic “knobs-into-holes” packing arrangement. The RMI/FANCM interface is shown to be critical for interaction between the components of the dissolvasome and the FA core complex. FANCM variants that substitute alanine for key interface residues strongly destabilize the complex in solution and lead to increased SCE levels in cells that are similar to those observed in *blm*- or *fancm*-deficient cells. This study provides a molecular view of the RMI/FANCM complex and highlights a key interface utilized in coordinating the activities of two critical eukaryotic DNA-damage repair machines.

Introduction

Cells have evolved numerous corrective pathways to ensure fidelity and stability of their genomes. Deleterious mutations in genes encoding DNA repair proteins can lead to several different diseases, many of which share common features such as formation of unusual chromosomal structures in affected cells and cancer predisposition in afflicted individuals. Fanconi anemia (FA), an autosomal recessive disease caused by mutations in any of 15 genes involved in the repair of DNA inter-strand crosslinks (ICLs), is an example of such a disease (1-3). FA patients suffer developmental abnormalities, progressive bone marrow failure, and a high incidence of cancer. FA cell lines are hypersensitive to agents that induce ICLs, exhibiting high levels of chromosomal instability as exemplified by radial chromosomes or, in some cases, by high levels of sister chromatid exchange (SCE) in mitotic cells. A core complex of 8 FA proteins, (FANCA, B, C, E, F, G, L, and M), cooperate in DNA damage recognition and in catalyzing monoubiquitination of the FANCD2/FANCI heterodimer in response to DNA damage. Ubiquitinated FANCD2/FANCI in turn recruits downstream factors to facilitate DNA repair (2).

Bloom syndrome (BS), similar to FA, is an autosomal recessive DNA-repair deficiency disease that exhibits chromosomal instability. BS results from mutation of a single gene, *BLM*, which encodes one of five human RecQ-family DNA helicases (4). BS patients exhibit proportional dwarfism, sun-sensitivity, and a broad predisposition to cancer (5, 6). BS cells display elevated levels of SCEs, which is consistent with the proposed role for the BLM helicase in the non-recombinogenic repair of double Holliday junction (dHJ) DNA structures that can arise during

repair. To catalyze this activity, BLM functions within a complex called the “dissolvasome” that also includes topoisomerase III α (Top3 α) and the RecQ-mediated genome instability, or RMI, subcomplex (RMI1/RMI2 heterodimer) (7-13). The dissolvasome is thought to drive HJ branch migration (catalyzed by BLM) and DNA cleavage of the resulting hemi-catenated DNA (catalyzed by Top3 α) in a reaction called “dissolution” (14). The RMI subcomplex stimulates these activities and stabilizes the dissolvasome (7, 8, 10).

Several studies have highlighted molecular connections between the FA core complex and the BLM dissolvasome. The initial finding linking the two complexes was identification of an ICL-induced super-complex called BRAFT (BLM, Replication Protein A, FANCM, and Top3 α) that included components of both complexes (15, 16). A subsequent study identified a direct link between a conserved 34 amino acid motif from FANCM (called MM2, for FANCM Motif 2) and the C-terminal OB-fold of RMI1 (RMI1 OB2) and Top3 α (Figure 3.1A) (17). FANCM is a member of the FA core complex that contains an N-terminal helicase domain and a degenerate C-terminal endonuclease domain; the MM2 sequence resides in the linker that connect these two domains (17, 18). Consistent with a connection between FANCM and BS, FANCM-deficient cells are characterized by high levels of SCEs, which is distinct from most other FA complementation groups (19) but similar to BS-derived cells. Deletion of MM2 from FANCM or substitution of two Phe residues in the FANCM MM2 sequence (Phe1232 and Phe1236) with Ala abrogates its binding to RMI1 and results in a DNA-damage-dependent increase in SCE formation (17). These observations lead to a model in which FANCM is responsible for

coordinating the activities of the FA complex and the BLM dissolvasome at damaged sites in the genome.

X-ray crystal structures of a protease-resistant “RMI core complex” comprising RMI1 OB2 bound to RMI2 have recently been reported (12, 13). Sequence changes at the RMI1/RMI2 interface generally destabilize the dissolvasome and increase cellular SCEs levels (12). However, one unusual interface variant, RMI2 Lys121Ala, supports normal dissolvasome assembly but leads to SCE levels similar to those observed in *rmi2*^{-/-} cells (12). Based on this result and the report that MM2 variants display SCE levels that are similar to those of BS cells (17), we hypothesized that FANCM binds at the RMI1/RMI2 interface and that RMI2 Lys121 is critical for proper formation of the MM2 binding site. In support of this hypothesis, we show that RMI subcomplex variants that include RMI2 Lys121Ala are defective in binding to a FANCM MM2 peptide *in vitro* and to full-length FANCM *in vivo*. Interestingly, the Top3α/FANCM interaction is no longer observed with the Lys121Ala RMI2 sequence change, supporting a model in which the interaction between Top3α and the FA core complex is mediated by the RMI/FANCM interface. To better understand this important molecular interface, we have determined the X-ray crystal structure of the FANCM MM2 peptide bound to the RMI core complex. The structure shows that FANCM MM2 binds to both RMI1 and RMI2 at a site involving RMI2 Lys121. The complex is formed through an extensive hydrophobic “knobs-into-holes” packing arrangement in which hydrophobic side chains from FANCM MM2 project into hydrophobic pockets on the RMI core complex. These MM2 “knob” residues include Phe residues previously identified as having roles in RMI subcomplex binding (17) and

several additional residues from FANCM. Sequence changes to the hydrophobic “knob” residues in FANCM abrogate binding to the RMI core complex *in vitro* and *in vivo*, leading to increases in cellular SCE formation. Collectively, these studies define the key molecular surface that links the FA core complex to the BLM dissolvasome and allow a deeper understanding of the mechanisms that connect these two diseases.

Results

A role for RMI2 Lys121 in linking the dissolvasome to FANCM. Previous studies have shown that the C-terminal OB domain of RMI1 (RMI1 OB2) interacts both with the MM2 region of FANCM and with RMI2 (Figure 3.1A) (7, 8, 12, 13, 17). We have confirmed the RMI/FANCM interaction *in vitro* by demonstrating that a glutathione S-transferase (GST)-MM2 fusion protein forms a complex with the RMI core complex (RMI1 OB2 in complex with RMI2) using analytical size-exclusion chromatography (Figure 3.2). To assess the stability of the RMI/MM2 complex, MM2 peptide was purified, labeled with fluorescein (producing F-MM2) and used in a fluorescence polarization (FP) assay (Figure 3.1B, 3.1C). F-MM2 was bound by the RMI subcomplex (full-length RMI1 in complex with RMI2) with an apparent dissociation constant ($K_{d, app}$) of ≤ 4.5 nM. Since the F-MM2 concentration is the same as the $K_{d, app}$, we could only establish an upper limit for the binding affinity. F-MM2 binding by the RMI core complex was very similar to that observed with the full-length RMI subcomplex, indicating that the RMI core complex is sufficient for binding to MM2 (Figure 3.1C).

Previous results have shown that the RMI2 variant Lys121Ala exhibits no defect in co-immunoprecipitation with other members of the BLM dissolvasome, but shows increases in cellular SCE levels (7, 12). We therefore tested a possible role for Lys121 from RMI2 in MM2 binding using the FP assay. An RMI core complex that included the RMI2 Lys121Ala variant bound F-MM2 with a $K_{d, app}$ of 360 ± 20 nM, which is ≥ 80 -fold weaker than the wild type RMI core complex (Figure 3.1C). Since the structure of the RMI core complex is unaltered by the Lys121Ala sequence change (12), this result strongly supports a role for Lys121 in FANCM

MM2 binding. To confirm the importance of RMI2 Lys121 in the context of full-length FANCM, Flag-tagged wildtype RMI2 or RMI2 Lys121Ala were expressed in *rmi2*^{-/-} DT40 cells and immunoprecipitated; co-precipitation of FANCM was detected by Western blotting. Consistent with a role for RMI2 Lys121 in FANCM binding, FANCM co-immunoprecipitated with wildtype Flag-RMI2 but not with Flag-RMI2 Lys121Ala (Figure 3.1D). Top3 α and RMI1 co-immunoprecipitated with both Flag-tagged RMI2 proteins as has been observed previously (7, 12). These data indicate that the Lys121Ala sequence change in RMI2 specifically abrogates binding to FANCM and identify a role for RMI2 in FANCM binding.

With its selective disruption of the FANCM/dissolvasome interaction, the RMI2 Lys121Ala variant is a useful tool for determining whether FANCM interacts directly with Top3 α , as has been suggested (17), or instead requires interaction with the RMI subcomplex to associate with Top3 α and other dissolvasome components. To determine whether interaction between FANCM and Top3 α depends on the RMI subcomplex, FANCM was immunoprecipitated from DT40 cells and FANCM, RMI1, and Top3 α levels were assessed by Western blotting. Top3 α and the RMI subcomplex co-immunoprecipitated with FANCM in extract from wildtype cells but not *rmi2*^{-/-} cells (Figure 3.3), indicating that RMI2 is essential for the association between the BLM dissolvasome and FANCM. We next performed co-immunoprecipitations in *rmi2*^{-/-} cells complemented with RMI2 or the RMI2 Lys121Ala variant in order to selectively disrupt the RMI/FANCM interaction (Figure 3.1E). RMI1 and Top3 α co-immunoprecipitation with FANCM was observed in cells complemented with wildtype RMI2 but not in cells complemented with RMI2 Lys121Ala (Figure 3.1E, longer exposures of experiments failed to

reveal the presence of RMI1, Flag-RMI2, and Top3 α in the cells complemented with RMI2 Lys121Ala as well). These results indicate that the RMI/FANCM interaction is necessary for FANCM interaction with the Top3 α protein (Figure 3.1F).

X-ray crystal structure of FANCM MM2 bound to the RMI core complex. A crystallographic approach was taken to better define the interaction between FANCM and the RMI core complex. Crystals of the RMI core complex bound to the FANCM MM2 peptide diffracted to 3.3-Å resolution, and the structure of the ternary complex was determined using molecular replacement (Table 3.1, Figure 3.4A). $F_o - F_c$ electron density maps revealed the presence of a single FANCM MM2 peptide (residues 1226 – 1237) bound to a surface comprised of elements from both RMI1 and RMI2 (Figure 3.4B). Additionally, electron density for the N-terminal region of RMI2 (residues 13-16) was observed in the ternary structure; this region was not resolved in earlier structures of the apo RMI core complex (12, 13) (Figure 3.4B and 3.5). These residues form a series of apparent hydrogen bonding interactions with the MM2 peptide.

The MM2/RMI interface is primarily defined by a knobs-into-holes packing arrangement in which alternating hydrophobic side chains from the MM2 peptide project into hydrophobic pockets on the RMI1/RMI2 surface (Figure 3.4B). The MM2 peptide is highly extended in the complex, docking a total of 5 hydrophobic “knob” residues onto the RMI core complex. Each of the MM2 knob residues is highly conserved among FANCM proteins, and Phe1232 and Phe1236 are known to be critical for FANCM binding to the RMI subcomplex (17). Two of the hydrophobic MM2 residues, Phe1232 and Leu1234, dock into a pocket formed at the

RMI1/RMI2 interface formed by RMI1 residues 514 – 517, 585, and 586 and RMI2 residues 90 and 120-121. The remaining MM2 knob residues bind to either RMI1 (Phe1228 and Val1230) or RMI2 (Phe1236).

Consistent with its importance in mediating the RMI/FANCM interaction, RMI2 residue Lys121 helps to form the hydrophobic pocket where MM2 residues Phe1232 and Leu1234 bind (Figure 3.4A, 3.4B). In structures of the apo RMI core complex, the ϵ -amino group of the Lys121 side chain is bound to the main chain carbonyl groups of residues 514 and 586 from RMI1 (Figure 3.4C) (12, 13). However, in the ternary complex with MM2, the Lys121 side chain is altered and its ϵ -amino group is bound to the side chain of RMI2 Thr123 and the carbonyl of RMI1 residue 586 (Figure 3.4D). This change is accompanied by structural changes in adjacent residues (120-124) in RMI2 (Figure 3.4C, 3.4D and 3.6). This overall movement appears to enhance the hydrophobicity of the binding pocket to better accommodate the MM2 peptide. These rearrangements suggest that Lys121 is important for conformational changes in the RMI subcomplex that help stabilize formation of the RMI/MM2 complex. Moreover, the hydrophobic portion of the Lys121 side chain is in proximity to Phe1232 (3.5 Å) and Leu1234 (3.3 Å) from the MM2 peptide, indicating a possible direct role in stabilizing the complex through hydrophobic contacts.

Assessing the importance of RMI/FANCM interactions *in vitro*. A competition peptide-binding assay was developed to test the effects of changes in the MM2 peptide sequence on the stability of the RMI/MM2 complex (Figure 3.7A). In this assay, the RMI core complex was

prebound to F-MM2 and unlabeled MM2 peptide variants were titrated into the RMI core complex/F-MM2 solutions. Concentration-dependent binding of the MM2 peptide variants to the RMI core complex was monitored as a decrease in FP that is observed as F-MM2 is liberated from the complex. Wild type MM2 peptide dissociated F-MM2 with an IC_{50} of 520 ± 50 nM (Figure 3.7B, 3.7C). A panel of MM2 variants in which hydrophobic knob residues were individually or multiply substituted for alanine were then tested to determine the impact of the sequence changes on binding the RMI core complex. The most dramatic effects were observed with single Ala variants of Phe1232 (IC_{50} $85,800 \pm 9,400$ nM) and Leu1234 (IC_{50} $57,600 \pm 4,900$ nM) and with the double Phe1232/Phe1236 Ala variant (no competition observed) (Figure 3.7B, 3.7C). Since Phe1232 and Leu1234 dock in the hydrophobic pocket formed at the RMI1/RMI2 interface, these dramatic effects coupled with the strong effects of the RMI2 Lys121Ala variant on complex formation argue that this interaction is a key anchor point in the RMI1/RMI2/FANCM ternary complex. The strong defect of the double Phe1232/Phe1236 variant was consistent with previous results in which the same sequence change in MM2 resulted in a defect in FANCM co-immunoprecipitation with BLM and RMI1 (17). The single MM2 Phe1228, Val1230, and Phe1236 Ala variants were defective in the competition assay as well, but with more modest defects than were observed with the other peptides.

To verify the competition binding results in a cellular context and with full-length FANCM, we performed co-immunoprecipitation experiments with FANCM variant proteins transiently transfected into HEK293 cells (Figure 3.8). The results strongly paralleled the competition

binding results, with the strongest defects observed for the Phe1232Ala and Leu1234Ala single site variants, and weaker defects for Phe1228Ala, Phe1236Ala, and Val1230Ala.

The RMI/FANCM interaction is important for suppressing cellular SCEs. To determine the cellular effects of destabilizing the RMI/FANCM interaction, FANCM variants with MM2 sequence changes were stably expressed in *fancm*^{-/-} DT40 cells and SCE levels were measured. In *fancm*^{+/+} DT40 cells, the mean number of SCEs per cell was 2.0 whereas in *fancm*^{-/-} cells the mean increased to 17.7 SCEs/cell (Figure 3.9). Expression of wildtype FANCM complemented the SCE phenotype in *fancm*^{-/-} cells, reducing the mean to 3.1 SCEs/cell. The SCE levels of FANCM variants were then measured in the same assay. Expression of variants with the most severe *in vitro* binding defects, Phe1232Ala and Leu1234Ala, led to mean SCEs/cell levels of 14.2 and 11.7, respectively. Expression of FANCM variants with intermediate binding defects, Phe1236Ala, Val1230Ala, and Phe1228Ala, showed mean SCE/cell levels of 9.7, 5.5, and 5.5, respectively. Thus the defects of MM2 variants in RMI/FANCM complex binding paralleled the SCE increases that accompanied the MM2 sequence changes. This observation strongly supports a model in which the interaction between FANCM and the RMI subcomplex through the interface identified in the structure is critical for coordinated activity of the FA and BLM dissolvasome complexes *in vivo*.

Discussion

Coordination of BLM dissolvasome and FA core complex activities is critical for suppressing potentially harmful SCE events that can arise from DNA damage repair processes. Here, we have defined the molecular interface that mediates this coordination, formed between the RMI subcomplex from the BLM dissolvasome and the FANCM protein from the FA core complex. These studies identify a previously unrecognized role for RMI2 in binding FANCM and show that the RMI/FANCM interface is required for the proper cellular assembly of FANCM with the BLM dissolvasome. Our results support a model in which the RMI/FANCM interaction provides the primary direct link between the BLM dissolvasome and the FA core complex to preserve genomic stability (Figure 3.1F, model).

Structure of the RMI/FANCM interface. Our structural and biochemical observations show that interaction between FANCM and the RMI core complex depends heavily upon docking of conserved hydrophobic side chains from FANCM into hydrophobic pockets on the RMI1/RMI2 surface (Figure 3.4). This knobs-into-holes packing arrangement utilizes an extended surface presented by both RMI1 and RMI2, with a key anchor site for FANCM binding located at a pocket formed at the interface between the two RMI proteins. Sequence changes to either the RMI1/RMI2 pocket (RMI2 Lys121) or in two FANCM “knob” residues (Phe1232 or Leu1234) that pack at this three-protein interface dramatically destabilize the ternary complex (Figure 3.4, 3.7, 3.8). The importance of this central site in RMI/FANCM complex stability is consistent with previously identified roles for FANCM Phe1232 in RMI1 binding and for FANCM Phe1232 and RMI2 Lys121 in the joint activities of the FA and BLM dissolvasome complexes

(7, 12, 17). In addition to this central interface position, several other important interactions involving FANCM residues Phe1228, Val1230, and Phe1236 were identified. These residues dock into hydrophobic pockets on RMI1 or RMI2 that are adjacent to the central anchor site. These secondary pockets stabilize the ternary complex and are likely to be important for the specificity of peptide binding.

Insights into the architecture of the BLM dissolvasome/FA core complex superstructure.

Our findings point to a primary role for the RMI/FANCM interface in mediating complex formation between the BLM dissolvasome and the FA core complex. An earlier study that identified the FANCM MM2 region as the RMI1 interaction site also demonstrated binding between an extended polypeptide including the MM2 region and purified Top3 α *in vitro* (17). However, immunoprecipitation experiments from cell lysates showed that the Flag-tagged RMI1 OB2 domain could co-precipitate with RMI2 and FANCM but not Top3 α (17), leaving the role of the possible Top3 α /FANCM interaction in directly mediating interaction between the larger BLM dissolvasome and the FA core complex unclear. Our results using the RMI2 Lys121Ala variant, which disrupts RMI1/RMI2 interaction with FANCM but not with components of the BLM dissolvasome (Figure 3.1D), help clarify the relative importance of the RMI/FANCM and possible Top3 α /FANCM interfaces in higher order complex formation. Top3 α co-immunoprecipitated with FANCM from cells expressing wildtype RMI2 but not from cells lacking RMI2 or expressing the RMI2 Lys121Ala variant (Figure 3.1E and 3.3). Thus, a direct interaction between the RMI proteins and FANCM is required for co-immunoprecipitation of Top3 α with FANCM, highlighting the primacy of the RMI/FANCM interface in higher order

complex formation. A direct interaction between FANCM and Top3 α in a super-complex formed through the RMI/FANCM interface is not excluded by these findings, but it appears to require the RMI/FANCM interface to form in cells.

Possible regulation of RMI/FANCM complex assembly. The RMI subcomplex forms a very stable interaction with the isolated MM2 peptide ($K_{d, \text{app}}$ of ≤ 4.5 nM) that could imply the proteins form a ternary complex under all cellular conditions. However, previously published data indicate that the BLM dissolvasome and FA core complexes only co-localize under conditions in which replication fork stalling is triggered, with co-localization being MM2-dependent (17). The mechanisms that regulate RMI/FANCM complex formation are presently unclear, but the interface could be regulated in a manner that facilitates conditional coordination of the activities of the complexes. One possibility is that FANCM and/or the RMI subcomplex undergo conformational changes that subsequently allow complex formation. Except for the highly conserved MM2 sequence, the central portion of FANCM is generally poorly conserved and predicted to be unstructured (Figure 3.1A) (17). It could be that FANCM folding occludes the MM2 sequence until it is triggered to expose the element through modification or ligand (DNA) binding to allow interaction with the dissolvasome. Conversely, regulation could occur through changes to the RMI subcomplex. Regulated assembly of the RMI subcomplex is not likely since it appears to be an obligatory heterodimer; the hydrophobic surface area between RMI1 and RMI2 is extensive and cellular RMI1 and RMI2 protein levels are dependent on each other (7, 12). However, full-length RMI1 is structurally dynamic, and there is a possibility that

the RMI1 structure is altered upon response to a cellular trigger in such a way that exposes the FANCM binding site (Figure 3.1A) (12).

A model for complex coordination and SCE suppression. A theme that emerges from this work and previous studies is that depletion of proteins in either the BLM dissolvasome or the FA core complex, or disruption of protein interfaces that stabilize each complex or link the two, lead to similar increases in SCE levels (Figure 3.9 and (7, 11, 12, 19, 20)). Taken together, these data suggest that the RMI/FANCM interface is critical in bridging the BLM dissolvasome and the FA core complexes, and that proper association of these two complexes is vital in SCE suppression. This indicates that the BLM dissolvasome and the FA core complex are highly interdependent and may have roles beyond ICL repair.

A model in which SCEs can result from improper coordination between the BLM dissolvasome and the FA core complex during repair of stalled replication forks is supported by these findings. In this model, FANCM binds to a damaged site in the genome and recruits the BLM dissolvasome through its interaction with the RMI subcomplex (Figure 3.1F). FANCM could act to remodel the replication fork, possibly in coordination with BLM and Top3 α (20-24). Furthermore, FANCM interacts with a large number of proteins that mediate many DNA repair pathways. It may be that FANCM binds to stalled replication forks that form as a consequence of many types of DNA damage, and in certain cases recruits the BLM dissolvasome to aid in replication fork remodeling whereas other FA proteins may only be recruited by FANCM when the damage is an ICL. Additional interactions among proteins implicated in FA and BS (*e.g.*

BLM and FANCI, a helicase that acts downstream of the FA core complex) could also aid in the process of replication fork remodeling, as FANCI-deficient cells also show increases in SCEs (25). Failure to properly orchestrate the activities of these complexes would result in an accumulation of SCE events. Although we have identified the primary molecular and functional basis underlying the RMI/FANCI interaction, how other FANCI interactions mediate genome stability remains to be determined.

Materials and Methods

Plasmids, Protein Overexpression, and Purification

Purification of the RMI complexes. The RMI core complex, RMI core complex with RMI2 Lys121Ala, and RMI subcomplex were purified as described previously (12).

Purification of the MM2 peptide. cDNA encoding the FANCM MM2 peptide (residues 1218 to 1251) was cloned into a modified pGEX4T-1 plasmid (Amersham) to express a glutathione S-transferase (GST)-MM2 fusion protein in which GST and MM2 are separated by thrombin (pKH0065) or thrombin and tobacco etch virus (TEV) protease cleavage sequences (pKH0066). TEV protease cleavage separates the MM2 peptide (with a residual N-terminal Gly residue) from the remainder of the fusion protein. Multiple silent mutations were introduced for efficient *E. coli* expression (pKH0069). In addition, codons for two N-terminal residues Lys-Trp were added upstream of the MM2 sequence (pKH0070). The Lys side chain is used for fluorescein isothiocyanate (FITC) labeling and the Trp is used for peptide concentration determination. The TEV-cleaved MM2 peptide sequence is: GKWEDIFDCSRDLFSVTFDLGFCSPDSDDEILEHTSD, where the underlined sequence is that of FANCM. MM2 Ala variants for FP were cloned using site-directed mutagenesis with the pKH0070 vector.

Rosetta 2 (DE3) *E. coli* cells transformed with pLysS (Novagen) and either pKH0069 or pKH0070 were grown at 30°C in LB medium supplemented with 100 µg/ml ampicillin and 50 µg/ml chloramphenicol. At an OD₆₀₀ of ~0.6 cells were induced to overexpress protein with the

addition of 0.5 mM isopropyl β -D-thiogalactopyranoside for 3.5 hours. Cells were harvested with centrifugation, frozen at -80°C , and resuspended at 4°C in lysis buffer (50 mM Tris-HCl, pH 7.5, 20 % sucrose, 600 mM KCl) supplemented with 1 mM phenylmethylsulfonyl fluoride and protease-inhibitor tablets (Roche). Cells were lysed by sonication and centrifuged for 30 min at 15,000 rpm. The soluble lysate was incubated with Glutathione Sepharose 4B resin (Amersham) for 45 min, poured into a column, washed with lysis buffer, and eluted with GS4B elution buffer (25 mM Tris-HCl pH 7.5, 100 mM NaCl, 20 mM reduced glutathione). The eluate was concentrated to ~ 1.5 mL and incubated with TEV protease overnight at 4°C . MM2 peptide was resolved on a Sephacryl S-100 column (Amersham) run at 0.2 mL/min in S-100 buffer (10 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT, 5 % glycerol). The MM2 peptides that were used in crystallization (from pKH0069) or labeled with FITC (from pKH0070) were further purified using reverse-phase HPLC and their masses confirmed with MALDI Mass Spectrometry. MM2 peptide variants for competition assays were not HPLC purified; these were dialyzed against 10 mM Na_2CO_3 (pH 9.3), lyophilized, and resuspended in water.

FITC labeling of MM2 (F-MM2)

FITC was reacted with ~ 0.5 mM MM2 peptide (with the sequence GKWEDIFDCSRDLFSVTFDLGFCSPDSDDEILEHTSD, where the underlined sequence is that of FANCM) at a 15-fold excess for 1 hour (in 100 mM Na_2CO_3 (pH 9.3), in the dark, at room temperature). Extensive dialysis against 10 mM Na_2CO_3 (pH 9.3) removed unreacted FITC. The calculated degree of labeling was ~ 0.8 label/MM2 peptide.

Size-exclusion Chromatography

Uncleaved GST-MM2 used in size-exclusion chromatography (from pKH0065) was purified as described for other experiments, except that the cleavage step was omitted and the GST-MM2 was purified using the Sephacryl S-100 column. The size-exclusion experiments were performed using an analytical Sephacryl S-200 column. The GST-MM2 and RMI core complex were run individually, and then mixed at a 1:1 molar ratio and run together. Individual fractions were electrophoresed on a 15 % SDS-PAGE gel.

Fluorescence polarization assays

FP was measured at 25 °C using a Panvera Beacon 2000 system (Invitrogen) with 490 nm excitation and 535 nm emission wavelengths for three replicates; the average FP value was plotted with one standard deviation of the mean shown as error. For direct binding experiments, dilutions of the RMI subcomplex, RMI core complex, or RMI core complex with RMI2 Lys121Ala (in 10 mM Tris pH 8.8, 1 mM DTT) were incubated for over 50 minutes at room temperature with 4-4.5 nM F-MM2 before measurement. Lower amounts of probe could not be tested due to high signal-to-noise. For competition experiments, 4-5 nM F-MM2 was pre-incubated for 50 minutes with 100 nM RMI core complex prior to the addition of MM2 variants; FP was measured after an additional 50 minute incubation. IC_{50} values are the concentrations of unlabeled peptides necessary for 50 % inhibition of F-MM2 binding. Data were fitted with a single binding-site model using the GraphPad Prism software package; lower baseline values were fixed at 27 mA to allow IC_{50} determination in cases where 100 % dissociation was not observed with the peptide concentrations tested.

RMI/MM2 complex crystallization and structure determination

RMI core complex (5 g/L, in 10 mM Tris pH 8.8, 1 mM DTT) was mixed at a 1:1 ratio with MM2 peptide (sequence GHMEDIFDCSRDLFSVTFDLGFCSPDSDDEILEHTSD) in the same buffer. The complex was mixed with mother liquor (200 mM Na₂SO₄, 5 % polyethylene glycol 3350) at a 1:1 (vol) ratio and crystals were formed by hanging drop vapor diffusion over a 6 month period. Crystals were transferred to a cryoprotectant solution (200 mM Na₂SO₄, 10 % polyethylene glycol 3350, 30 % ethylene glycol) and flash-frozen in liquid nitrogen.

X-ray diffraction data were indexed and scaled using HKL2000 (26). The structure of the RMI/MM2 complex was determined by molecular replacement using the apo RMI core complex structure (12) as a search model in the program Phaser (27) followed by rounds of manual fitting using Coot (28) and refinement using REFMAC5 (29). Coordinate and structure factor files have been deposited in the Protein Data Bank (PDB ID 4DAY).

Co-immunoprecipitation and SCE assays

For co-immunoprecipitation with wildtype and RMI2 Lys121Ala mutant, the Flag-tagged human RMI2 variants were stably expressed in chicken DT40 *rmi2*^{-/-} cells. Nuclear extracts were prepared as described previously (30). Co-immunoprecipitation experiments were performed using anti-FLAG M2-agarose (Sigma) as described (30).

For co-immunoprecipitation with FANCM, experiments were performed using anti-chicken FANCM antibody as described (20, 30). Immunoblotting shows levels of FANCM and Top3 α in

whole-cell lysates from various DT40 cells (wildtype, *rmi2*^{-/-} cells, and *rmi2*^{-/-} cells complemented with human WT or Lys121Ala RMI2).

For co-immunoprecipitations in HEK293 cells, different chicken FANCM variants were transiently transfected into HEK293 cells. Nuclear extracts were prepared, and co-immunoprecipitation was performed using a chicken FANCM antibody (30). Actin was used as an input control.

In order to measure SCE levels, chicken FANCM variants were stably expressed in chicken DT40 *fancm*^{-/-} cells. The SCE frequency was measured as previously described (12).

Acknowledgements

We thank Advanced Photon Source staff (LS-CAT beamline) and Ken Satyshur for assistance with data collection, and members of the Keck Lab for critical reading of this manuscript. We thank members of the Denu lab for their assistance with peptide purification and Mass Spectrometry. This work was funded by a grant from the NIH (GM068061, JLK) and by the Intramural Research Program of the National Institute on Aging (Z01 AG000657-08, WW). KAH was supported in part by an NIH training grant in Molecular Biosciences (GM07215).

Figure 3.1.

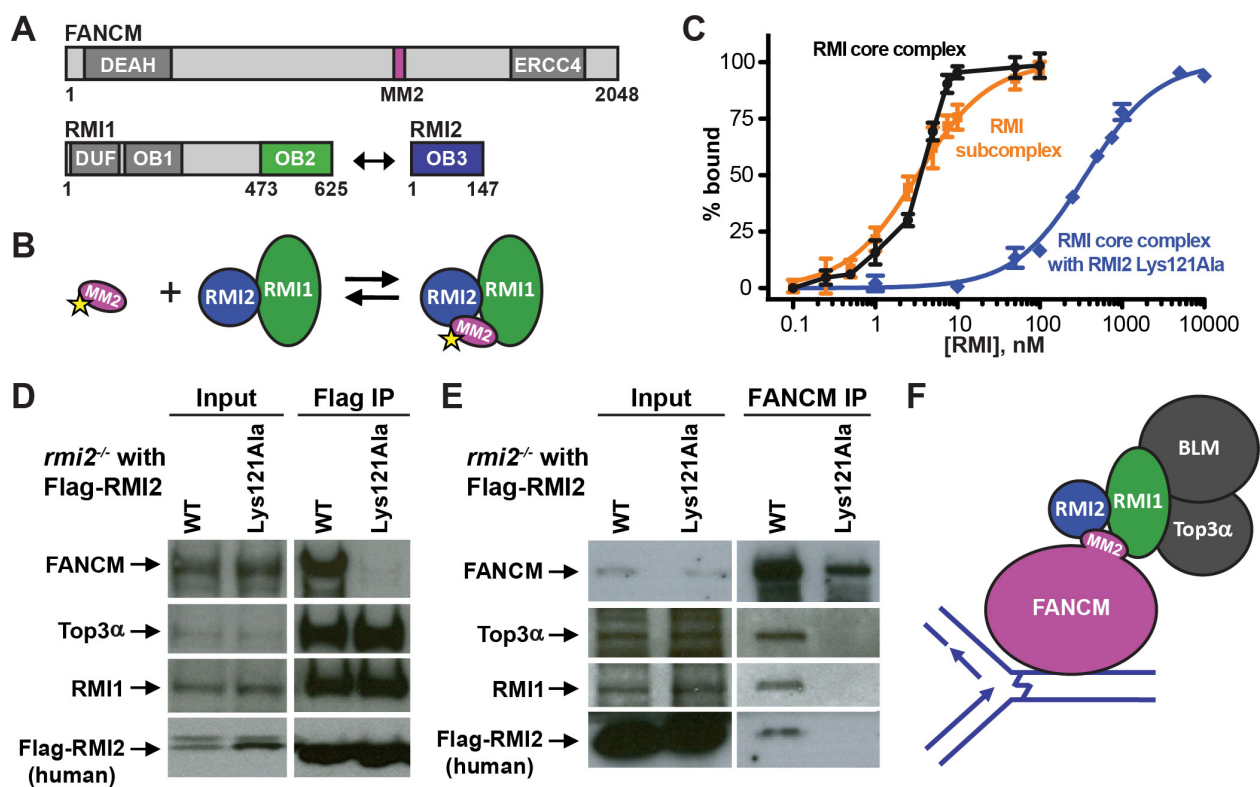


Figure 3.1. RMI2 Lys121Ala is defective in binding FANCM *in vitro* and *in vivo*. (A)

Schematic diagram of the domain structures of human FANCM, RMI1, and RMI2. The location of the MM2 portion of FANCM is shown in pink, and interacting OB domains comprising the RMI core complex are shown in green (RMI1 OB2) and blue (RMI2 OB3). Additionally, FANCM contains a DEAH helicase domain and degenerate ERCC4-like domain. RMI1 also contains a domain of unknown function, DUF, and an OB domain that interacts with Top3 α and BLM (OB1). **(B)** Scheme for the FP assay used to test binding in **(C)**. The MM2 peptide is FITC-labeled as indicated with the yellow star. **(C)** RMI2 Lys121Ala is defective in binding to MM2. Normalized binding curves for the RMI core complex (black), RMI subcomplex (orange), and RMI core complex with RMI2 Lys121Ala (blue) are shown. Experiments were done in triplicate and average FP value was plotted with one standard deviation of the mean shown as error. **(D)** RMI2 Lys121Ala is defective in FANCM binding *in vivo*. Protein input before co-immunoprecipitation (left), and the results of the co-immunoprecipitation (right), shows experiments with Flag-tagged RMI2 complemented into *rmi2*^{-/-} chicken DT40 cells, followed by Western to probe for the protein designated to the left of the gel. **(E)** The RMI subcomplex acts as a bridge between FANCM and Top3 α . Experiments were performed as in **(D)**, except FANCM was immunoprecipitated. The cell line is noted above each lane, and to the left the protein probed for in Western is designated. **(F)** Model of the FANCM/BLM dissolvasome interaction. FANCM at a site of DNA damage (depicted as an ICL) recruits the dissolvasome through the RMI subcomplex.

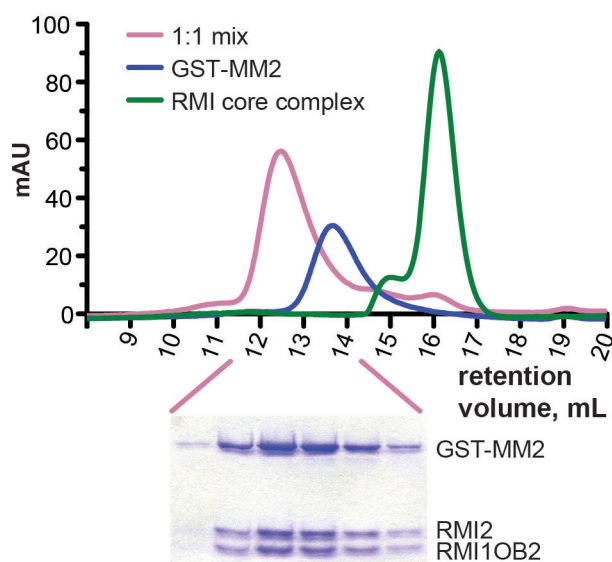
Figure 3.2.

Figure 3.2. Size exclusion chromatographic analysis confirms that the RMI core complex interacts with GST-MM2. The elution profiles for the RMI core complex (green), GST-MM2 (blue), and a ~1:1 mixture (pink) are shown. A 15 % SDS-PAGE gel shown below with the peak fractions of the RMI/GST-MM2 mixture confirms that a complex forms between the RMI core complex and GST-MM2. Fractions were collected every 0.5 mL with fractions corresponding to 11.5 – 14 mL as shown in the gel below.

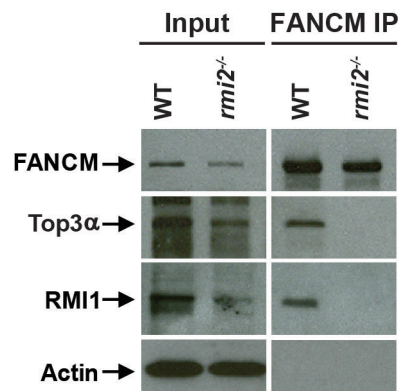
Figure 3.3.

Figure 3.3. The RMI subcomplex acts as a bridge between FANCM and Top3α. Protein input before co-immunoprecipitation (left), and the results of the co-immunoprecipitation with FANCM (right), shows experiments comparing wild-type and *rmi2^{-/-}* chicken DT40 cells. The cell line is noted above each lane and to the left the protein probed for in Western is designated. Actin was used as a loading control.

Figure 3.4.

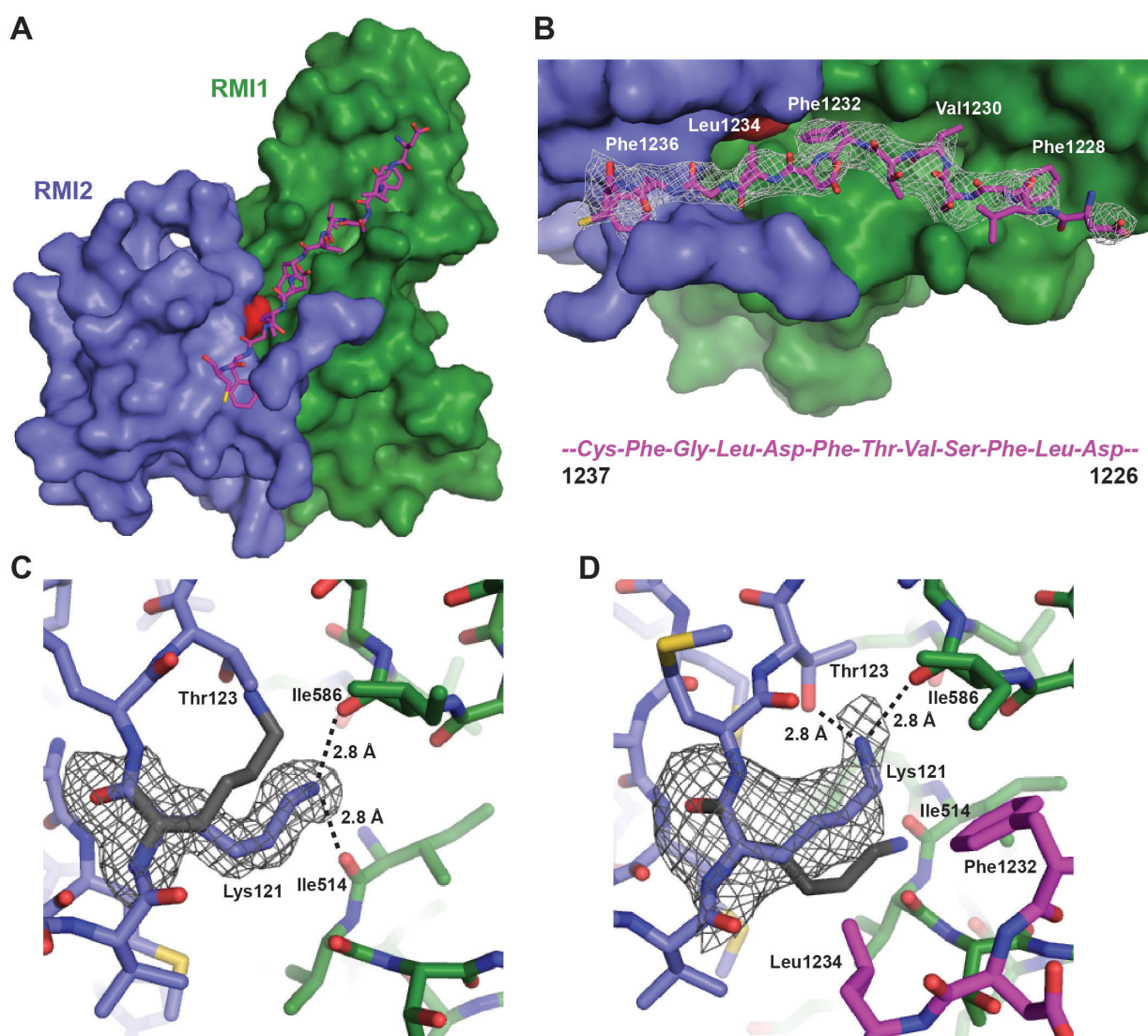


Figure 3.4. The structure of the RMI core complex bound to the MM2 peptide. (A) The overall structure of the RMI core complex with bound MM2 peptide. MM2 residues 1226 – 1237 are shown in pink. RMI1 is shown in green, RMI2 in blue, and RMI2 Lys121 is highlighted in red. (B) A closeup of the interface between the RMI core complex and the MM2 peptide. The peptide is shown with the C-terminus on the left with the corresponding residues shown below. Also shown is an $F_o - F_c$ electron density map (2σ) for the MM2 peptide. The same color scheme is used as in (A). (C) Shown in green (RMI1) and blue (RMI2) is the 1.55-Å resolution apo x-ray crystal structure (pdb 3mxn) with $F_o - F_c$ electron density map (2σ) for Lys121. Lys121 (grey) for the MM2-bound structure shows the alignment of the $C\alpha$ at Lys121. (D) The same is shown as in (C) except the MM2-bound structure is shown in color with Lys121 in grey from the apo structure. The $F_o - F_c$ electron density map (2σ) is shown for the MM2-bound Lys121.

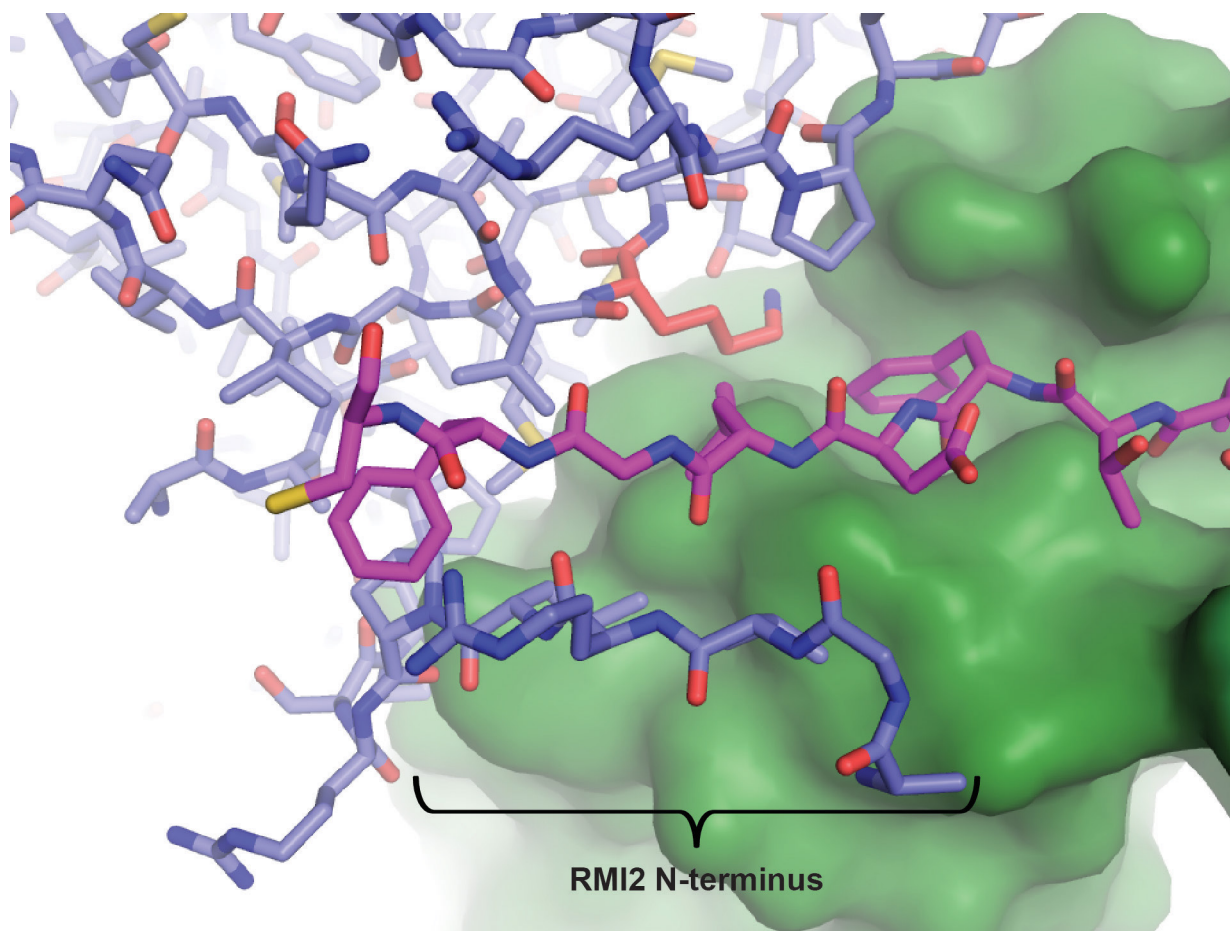
Figure 3.5.

Figure 3.5. The N-terminus of RMI2 orders to form β -sheet interactions with the MM2 peptide. The C-terminal portion of the peptide (pink) is shown interacting with the N-terminus of RMI2 (shown in blue). RMI1 is shown in green, and RMI2 Lys121 is also highlighted in red.

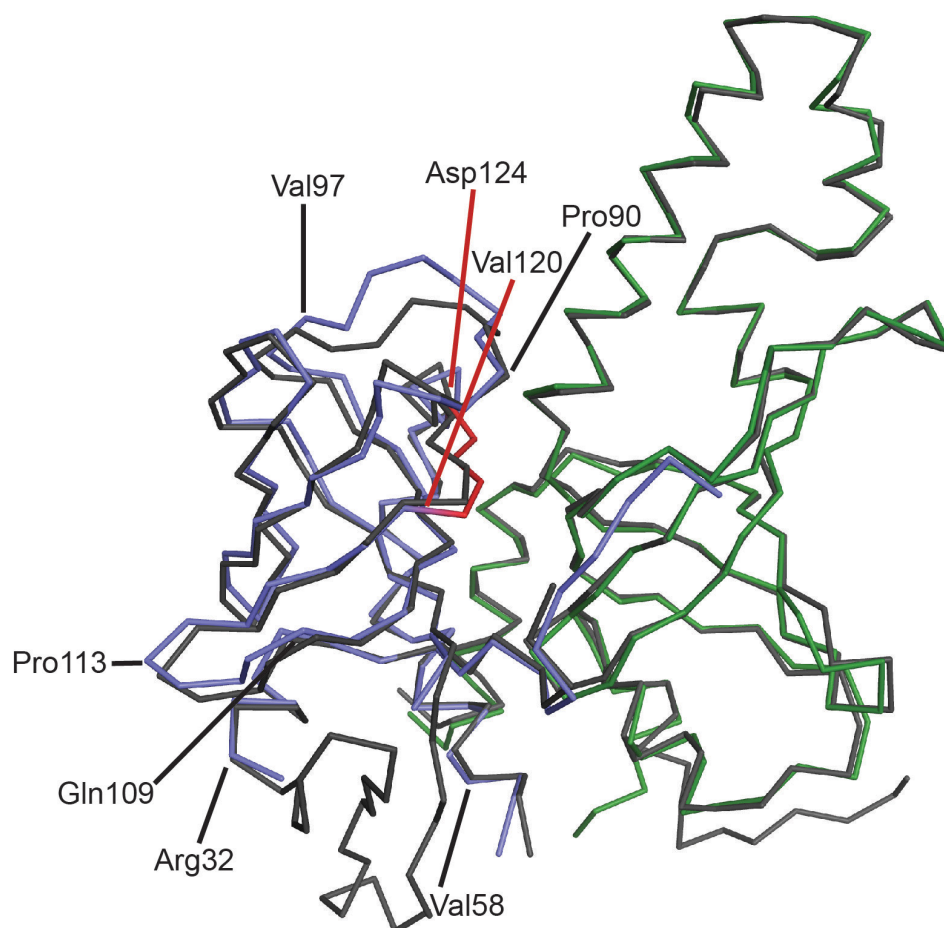
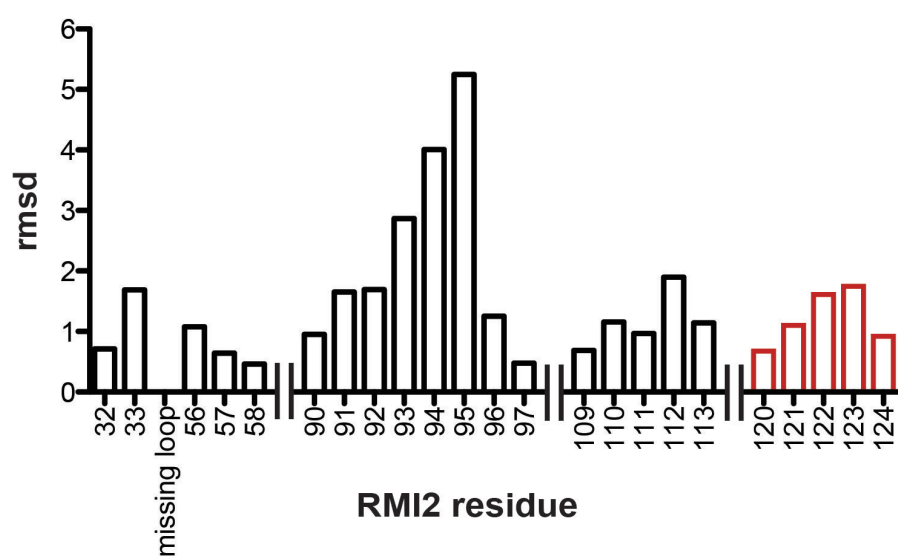
Figure 3.6.**A****B**

Figure 3.6. Differences between the apo and MM2-bound RMI core complex. (A) Mainchain C α residues from the apo (grey) and MM2-bound (RMI1 in green, RMI2 in blue) structures are aligned. Regions of movement are highlighted with lines pointing to the boundaries of these regions, and the rmsd for each of these residues is shown in (B). The region from RMI2 residues 120-124 is shown in red. **(B)** Four regions that contained the most movement between the two structures are plotted for RMI2. The rmsd for each individual RMI2 residue is shown. “Missing loop” accounts for the loop that was only observed in the apo structure (residues 34-55 were disordered in the MM2-bound structure). The region from residues 120-124 is shown in red, as in (A).

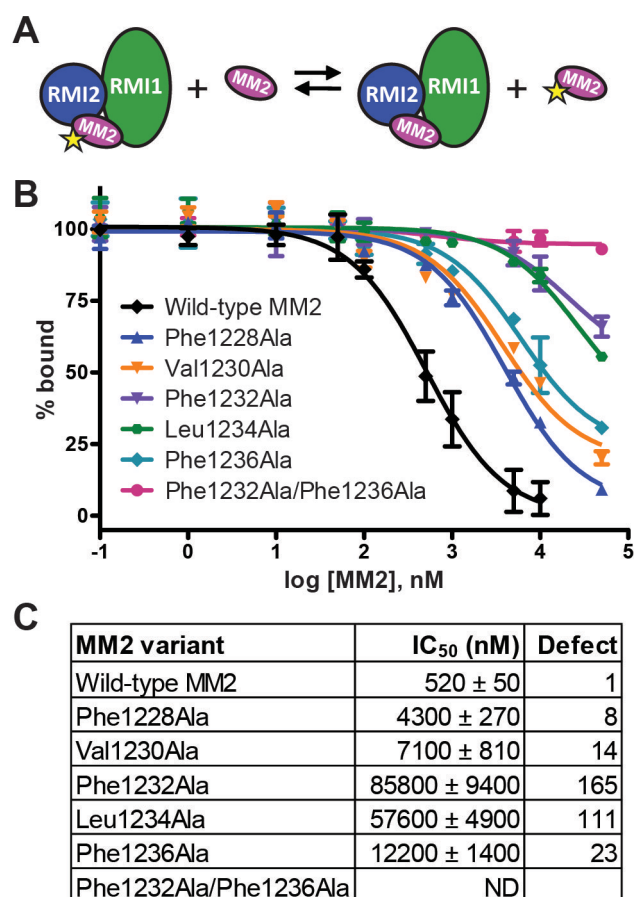
Figure 3.7.

Figure 3.7. MM2 variants have defects in binding the RMI core complex. (A) The scheme used in the FP competition assay. Unlabeled MM2 or variants are tested for their ability to compete with FITC-MM2 for binding to the RMI core complex. (B) Normalized FP competition results with different MM2 variants. Experiments were done in triplicate and average FP value was plotted with one standard deviation of the mean shown as error. (C) Table showing the IC₅₀ and competition defect for each MM2 variant.

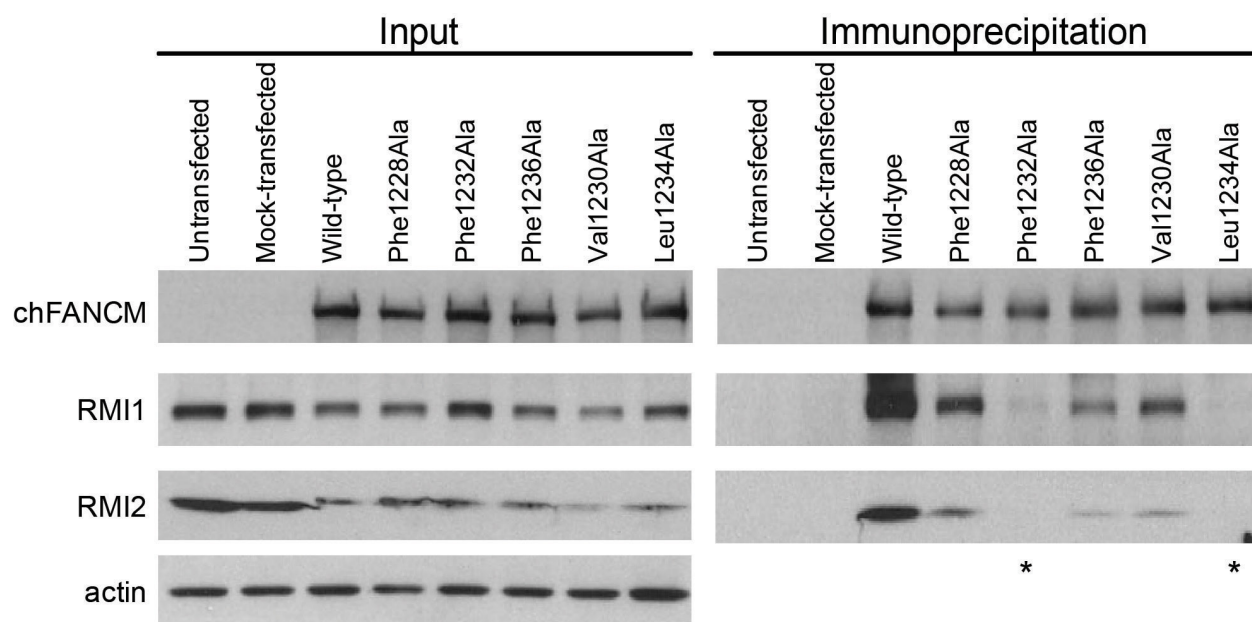
Figure 3.8.

Figure 3.8. Co-immunoprecipitations using various chicken FANCM variants in HEK293 cells. The left panel shows protein input before co-immunoprecipitation, and the right panel shows the results of the co-immunoprecipitation. Shown are co-immunoprecipitation experiments with Flag-tagged chicken FANCM and FANCM variants transiently transfected into HEK293 cells, followed by Western to probe for the protein shown to the left of the gel. Asterisks indicate the two residues with the most severe defect in co-immunoprecipitation.

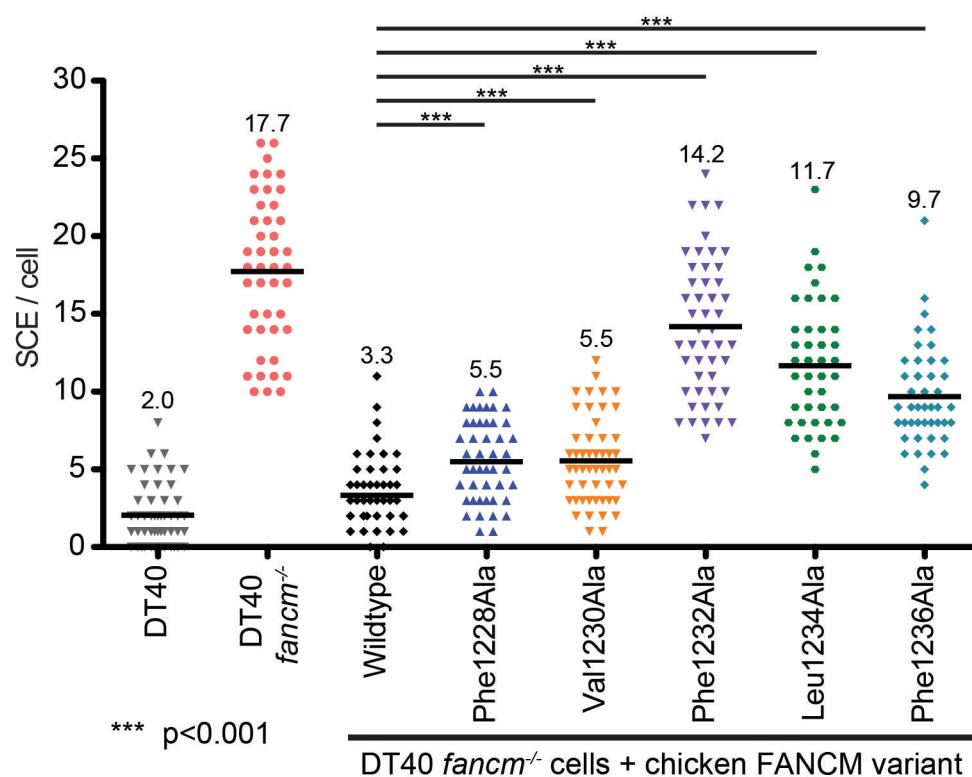
Figure 3.9.

Figure 3.9. FANCM variants with defects in binding increase SCEs. The number of SCEs per cell in chicken DT40 cells are shown as a histogram in which each point in a column is the number of SCE events in a single cell. The mean level of SCEs/cell is displayed as a black line and numerically at the top of the column. Statistical significance is indicated with *** (p<0.001). “DT40” denotes wild type DT40 cells, “*fancm*^{-/-}” denotes DT40 cells lacking FANCM, “wildtype” denotes *fancm*^{-/-} cells complemented with FANCM and the remainder denote the FANCM variants that have been substituted.

Table 3.1. X-ray data collection and structure determination statistics

Data Collection	
Wavelength, Å	0.97856
Resolution Range (high resolution bin), Å	50-3.30 (3.36-3.30)
Space Group	I2 ₁ 2 ₁ 2 ₁
Unit Cell (a, b, c (Å))	75.3, 96.4, 99.0
(α , β , γ (°))	90, 90, 90
Completeness (high resolution bin), %	99.7 (100)
Total/Unique Reflections	29106/5711
Redundancy (high resolution bin)	5.1 (5.1)
$\langle I/\sigma I \rangle$ (high resolution bin)	19.9 (2.4)
R_{sym}^{\dagger} , % (high resolution bin)	8.1 (58.7)
Refinement	
Resolution, Å	50-3.30
$R_{\text{work}}/R_{\text{free}}^{\ddagger}$, %	21.6/32.5
Rms deviations	
Bonds, Å	0.0102
Angles, °	1.38
Ramachandran statistics, %	
Most favored	84.3
Allowed	14.3
Generously allowed	1.3
Disallowed	0.0
# atoms	2098
$\langle B \text{ factor} \rangle$, Å ²	66.0

$^{\dagger} R_{\text{sym}} = \sum_j |I_j - \langle I \rangle| / \sum_j I_j$, where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

$^{\ddagger} R_{\text{work}}/R_{\text{free}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / |F_{\text{obs}}|$, where the working and free R factors are calculated by using the working and free reflection sets, respectively.

The free R reflections (5 % of the total) were held aside throughout refinement.

References

1. Moldovan GL & D'Andrea AD (2009) How the Fanconi anemia pathway guards the genome. *Annu Rev Genet* 43:223-249.
2. Kee Y & D'Andrea AD (2010) Expanded roles of the Fanconi anemia pathway in preserving genomic stability. *Genes Dev* 24(16):1680-1694.
3. Niedernhofer LJ, Lalai AS, & Hoeijmakers JH (2005) Fanconi anemia (cross)linked to DNA repair. *Cell* 123(7):1191-1198.
4. Ouyang KJ, Woo LL, & Ellis NA (2008) Homologous recombination and maintenance of genome integrity: cancer and aging through the prism of human RecQ helicases. *Mech Ageing Dev* 129(7-8):425-440.
5. Bachrati CZ & Hickson ID (2003) RecQ helicases: suppressors of tumorigenesis and premature aging. *Biochem J* 374(Pt 3):577-606.
6. German J (1993) Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine (Baltimore)* 72(6):393-406.
7. Xu D, Guo R, Sobeck A, Bachrati CZ, Yang J, Enomoto T, Brown GW, Hoatlin ME, Hickson ID, & Wang W (2008) RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. *Genes Dev* 22(20):2843-2855.
8. Singh TR, Ali AM, Busygina V, Raynard S, Fan Q, Du CH, Andreassen PR, Sung P, & Meetei AR (2008) BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvasome. *Genes Dev* 22(20):2856-2868.
9. Raynard S, Bussen W, & Sung P (2006) A double Holliday junction dissolvasome comprising BLM, topoisomerase IIIalpha, and BLAP75. *J Biol Chem* 281(20):13861-13864.
10. Wu L, Bachrati CZ, Ou J, Xu C, Yin J, Chang M, Wang W, Li L, Brown GW, & Hickson ID (2006) BLAP75/RMI1 promotes the BLM-dependent dissolution of homologous recombination intermediates. *Proc Natl Acad Sci U S A* 103(11):4068-4073.
11. Yin J, Sobeck A, Xu C, Meetei AR, Hoatlin M, Li L, & Wang W (2005) BLAP75, an essential component of Bloom's syndrome protein complexes that maintain genome integrity. *Embo J* 24(7):1465-1476.
12. Hoadley KA, Xu D, Xue Y, Satyshur KA, Wang W, & Keck JL (2010) Structure and cellular roles of the RMI core complex from the Bloom syndrome dissolvasome. *Structure* 18(9):1149-1158.

13. Wang F, Yang Y, Singh TR, Busygina V, Guo R, Wan K, Wang W, Sung P, Meetei AR, & Lei M (2010) Crystal structures of RMI1 and RMI2, two OB-fold regulatory subunits of the BLM complex. *Structure* 18(9):1159-1170.
14. Wu L & Hickson ID (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* 426(6968):870-874.
15. Meetei AR, Sechi S, Wallisch M, Yang D, Young MK, Joenje H, Hoatlin ME, & Wang W (2003) A multiprotein nuclear complex connects Fanconi anemia and Bloom syndrome. *Mol Cell Biol* 23(10):3417-3426.
16. Wang W (2007) Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet* 8(10):735-748.
17. Deans AJ & West SC (2009) FANCM connects the genome instability disorders Bloom's Syndrome and Fanconi Anemia. *Mol Cell* 36(6):943-953.
18. Meetei AR, Medhurst AL, Ling C, Xue Y, Singh TR, Bier P, Steltenpool J, Stone S, Dokal I, Mathew CG, Hoatlin M, Joenje H, de Winter JP, & Wang W (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet* 37(9):958-963.
19. Rosado IV, Niedzwiedz W, Alpi AF, & Patel KJ (2009) The Walker B motif in avian FANCM is required to limit sister chromatid exchanges but is dispensable for DNA crosslink repair. *Nucleic Acids Res* 37(13):4360-4370.
20. Yan Z, Delannoy M, Ling C, Dae D, Osman F, Muniandy PA, Shen X, Oostra AB, Du H, Steltenpool J, Lin T, Schuster B, Decaillet C, Stasiak A, Stasiak AZ, Stone S, Hoatlin ME, Schindler D, Woodcock CL, Joenje H, Sen R, de Winter JP, Li L, Seidman MM, Whitby MC, Myung K, Constantinou A, & Wang W (2010) A histone-fold complex and FANCM form a conserved DNA-remodeling complex to maintain genome stability. *Mol Cell* 37(6):865-878.
21. Singh TR, Saro D, Ali AM, Zheng XF, Du CH, Killen MW, Sachpatzidis A, Wahengbam K, Pierce AJ, Xiong Y, Sung P, & Meetei AR (2010) MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Mol Cell* 37(6):879-886.
22. Ralf C, Hickson ID, & Wu L (2006) The Bloom's syndrome helicase can promote the regression of a model replication fork. *J Biol Chem* 281(32):22839-22846.
23. Davies SL, North PS, & Hickson ID (2007) Role for BLM in replication-fork restart and suppression of origin firing after replicative stress. *Nat Struct Mol Biol* 14(7):677-679.

24. Luke-Glaser S, Luke B, Grossi S, & Constantinou A (2010) FANCM regulates DNA chain elongation and is stabilized by S-phase checkpoint signalling. *Embo J* 29(4):795-805.
25. Suhasini AN, Rawtani NA, Wu Y, Sommers JA, Sharma S, Mosedale G, North PS, Cantor SB, Hickson ID, & Brosh RM, Jr. (2011) Interaction between the helicases genetically linked to Fanconi anemia group J and Bloom's syndrome. *Embo J* 30(4):692-705.
26. Otwinowski Z & Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology*, eds Carter CW, Jr. & Sweet RM (Academic Press, New York), Vol 276, pp 307-326.
27. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.
28. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
29. Winn MD, Isupov MN, & Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 1):122-133.
30. Guo R, Xu D, & Wang W (2009) Identification and analysis of new proteins involved in the DNA damage response network of Fanconi anemia and Bloom syndrome. *Methods* 48(1):72-79.

CHAPTER 4

Crystal structure of a DNA-bound bacterial RecQ helicase reveals substrate-dependent conformational changes

Kelly A. Manthei performed protein purification, structure determination, fluorescence polarization, and SAXS experiments. Morgan C. Hill performed biochemical experiments with EcRecQ His156Ala and Gln322Ala. Jordan E. Burke assisted with data collection for SAXS and processed SAXS data.

Summary

RecQ DNA helicases are central genome maintenance enzymes found in nearly all organisms. We have determined the 2.4-Å resolution structure of the RecQ catalytic core from *Chronobacter sakazakii* (Cs) in complex with partial duplex DNA that contains a 3' single-stranded DNA (ssDNA) extension. The structure reveals direct roles for the winged-helix domain in binding duplex DNA, and for the helicase domain in binding ssDNA. Relative to the apo *Escherichia coli* RecQ catalytic core structure, the winged-helix domain of CsRecQ dramatically reorients when bound to DNA and is positioned in a similar location to that observed in the human RecQ1 catalytic core structure. This domain arrangement is accompanied by a ~90 degree bend in the DNA, which could aid in DNA unwinding. Unwound ssDNA is bound by an aromatic-rich loop (ARL) in the helicase domain that has been previously shown to be critical in coupling ATP hydrolysis to DNA-binding/unwinding. Biochemical studies confirmed the importance of several residues involved in DNA binding and unwinding, whereas a SAXS approach verified the DNA-dependent mobility in the winged-helix domain. These studies reveal mechanistic insights into large- and small-scale movements within the helicase that help coordinate ATP hydrolysis to DNA unwinding. Our observations also highlight surprising differences between prokaryotic and eukaryotic RecQs.

Introduction

Helicases are molecular motor proteins that use the energy derived from nucleoside triphosphates (NTPs) to catalyze the separation of duplex DNA and RNA. They are essential in nearly all aspects of nucleic acid metabolism, including DNA replication, recombination, repair, transcription, and translation (1). Helicases have been grouped into six superfamilies, with the two largest, SF1 and SF2, classified by the presence of seven conserved helicase motifs: I, Ia, II, III, IV, V, and VI (2). Structural studies have revealed that these motifs line the cleft formed between the two RecA-like helicase lobes (1). Motifs I and II correspond to the Walker A and B motifs, respectively, and play roles in NTP binding and hydrolysis. The other motifs are not as well conserved between SF1 and SF2 helicases, and help define the two superfamilies (1, 2).

The RecQ-family of DNA helicases are SF2 helicases that are critical for maintaining genomic stability from bacteria through humans (3). The founding RecQ-family member is *Escherichia coli* RecQ (EcRecQ), which was discovered as part of the RecF recombination pathway (4). Bacteria and yeast typically code for a single RecQ protein, whereas humans have five: RecQ1, BLM, WRN, RecQ4, and RecQ5. Three recessive genetic diseases are caused by mutations in the *BLM*, *WRN*, and *RecQ4* genes: Bloom syndrome, Werner's syndrome, and Rothmund-Thomson syndrome, respectively (5-7). Most RecQ proteins share a similar domain architecture, comprised of a helicase domain which contains two RecA-like helicase lobes, a "RecQ C-terminal" (RQC) domain, and a "helicase and RNaseD C-terminal" (HRDC) domain (8, 9) (Figure 4.1, 4.2). The RQC domain is further defined by a Zn^{2+} -binding scaffold and a winged-helix (WH) subdomain (9).

Structures of the apo and ATP γ S EcRecQ catalytic core, which includes the helicase and RQC domains, were solved in 2003 (9). These structures were followed by the structure of a human RecQ, RecQ1, which displayed some notable differences compared to EcRecQ (10). First, the location of the WH domain relative to the helicase and Zn²⁺-binding domains differed in the two proteins: in EcRecQ, it is almost directly below the second helicase lobe, perpendicular to the helicase domain, whereas in RecQ1 it has rotated almost 90° past the second helicase lobe (Figure 4.2B, 4.2C). Second, there is a β -hairpin within the WH domain that varies in length between human and bacterial WH structures, with the β -hairpin in EcRecQ being the shortest (9-11). Biochemical studies of β -hairpin variants have shown that both RecQ1 and WRN rely on hydrophobic residues at the tip of this β -hairpin to separate DNA, however, sequence changes to the β -hairpin in EcRecQ do not impair its helicase function (9-12). It remains unknown if bacterial RecQs use a similar or different mechanism in separating duplex DNA.

While RecQs are SF2 helicases, the RecQ helicase domain also shares significant structural similarities with SF1 helicases such as PcrA, Rep, and UvrD (9, 13, 14). In these helicases, an aromatic rich region in motif III has been shown to be important for coupling ATP hydrolysis to DNA unwinding (14-21). Structural studies with PcrA and Rep have demonstrated that aromatic residues from this region are able to stack with ssDNA (13, 14). Ultimately, the current model suggests that a motif III glutamine residue directly N-terminal to the aromatic rich region acts as a sensor for detecting ATP versus ADP and signaling this change through differences in DNA-binding affinity (13, 17, 18). We have previously shown that a similar aromatic rich loop (ARL) exists in RecQ helicases between motifs II and III, and in EcRecQ is responsible for coupling

ATP hydrolysis to DNA unwinding (22). However, in apo EcRecQ this ARL is not in an orientation that would allow for coordination between ATP binding and DNA unwinding, and so many questions remain as to the function of the ARL in RecQ proteins (9).

In this report, we describe the structure of the first bacterial helicase bound to a partial-duplex DNA substrate with a 3' ssDNA overhang. Surprisingly, the WH moves 86° relative to the position observed in the apo EcRecQ catalytic core structure to an orientation analogous to the RecQ1 structure. This change could indicate that, aside from β -hairpin differences, bacterial and human RecQs utilize a similar mechanism to unwind DNA. Furthermore, the ARL moves to stack with ssDNA in a striking parallel to previously described PcrA and Rep structures (13, 14). Mutagenesis studies confirm the importance of certain residues in binding and unwinding DNA, and SAXS was used to confirm the WH domain movement in solution. These studies lead to a model that incorporates small- and large-scale movement for bacterial RecQs in binding and unwinding DNA, and highlights similarities with human RecQs that may help to define a common mechanism for all RecQ helicases.

Results

X-ray crystal structure of the CsRecQ catalytic core in complex with DNA. Although our lab has previously solved the structure of the EcRecQ catalytic core, attempts at crystallizing the EcRecQ in complex with DNA were unsuccessful (9). Therefore, a new approach was taken to crystallize RecQ proteins from similar bacterial species; since unconserved surface residues are variable, these changes might allow for better crystal packing. The RecQ catalytic core domain from *Chronobacter sakazakii* (CsRecQ) (85% identical to EcRecQ) was crystallized in complex with a DNA hairpin containing a 10 base pair region with a 10 base ssDNA 3' overhang (Figure 4.1). Crystals of the CsRecQ catalytic core DNA complex diffracted to 2.4-Å resolution, and the structure was determined using molecular replacement (Figure 4.3A). The DNA-bound CsRecQ structure was refined with good bond geometry and crystallographic statistics (Table 4.1). The overall structures of individual domains in apo EcRecQ were similar to DNA-bound CsRecQ. However, there are two striking differences. The first is that the WH domain of CsRecQ had rotated 86° in comparison with the apo EcRecQ structure due to DNA binding (23) (Figure 4.3B). Overall, the orientation of the WH is more reminiscent of the structure of human RecQ1 (10) (Figure 4.3C). The second difference is a rearrangement of the ARL that is coincident with DNA binding.

F_o-F_c electron density maps revealed a direct role for the WH domain in dsDNA binding, whereas the two helicase lobes bound the 3' ssDNA (Figure 4.3A). The movement of the WH domain induces an almost 90° bend in the DNA as it transitions from ds to ssDNA. Electron density was observed beginning at the third base of the 5' end of the DNA, which forms eight

canonical base pairs before reaching the DNA hairpin (Figure 4.4). The DNA hairpin, with the sequence GTAA, forms a sheared GA base pair between the guanine at 11 and adenine at 14. Electron density was not observed for bases 12 or 13 of the hairpin. Following the hairpin, the DNA is duplex until reaching bases 23 and 24, were paired in the DNA substrate but are now split; therefore, two bases were unwound by the helicase. The cytosine at position 23 is located in the 90° turn of the DNA (Figure 4.5). Bases 24-26 stack together and are pointed into a pocket formed at the interface of the two helicase lobes and the Zn^{2+} binding platform. This stacking continues with base 26 stacking with Phe158 of the ARL. The rest of the ssDNA bases are flipped out and oriented away from the protein. Interestingly, the end of the 3' ssDNA extends into an adjacent symmetrically related molecule, where the penultimate adenine at position 33 binds at the ATP binding site (Figure 4.6). This appears to facilitate crystal packing, and may be relevant to DNA-dependent oligomerization of RecQ (see Discussion).

Interactions between CsRecQ and DNA. As expected for a protein that acts without sequence specificity, most of the DNA contacts are to the phosphate backbone. The WH domain of CsRecQ aligns with an rmsd of 1.8 Å to the structure of the WRN WH bound to DNA, and binds DNA in a similar fashion (24) (Figure 4.4B). As in the WRN structure, the loop between helices $\alpha 20$ and $\alpha 21$ of the WH (residues 441-446) is located in the major groove of the DNA, and makes many apparent hydrogen bond contacts with the phosphates of the 3' side of the DNA (Figure 4.4A). Residues at the N-terminus of helix $\alpha 20$ (residues 432-434) also make contacts to the phosphate backbone of the 5' side of the DNA. Residues in the β -hairpin, specifically His489, also appear to make contacts to the phosphate backbone of the DNA. Additionally, a

salt bridge may be conserved between Arg446 (Arg993 in WRN) and the phosphate backbone of the 5' side of the DNA. Furthermore, residues of the loop between $\beta 8$ and $\alpha 11$, as well as within helix $\alpha 11$ in the second helicase lobe (residues 220-225) make apparent hydrogen bond contacts with the phosphate backbone of the 5' side of the DNA.

Several additional contacts are made at the ss-dsDNA junction. C23 appears to stack with Trp347, and then the ssDNA moves past Trp347 to complete a 90° turn (Figure 4.4B). Arg315 appears to make contacts with the base of G24, this interaction may only take place when certain bases (purines) are in this position. Arg246, Arg275, and Thr293 are conserved residues that appear to make important contacts with the phosphate backbone.

Reorientation of the ARL. The ARL is encoded between motifs II and III (Trp154-Arg159) and is conserved among RecQs. With DNA-bound the ARL is positioned to allow Phe158 to base-stack with the ssDNA, while Trp154 T-stacks with Phe158 (Figure 4.7A). This orientation of aromatic residues within the ARL is strikingly similar to SF1 helicases PcrA and Rep that contain an ARL as part of motif III (13, 14) (Figure 4.7B). Previously, an EcRecQ Phe158Leu variant was shown to possess DNA-independent ATP hydrolysis activity (22). In the CsRecQ structure, DNA binding induces multiple changes that are only observed in the DNA-bound structure, and help lock the ARL in this new orientation. These new interactions are formed between His156 and Asp191, and Arg159 and Glu124 (Figure 4.7A).

Furthermore, reorientation of the ARL also appears to play a role in communication between the two helicase lobes. Other SF2 helicases have a conserved histidine in motif II that interacts with a conserved glutamine in motif VI (25-27). In the crystal structure of EcRecQ bound to ATP- γ S, Gln322 of motif VI is 2.7 Å away from His156 of the ARL, but the histidine in motif II (His149) is too far away to interact. However, when the ARL moves in response to DNA, His156 is moved more than 15 Å away from Gln322, and instead, His149 of motif II is able to interact with Gln322 (Figure 4.7C). His149 and Gln322 become 3.2 Å apart in the DNA-bound CsRecQ structure and in this way allow communication between the two helicase lobes. Furthermore, Gln322 is pointed towards the ATP binding site and does not interact with either histidine in the apo EcRecQ structure, indicating that Gln322 may act as an important switch in ATP- and DNA-binding (Figure 4.7C).

Biochemical studies confirm the importance of certain residues and interactions. A panel of EcRecQ alanine variants were made based on the CsRecQ structure to test the crystallographic model. The different variants test the importance of reordering the ARL (Glu124, His156, Asp191), communication between the helicase lobes (His149, Gln322), dsDNA binding (Arg446), and ssDNA binding (Trp347, Arg246, Arg275, Thr293, Arg315). Mutations in the ARL (Trp154Leu, Phe158Leu, Arg159Leu) have been previously studied and were shown to have various defects, especially in their ability to unwind DNA (22). These new variants were tested for their ATPase activity, and their ability to bind and unwind a dsDNA with a 3' ssDNA overhang.

Previously, Phe158Leu was shown to have DNA-independent ATPase activity, indicating that the variant uncouples ATP-hydrolysis and DNA-unwinding (22) (Table 4.2). The panel of variants described above were tested using a spectrophotometric ATPase assay (28). These activities were then compared to that of EcRecQ, which has weak ATPase activity that is greatly stimulated in the presence of DNA (29, 30). A homopolymer cofactor (dT₂₈) was added to the reactions to allow measurement of DNA-dependent ATP hydrolysis with a homogenous DNA structure that wouldn't be unwound by the helicase, since ssDNA and dsDNA differ in their ability to stimulate ATP hydrolysis by RecQ (29, 30). Using this assay, the His156Ala EcRecQ variant was also observed to have an increase in DNA-independent ATPase activity ($43.9 \pm 1.1 \text{ min}^{-1}$), although its activity was not as high as the Phe158Leu variant ($86.3 \pm 8.7 \text{ min}^{-1}$) (Table 4.2). Both of these variants' ATPase activity is stimulated by DNA, however, not to nearly the same level as WT EcRecQ (Table 4.2). Furthermore, the Gln322Ala variant had very little DNA-independent ATPase activity ($0.7 \pm 0.4 \text{ min}^{-1}$) that was only stimulated 13-fold in the presence of DNA. Overall, these data indicate that the ARL plays a critical role in coupling ATP-hydrolysis and DNA-unwinding, and that communication between Gln322 and the ARL may facilitate this coupling.

Next, the EcRecQ variants were tested for their ability to bind a dsDNA with a 3' ssDNA overhang, which is similar to the substrate used in the structural studies. This substrate contains a duplex region of 18 bp and a 12 nt 3' overhang, and is labeled with fluorescein (F-3prOH). WT EcRecQ binds F-3prOH with an apparent dissociation constant ($K_{d, \text{app}}$) of $0.4 \pm 0.1 \text{ nM}$.

Overall, ARL variants do not have a defect in DNA binding affinity, although the Glu322Ala variant had an almost 20-fold defect in binding F-3prOH (Table 4.3, Figure 4.8). Arg246 and Arg275, which appear to bind ssDNA, have also been tested (Figure 4.5, 4.8). EcRecQ variant Arg246Ala had a strong defect in binding F-3prOH, suggesting Arg246 plays a critical role in DNA binding. Preliminary data suggests that Arg275Ala has no binding defect.

Finally, these variants were tested for their ability to unwind radiolabeled 3prOH in the presence of 1mM ATP/MgCl₂. Helicase activity requires that the enzyme can both bind DNA and coordinate ATP hydrolysis to strand separation. Consistent with previous results, WT RecQ was able to unwind 3prOH in a concentration-dependent manner and displayed half-maximal unwinding at ~0.005 nM EcRecQ (29). As was previously reported, neither the Trp154Leu nor Arg159Leu variants were able to unwind 3prOH at the highest enzyme concentration tested (100 nM), whereas the Phe158Leu variant required ~1000-fold more enzyme to reach 50% 3prOH unwound (22). The His156Ala variant required ~50-fold more enzyme to reach 50% 3prOH unwound, whereas the Gln322Ala variant was able to unwind some 3prOH, but it never reached saturation. These data indicate that the ARL and communication between the two helicase lobes is critical for the helicase to efficiently unwind DNA.

SAXS confirms movement of the WH domain. Since the EcRecQ and CsRecQ crystal structures revealed different orientations of the WH domain, SAXS was used to confirm that this domain movement was not a crystallographic artifact or a fundamental difference between the two bacterial proteins. The global structure of the CsRecQ catalytic core in solution was

determined using SAXS, and the apo EcRecQ catalytic core structure was fit into the model generated from the SAXS data (Figure 4.9). The CsRecQ experimental scattering curve was then compared to the predicted scattering curve from the EcRecQ structure generated using CRY SOL (31) (Figure 4.9B). The data agree remarkably well, which indicates that in solution, the catalytic core of both RecQs are in the same orientation as described by the EcRecQ structure. This gives strong evidence that, at least for the catalytic core, the WH indeed moves $\sim 90^\circ$ as it binds to DNA. SAXS was also performed on the CsRecQ catalytic core in complex with the DNA used in crystallization, however, the catalytic core is known to bind DNA with ~ 100 -fold less affinity, and this greatly complicated data analysis. To circumvent this problem, full-length CsRecQ was purified and SAXS was performed in the presence and absence of the DNA used in crystallization. However, the addition of the HRDC domain appeared to complicate data analysis and modeling even further. As such, only the data for the CsRecQ catalytic core is shown.

Discussion

RecQ DNA helicases are ubiquitously conserved from bacteria to humans, and play a role in DNA replication, recombination, and repair (3). They are SF2 helicases that also contain a similar ARL as PcrA-like SF1 helicases, and the mechanism that incorporated this ARL has been unclear thus far. Here, we report the structure of a bacterial RecQ helicase bound to DNA. Surprisingly, the WH domain moves as a consequence of DNA-binding. Furthermore, the ARL moves and stacks with the ssDNA in a similar manner as SF1 helicases PcrA and Rep (13, 14). Finally, biochemical studies confirm the importance of the ARL and communication between motifs II and VI in RecQ function.

Comparison between DNA-bound CsRecQ and human RecQ structures. The WH domain movement induced by DNA-binding causes the CsRecQ structure to align much more closely with the RecQ1 structure. However, the Zn^{2+} -binding domain is in a slightly different orientation, and the WH rotates even farther than the location of the RecQ1 WH (Figure 4.3C). Interestingly, an unpublished structure of RecQ1 bound to DNA has the WH in a similar orientation as the apo structure (reviewed in (12)). However, in the RecQ1 structure, the DNA appears to take a slightly different path across the WH as compared to CsRecQ. This difference could be the result of crystal packing or a reflection of the ability of the WH to bind many different DNA substrates, utilizing the apparent flexibility of this domain.

In contrast to the RecQ1 WH, the DNA-bound WRN WH aligns with the CsRecQ structure very well (Figure 4.4B). However, the WRN WH, as well as the RecQ1 WH, has a longer β -hairpin

that contains an aromatic residue at the tip that has been shown to be critical in DNA unwinding (10, 11). The shorter β -hairpin in EcRecQ has a histidine in a similar position that mutational studies revealed plays no role in DNA unwinding (10). Even removal of three residues at the tip to further shorten the EcRecQ β -hairpin has little impact on helicase function (10). Interestingly, the BLM protein seems to contain aspects from both β -hairpins; it lacks an aromatic residue at the tip similarly to bacterial RecQs, however, the length of the β -hairpin is more comparable in length to other eukaryotic RecQs. Therefore, different RecQ proteins may use different mechanisms, with only a subset of RecQs relying on the β -hairpin in DNA unwinding. This may lead to bacterial RecQs being more similar to certain eukaryotic RecQ proteins, and less similar to others, since higher eukaryotes often encode multiple RecQs.

The ARL is critical in coupling ATP hydrolysis to DNA unwinding. We have previously described evidence that the ARL played a similar role to the aromatic rich region in motif III of some SF1 helicases such as PcrA, Rep, and UvrD (14-22). Here we show that indeed the ARL stacks with the ssDNA in a similar manner as DNA-bound PcrA and Rep (13, 14) (Figure 4.7A, 4.6B). With the ARL being a conserved feature of RecQ helicases and would therefore be expected to play a similar role in all RecQs. In the PcrA-like SF1 helicases, there is also a motif III glutamine residue that acts as a sensor for detecting ATP versus ADP and a model exists where this sensor communicates to the aromatic rich region within motif III, causing changes DNA-binding affinity (13, 17, 18). In RecQs there is no motif III glutamine, and instead, a similar sensor may exist in a conserved feature of SF2 helicases, where a histidine in motif II contacts a glutamine in motif VI. For bacterial RecQs, this interaction has only been observed in

the presence of DNA. In fact, three states have been observed for Gln322 depending on the presence of ATP and/or DNA (Figure 4.7C). The location of Gln322 suggests that it may signal to motif II and the ARL only in the presence of ATP. Supporting this, a similar feature has been observed in HCV (Hepatitis C virus) helicase NS3, where the interaction is between His293 of motif II and Gln460 of motif VI (25). In the apo NS3 structure, Gln460 is not interacting with His293 and pointed towards the location of where the phosphates of ATP are presumed to bind, whereas in the RNA-bound structure, Gln460 is 3.7 Å away (25, 32). In summary, RecQs appear to use a hybrid of features from both SF1 and SF2 helicases in order to couple ATP hydrolysis to DNA unwinding. Exactly how this coordination occurs will require future studies.

Crystal packing may offer hints at an oligomeric state. The subject of the oligomeric state of RecQ DNA helicases has been debated for many years. RecQs have been reported with differing oligomeric states, and these often change depending on the addition of ATP and/or DNA. EcRecQ has been reported to hydrolyze ATP with a Hill coefficient of 3 in dsDNA unwinding, which indicates that three ATPase sites on three RecQs may act together in unwinding (33). Other studies have reported EcRecQ acting as a monomer, or that it is dependent on the length of the 3' overhang (34, 35). A recent single molecule study described the presence of two monomers on DNA, and that one monomer appears to inhibit the ATP hydrolysis activity of the other (36). One monomer would bind as shown in the CsRecQ structure, and the second monomer would bind to the ssDNA ahead of the first monomer. Our observation that the ssDNA extends into the ATP binding site of the next RecQ could explain the inhibition of ATP hydrolysis, since one site is blocked by DNA (Figure 4.6). However, we find very little

conservation at the interfaces observed in our DNA-bound structure. Furthermore, our DNA-bound SAXS studies show no evidence for a dimer in solution. Thus, more studies will be required in order to elucidate the oligomeric state of RecQs.

Future experiments

This chapter is incomplete: cloning is almost complete on the panel of variants described in the results section, and future biochemical studies will be aimed at testing the entire panel with all three biochemical experiments described (ATPase, DNA-binding via FP, and helicase assays).

Experimental Procedures

Purification of RecQ proteins. CsRecQ was cloned from genomic DNA (ATCC BAA-894D-5) and placed into pET28b for protein overexpression. The catalytic core was purified from full-length CsRecQ in pET28b through extensive thrombin cleavage, which removes the HRDC. To avoid removal of the HRDC from full-length protein, a TEV cleavage site was added to the N-terminal region of the protein and TEV was used instead for removal of the His-tag, which does not cleave the HRDC. Plasmids were transformed into Rosetta 2 (DE3) E. coli transformed with pLysS (Novagen). Cells were grown at 37°C in Luria Broth supplemented with 50 µg/ml kanamycin and 50 µg/ml chloramphenicol. At an OD₆₀₀ of ~0.6 cells were induced to overexpress protein with 1 mM isopropyl β-D-thiogalactopyranoside at ~37°C for 3.5 hours. Protein was purified as described previously for EcRecQ, with changes made to His-tag removal: for the catalytic core, thrombin cleavage was allowed to proceed overnight to remove the HRDC, and for the full-length CsRecQ, TEV was used overnight instead of thrombin to retain the HRDC. The CsRecQ proteins were further purified on a Sephacryl S-100 column (GE Healthcare) to ensure protein homogeneity.

EcRecQ full-length proteins were purified as previously described, with short thrombin cleavage reactions to ensure the HRDC would not be removed (22). The only change is that a HiPrep Q FF column (GE Healthcare) was used instead of a MonoQ.

CsRecQ/DNA complex crystallization and structure determination. CsRecQ catalytic core at 7 mg/mL in minimal buffer (1 M ammonium acetate, 10 mM Tris pH 8) was mixed at

approximately a 1 to 1.2 (protein to DNA) ratio with the self-annealed hairpin DNA with the sequence CGGTAGAATTGTAAAATTCTACCGGTGCCTTACT (resuspended in 10 mM Tris pH 8). The complex was mixed with mother liquor (70 mM NaOAc, 24% glycerol, 11% PEG 4000) at a 1:1 (vol) ratio and crystals were formed by hanging drop vapor diffusion. Crystals only formed in the presence of the DNA substrate. Crystals were flash-frozen in liquid nitrogen.

X-ray diffraction data were indexed and scaled using HKL2000 (37). The structure of CsRecQ bound to DNA was determined by molecular replacement using the apo RecQ structure (9) as a search model in the program Phaser (38) followed by rounds of manual fitting using Coot (39) and refinement using REFMAC5 (40) and PHENIX (41). Coordinate and structure factor files will be deposited in the Protein Data Bank.

ATPase assays. To compare DNA-dependent ATP hydrolysis profiles, purified RecQ variants were mixed with 0–1000 nM dT₂₈ in 20 mM HEPES, pH 8.0, 50 mM NaCl, 1 mM β -mercaptoethanol, 2 mM MgCl₂, 0.1 g/L bovine serum albumin, 1 mM ATP, at 25°C. Wild type (WT) or variant RecQ protein concentrations were 1, 5, or 50 nM. To compare the effects of ATP concentration on ATP hydrolysis roles, 50 nM of each variant protein was mixed with 0–4.55 mM ATP/MgCl₂ in 20 mM HEPES, pH 8.0, 50 mM NaCl, 1 mM β -mercaptoethanol, 0.1 g/L bovine serum albumin, 100 nM dT₂₈, at 25°C. Both ATPase assay variations also included an ATP regeneration system that converts ADP to ATP in a reaction that is coupled to the conversion of NADH to NAD⁺ (28). This coupled reaction can be detected spectrophotometrically by observing the decrease of A_{340nm} due to NADH oxidation. Steady-

state $\Delta A_{340\text{nm}}/\Delta t$ rates were measured and converted to $\Delta[\text{ATP}]/\Delta t$ to determine the ATPase rates and were normalized to the concentration of RecQ present in each reaction. K_m^{DNA} , k_{max}^{DNA} , K_m^{ATP} , and k_{cat}^{ATP} values were derived by fitting ATPase activity resulting from DNA or ATP titrations to the Michaelis-Menten equation (Curve Expert), $\text{rate} = (V_{\text{max}} \times [\text{ATP}]) / (K_m + [\text{ATP}])$. Intrinsic ATPase rate is reflected by the constant k_{min}^{DNA} , and is measured as the rate of ATP hydrolysis in the absence of DNA cofactor. Measurements are reported in triplicate and error bars and uncertainties represent one standard deviation of the mean. Uncertainties reported for all values are one standard deviation of the mean.

DNA Binding Assay. A 3'-fluorescein-conjugated 30 base ssDNA (oligo 1) was the DNA substrate for assays measuring binding to ssDNA. 3'-fluorescein-conjugated oligo 1 was annealed to non-labeled oligo 2 to create the 3' overhang substrate F-3prOH. RecQ proteins were diluted serially from 1,000 nM to 0.01 nM into 20 mM Tris, pH 8.0, 50 mM NaCl, 4% v/v glycerol, 1 mM β -mercaptoethanol, 0.1 g/L bovine serum albumin. Dilutions of RecQ proteins were incubated for 20 min at room temperature with 200-600 pM F-3prOH in a total volume of 100 μL . The fluorescence polarization of each sample was measured at 25°C with a Beacon 2000 fluorescence polarization system. Measurements are reported in triplicate and error bars/uncertainties represent one standard deviation of the mean. Uncertainties reported with apparent K_d values are one standard deviation of the mean.

Helicase assays. The DNA substrate was created by phosphorylating the 5' end of oligo 1 in a T4 polynucleotide kinase (PNK) reaction with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ and annealing phosphorylated oligo 1

to oligo 2 by boiling and slowly cooling an equimolar mixture, resulting in a radiolabeled 3prOH molecule. The substrate was purified via native PAGE on a 12% gel followed by electroelution and dialysis against 20 mM Tris, pH 8.0, 50 mM NaCl. Purified RecQ variants were incubated with substrate (~1 nM, molecules) in 20 mM Tris, pH 8.0, 50 mM NaCl, 1 mM β -mercaptoethanol, 1 mM MgCl_2 , 1 mM ATP, 0.1 g/L bovine serum albumin, 4% v/v glycerol for 30 min at 25°C with protein concentrations between 0.0001 and 100 nM. Reactions were terminated by addition of 11% v/v glycerol, 0.28% SDS (to denature RecQ) and 5 ng unlabeled oligo 1 (to prevent reannealing of the unwound radiolabeled DNA). The helicase products were analyzed by PAGE on a 12% non-denaturing gel, dried onto Whatman paper, visualized using a phosphorimager, and quantified (ImageQuant).

Small angle X-ray scattering data collection. SAXS data were collected at the National Magnetic Resonance Facility at Madison (NMRFAM) on a Bruker Nanostar at a wavelength of 1.54 Å and a sample-detector distance of 67.7 cm resulting in a momentum transfer range of $0.01 < q < 0.40 \text{ Å}^{-1}$. Data were collected for buffer and protein at 25°C for 2-5 hours (NMRFAM) in 20 mM Tris pH 8, 300 mM NaCl, 5% glycerol, 1 mM β -mercaptoethanol. Buffer subtraction was adjusted for the excluded volume of the protein. The predicted scattering intensity at $q = 0 \text{ Å}^{-1}$ and the radius of gyration (R_g) were determined by Guinier analysis (42) and compared between 2 concentrations (2.5-5 mg/ml) to detect interparticle interference.

SAXS analysis and modeling. SAXS data were processed using Primus and Gnom (ATSAS) (43) to determine the R_g and maximum dimension (D_{max}) of the protein and to obtain regularized

scattering amplitudes extrapolated to $q = 0 \text{ \AA}^{-1}$ and the pair distance distribution function (PDDF) for each sample. The latter provides a profile of all the pair distances within the molecule and is therefore descriptive of the overall shape. When computing the PDDF, D_{max} was adjusted by increments of 1-2 \AA and optimized based on the following criteria: 1) A smooth drop-off to zero probability at D_{max} 2) Agreement between the R_g measured using the Guinier transform and the R_g extracted from the PDDF and 3) Agreement between the regularized scattering determined from the PDDF and the experimental data. Furthermore, the regularization parameter, α , was kept below 5 to avoid over-smoothing. The resulting regularized scattering amplitudes were employed to calculate 10 replicate *ab initio* dummy atom models using DAMMIF (44). The quality and uniqueness of the results were further assessed by agreement between replicate models as quantitated by the mean normalized spatial discrepancy (NSD) (45).

Models were superimposed with one another using the Supcomb20 algorithm (46). Existing crystal structures (PDB ID 1OYW) were manually docked into the dummy atom models and the scattering of the resulting all-atom model was predicted using the FOXS web server (47).

Acknowledgments

We thank Advanced Photon Source staff (LS-CAT beamline) and Ken Satyshur for assistance with data collection, and members of the Keck Lab for critical reading of this manuscript. This work was funded by a grant from the NIH (GM068061, JLK). KAM was supported in part by an NIH training grant in Molecular Biosciences (GM07215). This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH grants P41RR02301 (BRTP/ NCRR) and P41GM66326 (NIGMS). Additional equipment was purchased with funds from the University of Wisconsin, the NIH (RR02781, RR08438), the NSF (DMB-8415048, OIA-9977486, BIR-9214394), and the USDA.

Figure 4.1

			0	Q672R	I		
EcRecQ	1	MAQAEVLNLESGAKQV	LQETFGYQQFRPGQ	EIIIDTVLSGRDCLV	VMPTGGGKSLCYQIP	60	
CsRecQ	1	MAQAEVYSQETLAKQV	LQETFGYQQFRPGQ	ATIIDAVLEGRDCLV	VMPTGGGKSLCYQIP	60	
BLM	643	RFQSLSFPHTKEMMKI	FHKKFGLHNFRTN	GLEAINAALLGEDCFI	LMPTGGGKSLCYQLP	702	
		*:	.	: : : : *	: : : *	* : : *	: : : : : *
			Ia				
EcRecQ		ALLNGLTIVVVSPLI	SLMKDQVDQLQANGVAAACLN	--STQTREQQLEVMTGCRTGQIRL	118		
CsRecQ		ALVKTGLTIVVVSPLI	SLMKDQVDQLQANGVAAACLN	--STQSREEQQAVLAGCRTGQVRL	118		
BLM		ACVSPGVTIVVISPLR	SLIVDQVQKLTSLDIPATYLTGDKTDSEATNIYLQLSKKDPPIKL	762			
		* :	* : : : *	* : : *	: : : *	: : *	: : : *
			II	ARL			
EcRecQ		LYIAPERLMLDN----	FLEHLAHWNPVLLAVDEAHCISQ	WGHDFRPEYAAALGQLRQRF	173		
CsRecQ		LYIAPERLMLDN----	FIDTLGYWDLAMVAVDEAHCISQ	WGHDFRPEYAAALGQLRARF	173		
BLM		LYVTPEKICASNRLISTLENLYERKLLAR	FVIDEAHCVSQ	WGHDFRQDYKRMNMLRQKFP	822		
		* : : : *	: :	: : : : *	: : : *	: *	: : *
			III	I841T	C878R		
EcRecQ		TLPEMALTATADDTT	RQDIVRLGLNDPLIQISSFDRPNIRYMLMEKFKP	---LDQLMRY	230		
CsRecQ		AVPEMALTATADDTT	RRDIVRLGLDDPLIEISSFDRPNIRYMLMEKFKP	---LDQLMRY	230		
BLM		SVPMALTATANPRVQKD	ILTQLKILRPQVFSMSFNRHNLKYYVLPKKPKKVAFD	CLEWI	882		
		: : *	: : : *	: : *	: : *	: *	: *
			G891E IV	C901Y			
EcRecQ		VQEQRGKSGI	IYCNSRAKVEDTAARLQSKGISAAAYHAGLENNVRADVQEKFQ	-RDDLQI	289		
CsRecQ		VQEQRGKSGI	IYCNSRAKVEDTAARLQSRGISAAAYHAGLEHEVRASVQEKFQ	-RDDLQI	289		
BLM		RKHHPYDSGI	IYCLSRRECDTMADTLQDGLAALAYHAGLSDSARDEVQKWINQDGCQV	942			
		: :	: : *	: : *	: : *	: : *	: : *
			V G952V	H963Y	VI		
EcRecQ		VVATVAFGMGINKPNV	RFVVHFDIPRNIESYYQETGRAGR	DGLPAEAMLFYDPADMAWLR	349		
CsRecQ		VVATVAFGMGINKPNV	RFVVHFDIPRNIESYYQETGRAGR	DGLPAEAMLFYDPADMAWLR	349		
BLM		ICATIAFGMGIDKPDV	RFVIHASLPKSVEGYQESGRAGR	DGEISHCLLFYTYHDVTRLK	1002		
		: *	: : *	: : *	: : *	: : *	: : *
			C1036F	C1055G/s/R			
EcRecQ		RCLEEKPGQLQDIERHKLNAMGAFAEAQ	---TCRRLVLLNYFGEG	--RQPCG	---N	399	
CsRecQ		RCLEEKAPGPLQDIERHKLNAMGAFAEAQ	---TCRRLVLLNYFGEG	--RQAPCG	---N	399	
BLM		RLIMMEKDGNNHTRETHFNLYSMVHYCENTITE	CRRIQLLAYFGENGFNPDF	CKKHDPVS	1062		
		* :	: *	: *	: *	: *	: *
			D1064V C1066Y				
EcRecQ		CDICLDPPK	-----QYDGSTDAQIALSTIGRVN	--QRFGMGYVVEVIRGAN	443		
CsRecQ		CDICLDPPR	-----RYDGLVDAQKALSAIARVE	--QRFGMGYVVEVIRGAN	443		
BLM		CDNCKCTKDYKTRDVTDDVKSIVRFVQEHSSSQGMRN	IKHVGPSGRFTMNMLVDIFLGSK	1122			
		** *	: :	: : *	: : *	: : *	: : *
EcRecQ		NQRIIRDYGHDKLVYGMGRDKSHEHWVSVIRQLIHLGLVTQ	NIAQHSALQLTEAARPVLR	503			
CsRecQ		NQRIRELGHDKLVYGIGRDQSQEHWVSVIRQLIHLGVVTQ	NIAQHSALQLTEAARPFLR	503			
BLM		SAKIQSGIFGKGSAYSRRHNAERLFKKLILDKILDEDLYINANDQAIAYV	MLGNKAQTVLN	1182			
		. : *	: *	: *	: *	: *	: *
EcRecQ		GESSLQLAVPR	----IVALKPKAMQKSF	GGNYDRKLF	AKLRKLRKSIAD	SNVPPYVFN	559
CsRecQ		GEAPIMLAVPR	----VAALKPRVAQKSYG	GGNYDRKLF	AKLRKLRKAIAD	ENIPPYVFN	559
BLM		GNLKVD	FMETENSSSVKKQKALVAKVSQREEMVKKCLGELTEVCKSLGKVF	GVHYFNIFN	1242		
		* :	: *	: *	: *	: *	: *
EcRecQ		DATLIEMAEQMPITASEMLS	VNGVGMRKLERFGKPFMALIRAHVDGDDEE	609			
CsRecQ		DATLIEMAEQMPISPGEM	LINGVGRKLERFGRPFMTLIREHVDGDDEE	609			
BLM		TVTLLKLAESLSSDPEVLLQIDGVTEDKLEKYGA	EIVISVLQKYSEWTPA	1292			
		. **	: : *	: : *	: : *	: : *	: : *

Figure 4.1. Sequence alignment between EcRecQ, CsRecQ, and BLM. The alignment was performed using Clustal (48). Invariant residues are shown with an asterisk below the alignment, highly conserved residues with a colon, and more weakly conserved residues with a period. Bloom syndrome missense mutations are shown in pink (49-53). Coloring of the CsRecQ sequence denote the helicase (blue and red), Zn²⁺-binding (yellow), WH (green), and HRDC (black) domains.

Figure 4.2

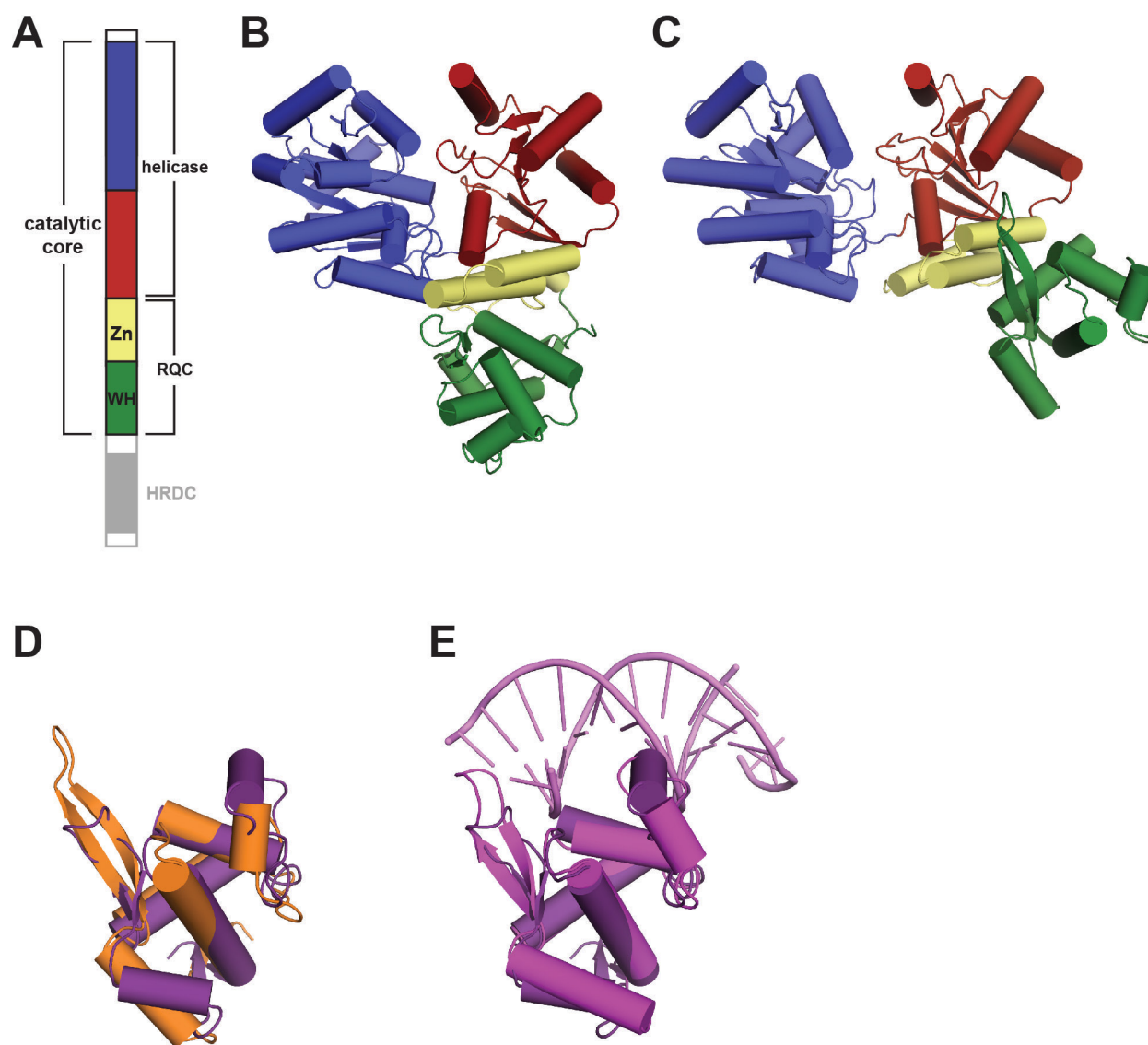


Figure 4.2. Previous RecQ structures. (A) Domain architecture of RecQ DNA helicases. Abbreviations used: RQC = RecQ-conserved, HRDC = helicase and RNaseD C-terminal, Zn = Zinc binding domain, WH = winged helix. (B) The structure of the catalytic core of apo EcRecQ (9). The domains are colored as shown in (A). (C) The structure of the catalytic core of apo RecQ1 (10). The domains are colored as shown in (A). (D) An overlay of the WH domains of EcRecQ (purple) and RecQ1 (orange). The β -hairpin is shown in the upper left. (E) An overlay of the WH domains of EcRecQ (purple) and WRN bound to dsDNA (WRN: magenta, DNA: pink) (11). The β -hairpin is shown in the upper left in position to act as a pin to separate the DNA.

Figure 4.3

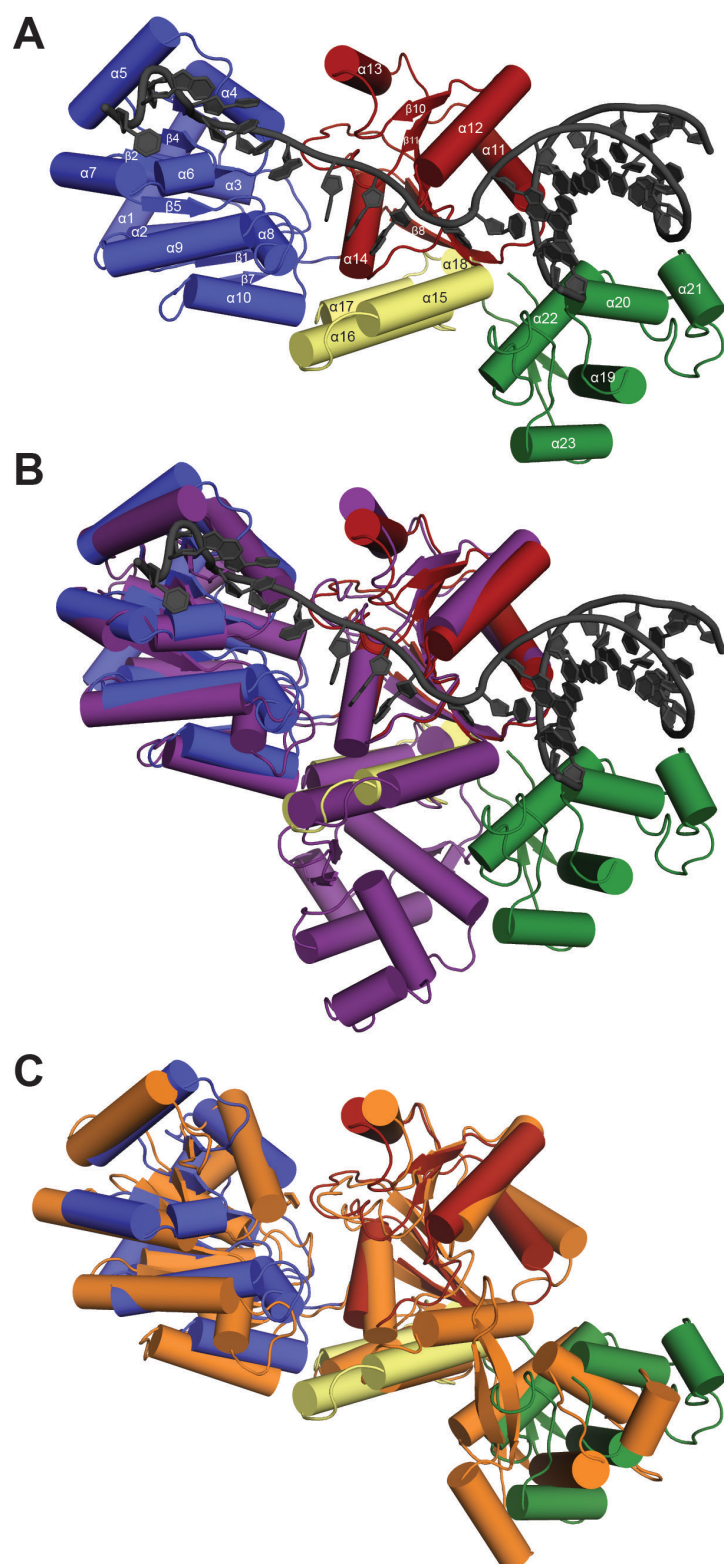


Figure 4.3. The structure of CsRecQ bound to DNA. **(A)** The overall structure of CsRecQ with DNA shown in gray. Helicase lobe 1 is depicted in blue, helicase lobe 2 in red, the Zn^{2+} -binding domain is shown in yellow, and the WH domain in green. **(B)** Alignment of the DNA-bound CsRecQ structure with the apo EcRecQ structure (shown in purple) (9). This highlights the movement of the WH domain. **(C)** Alignment of the DNA-bound CsRecQ structure (DNA hidden) with the apo human RecQ1 structure (shown in orange) (10). This alignment highlights the similarities between the two structures once the WH moves.

[illegible]

Figure 4.4. dsDNA binds to the WH domain similar to WRN. (A) Schematic representation of the dsDNA in the structure. Potential interactions within 3.3 Å are depicted with dashed lines. Water-mediated contacts are denoted with H₂O. The dashed line at the left indicates where the dsDNA connects to the ssDNA region. (B) An image of the CsRecQ WH with dsDNA. Also shown is an F_o–F_c electron density map (2 σ) for the dsDNA. Arg446 is represented in a stick model. (C) Alignment with the WRN WH domain. CsRecQ is shown in green with DNA in gray, while the WRN WH is shown in magenta with DNA in pink. Arg446 (CsRecQ), which is the same as Arg993 in WRN, are both represented as sticks: this is the only contact between a side chain and the phosphate backbone. The WRN aromatic residue that acts as a pin is also depicted as sticks (Phe1037).

B

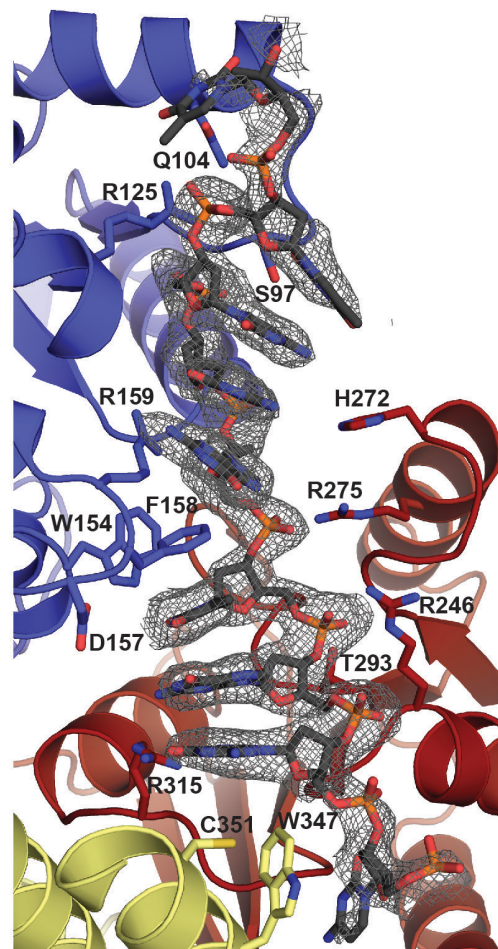
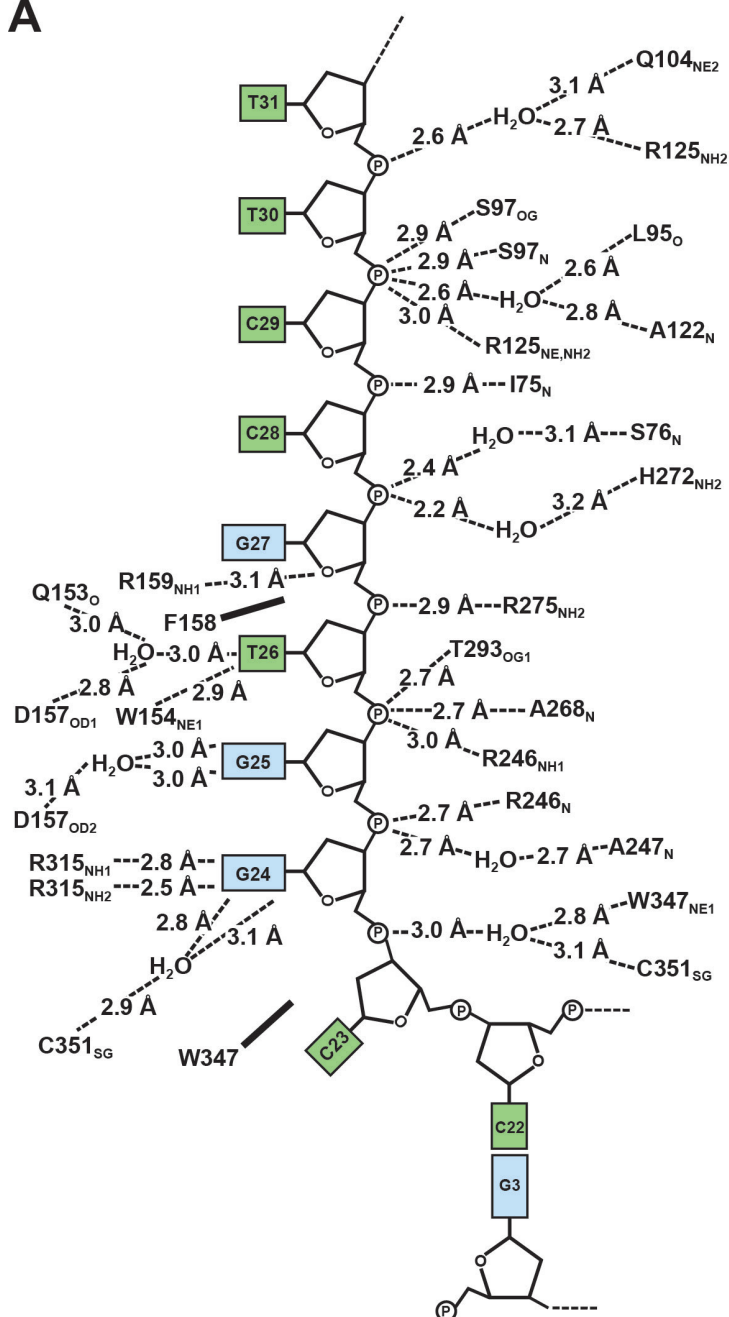
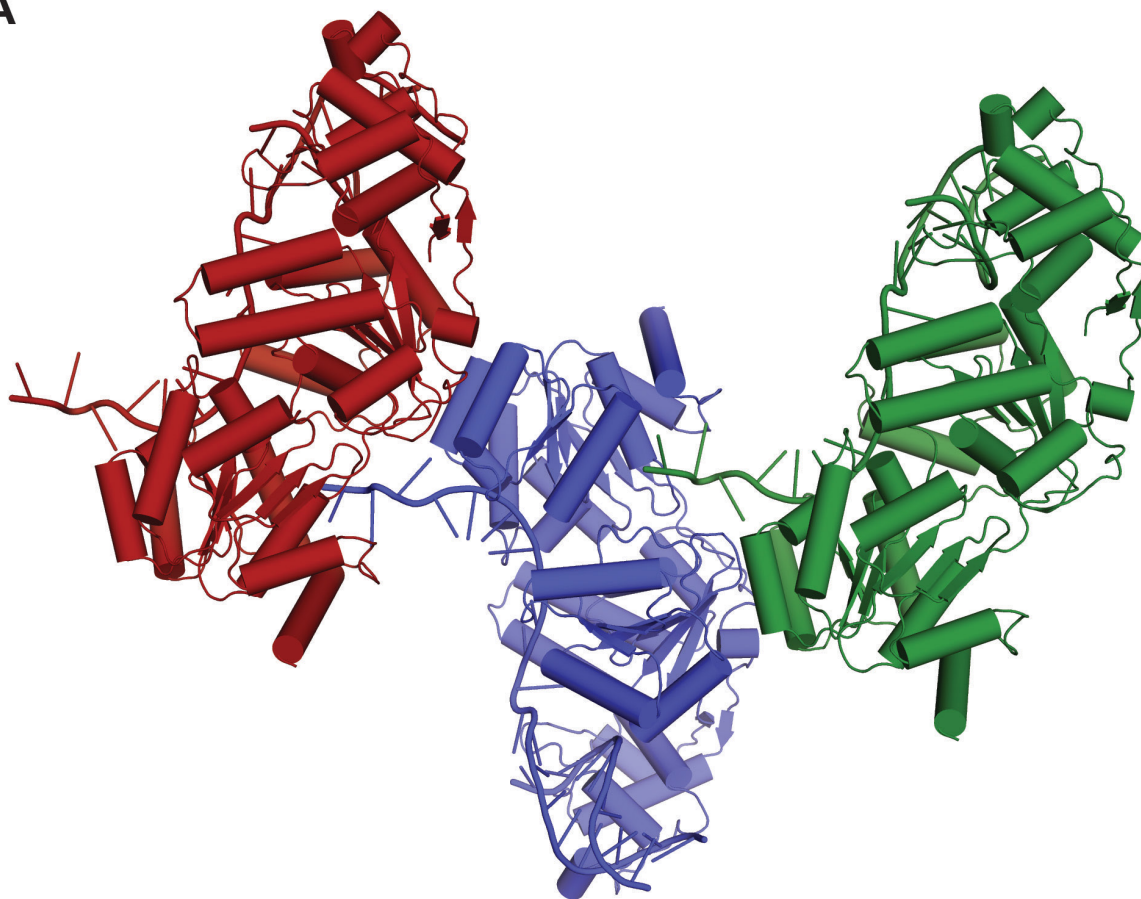


Figure 4.5. Details of the ssDNA binding interface. (A) Schematic representation of the ssDNA in the structure. Potential interactions within 3.3 Å are depicted with dashed lines. Water-mediated contacts are denoted with H₂O. The dashed lines at the beginning and end indicates where the ssDNA continues for two more bases (top), and where dsDNA connects to the ssDNA region (bottom). The solid lines indicate aromatic residues that stack with the DNA bases. (B) An image of the ssDNA from the structure with key residues shown as sticks and labeled. Also shown is an F_o–F_c electron density map (2 σ) for the ssDNA.

Figure 4.6

A



B

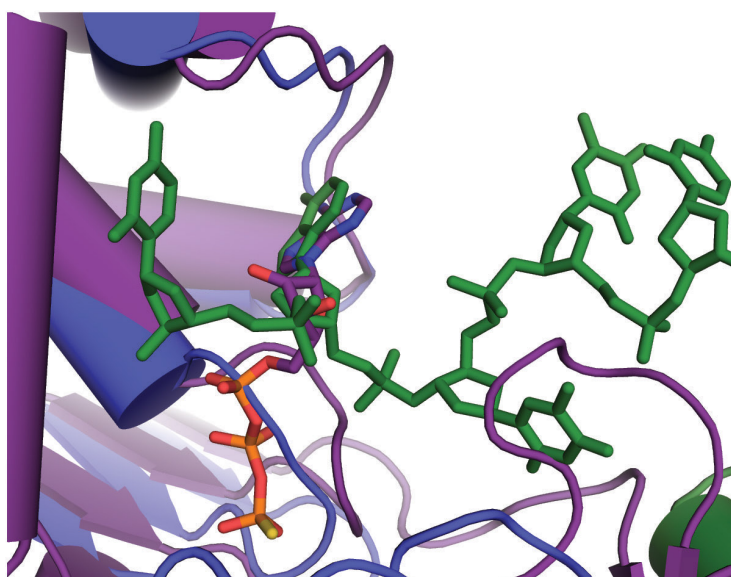


Figure 4.6. DNA connects the symmetry mates. **(A)** DNA-bound CsRecQ (blue) shown with two symmetry mates (green and red). The DNA from one symmetry mate reaches into the next. **(B)** The DNA from the symmetry mate binds at the ATP binding site. An adenine at position 33 binds to the ATP site (DNA shown in green). The DNA continues with one more base at position 34. The ATP binding site for CsRecQ (blue) is overlaid with the ATP binding site from EcRecQ, with bound ATP γ S shown in purple.

Figure 4.7

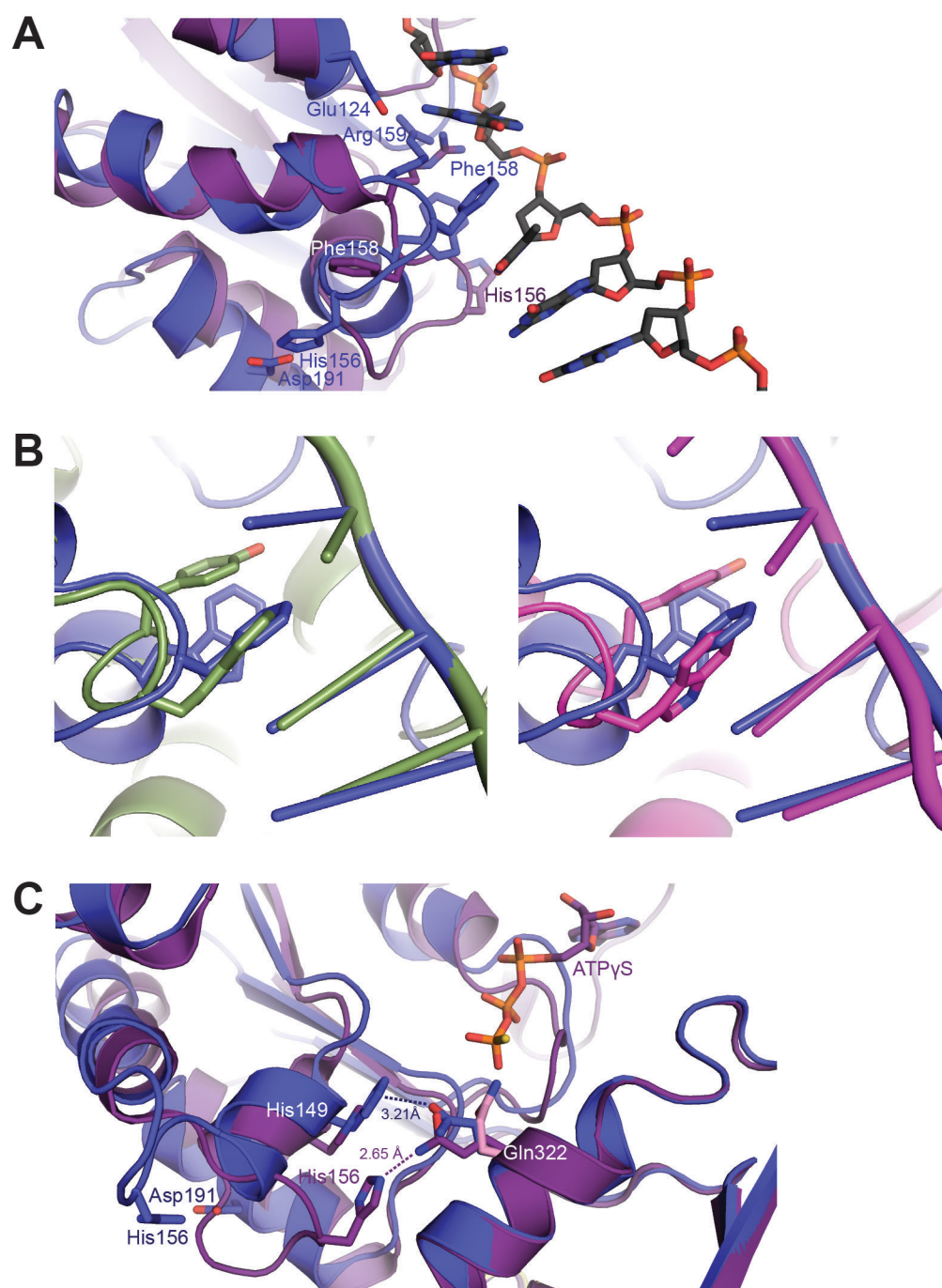


Figure 4.7. ARL movement in response to DNA binding. **(A)** EcRecQ and CsRecQ structure comparison highlight ARL movement. EcRecQ, shown in purple, aligned to the DNA-bound CsRecQ structure, shown in blue (9). The DNA is shown in gray. Important residues within the ARL and that help lock the ARL in place are labeled in the same color as their appropriate structure, with the exception of the Phe158 from EcRecQ, which is white. **(B)** The ARL movement is similar to PcrA (left panel, green) and Rep (right panel, magenta) (13, 14). In both panels, CsRecQ is shown in blue. PcrA and Rep were aligned manually by lining up the DNA from all three structures. All three structures have two aromatic residues in similar locations, however only one stacks with the DNA (Phe158 for CsRecQ, Trp259 for PcrA, Trp250 for Rep). **(C)** The ARL movement induces a motif II-VI interaction in CsRecQ. EcRecQ and CsRecQ are depicted in the same colors as in (A) and the region around Gln322 was used to align the structures. EcRecQ is shown with bound ATP γ S (purple), while only Gln322 from apo EcRecQ is depicted (pink) (9). The movement of Gln322 is highlighted: in the apo state, Gln322 (pink) is oriented towards the empty ATP binding site, while in the ATP γ S-bound structure, Gln322 (purple), is 2.65 Å from His156, while in the DNA-bound state, Gln322 (blue), is 3.21 Å from His149 of motif II. In the DNA-bound state, the ARL movement causes His156 to move away and it is able to interact with Asp191.

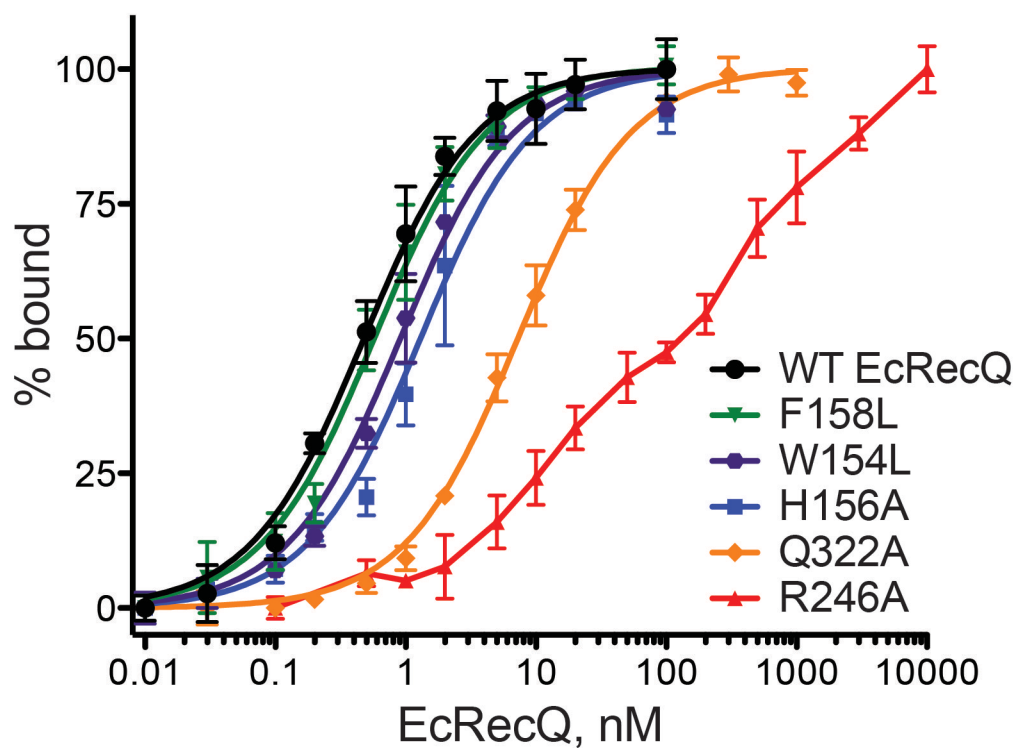
Figure 4.8

Figure 4.8. EcRecQ and variants binding to F-3prOH. Normalized binding curves for the WT EcRecQ and EcRecQ variants as indicated. Experiments were done in triplicate and average FP value was plotted with one standard deviation of the mean shown as error.

Figure 4.9

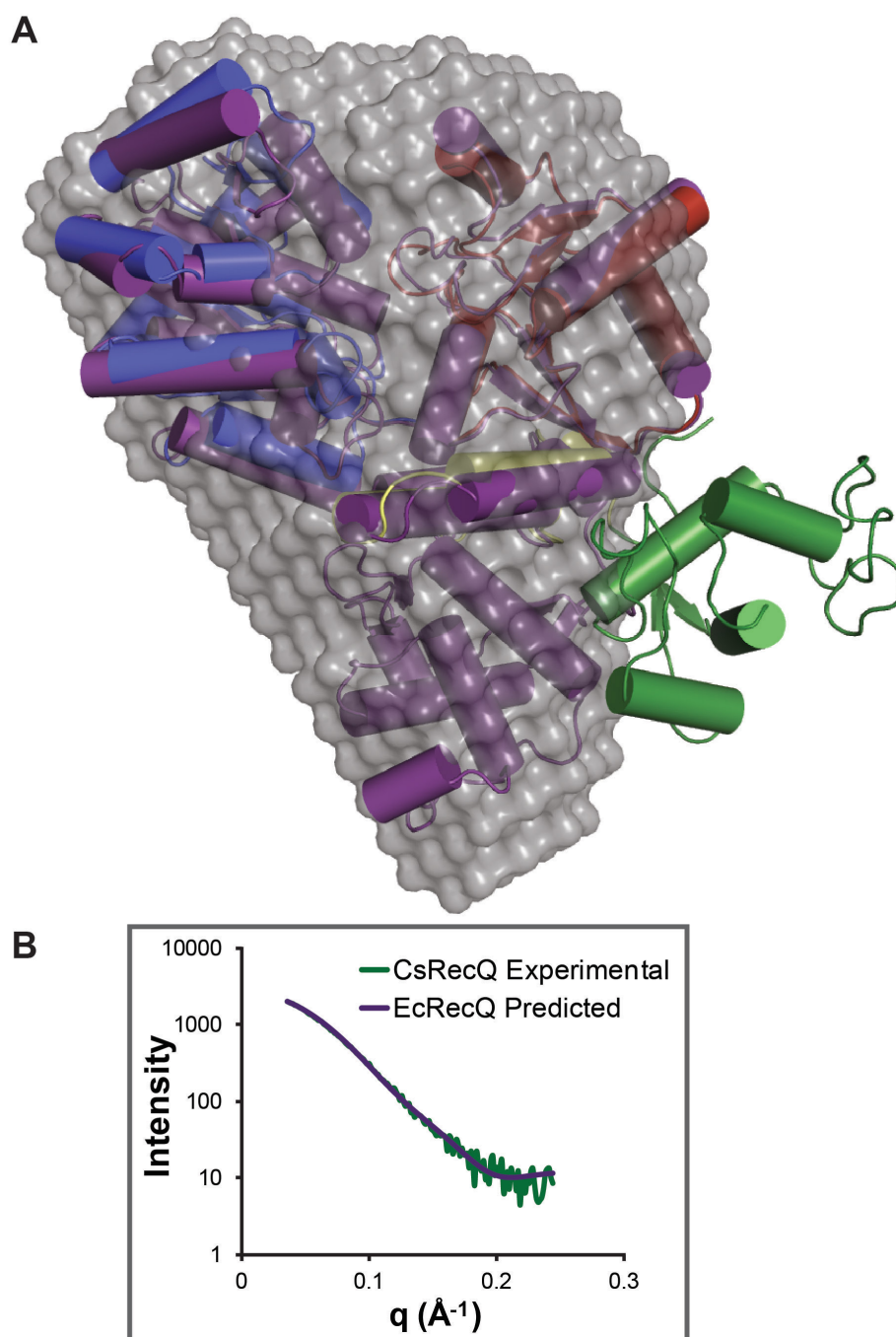


Figure 4.9. The WH reorients upon DNA-binding. (A) SAXS was used to confirm the orientation of the WH in apo CsRecQ. EcRecQ (purple) aligned with the CsRecQ x-ray crystal structure highlights the movement of the WH. The averaged dummy atom model of the experimental data (B) is shown in gray. (B) The experimental scattering data compared to the data predicted from the apo EcRecQ crystal structure using CRY SOL (31). 10 models were generated with a normalized spatial discrepancy (NSD) of 0.73.

Table 4.1. X-ray data collection and structure determination statistics

CsRecQ / DNA	
Data Collection	
Wavelength, Å	0.97872
Resolution Range (high resolution bin), Å	50-2.40 (2.44-2.40)
Space Group	P2 ₁ 2 ₁ 2
Unit Cell (a, b, c (Å))	78.3, 94.0, 100.2
(α , β , γ (°))	90, 90, 90
Completeness, %	99.9 (100)
Total/Unique Reflections	225934/29764
Redundancy	7.6 (7.7)
$\langle I/\sigma I \rangle$	27.7 (4.0)
R_{sym}^{\dagger} , %	9.1 (52.8)
Refinement	
Resolution, Å	40-2.40
$R_{\text{work}}/R_{\text{free}}$, %	18.4/23.6
Rms deviations	
Bonds, Å	0.0087
Angles, °	1.22
Ramachandran statistics, %	
Most favored	92.2
Allowed	7.6
Generously allowed	0.2
Disallowed	0.0
# atoms	
Protein	4032
DNA	594
Solvent	303
Zn ²⁺	1
$\langle B \text{ factor} \rangle$, Å ²	
Protein	46.53
DNA	96.07
Solvent	47.88
Zn ²⁺	21.00

$R_{\text{sym}}^{\dagger} = \frac{\sum_j |I_j - \langle I \rangle|}{\sum_j I_j}$, where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

$R_{\text{work}}/R_{\text{free}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$, where the working and free R factors are calculated by using the working and free reflection sets, respectively.

The free R reflections (5% of the total) were held aside throughout refinement.

Table 4.2. DNA-independent and -dependent ATPase activity

EcRecQ variant	Function	k_{\min}^{DNA} (min⁻¹)	k_{\max}^{DNA} (min⁻¹)	Stimulation
WT		6.5 ± 0.5	1260 ± 120	194
Trp154Leu (22)	in ARL	13.4 ± 0.9	850 ± 30	63
Phe158Leu (22)	in ARL	86.3 ± 8.7	610 ± 50	7
Arg159Leu (22)	in ARL	11.9 ± 0.3	1170 ± 20	98
His156Ala	in ARL	43.9 ± 1.1	1610 ± 20	37
Glu124Ala	ARL movement	NT	NT	
Asp191Ala	ARL movement	NT	NT	
His149Ala	His-Gln interaction	NT	NT	
Gln322Ala	His-Gln interaction	0.7 ± 0.4	9 ± 0.3	13
Arg446Ala	dsDNA binding	NT	NT	
Trp347Ala	ssDNA binding	NT	NT	
Arg246Ala	ssDNA binding	NT	NT	
Arg275Ala	ssDNA binding	NT	NT	
Thr293Ala	ssDNA binding	NT	NT	
Arg315Ala	ssDNA binding	NT	NT	

NT = not tested yet

Table 4.3. Binding to DNA substrate F-3prOH

EcRecQ variant	Function	K_{d, app} F-3prOH (nM)
WT		0.4 ± 0.1
Trp154Leu (22)	in ARL	0.8 ± 0.2
Phe158Leu (22)	in ARL	0.5 ± 0.1
Arg159Leu (22)	in ARL	1.7 ± 0.3
His156Ala	in ARL	1.5 ± 0.3
Glu124Ala	ARL movement	NT
Asp191Ala	ARL movement	NT
His149Ala	His-Gln interaction	NT
Gln322Ala	His-Gln interaction	7.4 ± 1.2
Arg446Ala	dsDNA binding	NT
Trp347Ala	ssDNA binding	NT
Arg246Ala	ssDNA binding	~50-100
Arg275Ala	ssDNA binding	1.2 ± 0.1
Thr293Ala	ssDNA binding	NT
Arg315Ala	ssDNA binding	NT

NT = not tested yet

References

1. Singleton MR, Dillingham MS, & Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annual review of biochemistry* 76:23-50.
2. Gorbalenya AE & Koonin EV (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Current Opinion in Structural Biology* 3(3):419-429.
3. Manthei KA & Keck JL (2013) The BLM dissolvasome in DNA replication and repair. *Cellular and molecular life sciences : CMLS*.
4. Nakayama H, Nakayama K, Nakayama R, Irino N, Nakayama Y, & Hanawalt PC (1984) Isolation and genetic characterization of a thymineless death-resistant mutant of *Escherichia coli* K12: identification of a new mutation (recQ1) that blocks the RecF recombination pathway. *Molecular & general genetics : MGG* 195(3):474-480.
5. German J (1993) Bloom syndrome: a mendelian prototype of somatic mutational disease. *Medicine (Baltimore)* 72(6):393-406.
6. Yu CE, Oshima J, Fu YH, Wijsman EM, Hisama F, Alisch R, Matthews S, Nakura J, Miki T, Ouais S, Martin GM, Mulligan J, & Schellenberg GD (1996) Positional cloning of the Werner's syndrome gene. *Science* 272(5259):258-262.
7. Kitao S, Shimamoto A, Goto M, Miller RW, Smithson WA, Lindor NM, & Furuichi Y (1999) Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. *Nat Genet* 22(1):82-84.
8. Morozov V, Mushegian AR, Koonin EV, & Bork P (1997) A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem Sci* 22(11):417-418.
9. Bernstein DA, Zittel MC, & Keck JL (2003) High-resolution structure of the *E.coli* RecQ helicase catalytic core. *EMBO J* 22(19):4910-4921.
10. Pike AC, Shrestha B, Popuri V, Burgess-Brown N, Muzzolini L, Costantini S, Vindigni A, & Gileadi O (2009) Structure of the human RECQL1 helicase reveals a putative strand-separation pin. *Proc Natl Acad Sci U S A* 106(4):1039-1044.
11. Kitano K, Kim SY, & Hakoshima T (2010) Structural basis for DNA strand separation by the unconventional winged-helix domain of RecQ helicase WRN. *Structure* 18(2):177-187.
12. Vindigni A, Marino F, & Gileadi O (2010) Probing the structural basis of RecQ helicase function. *Biophysical chemistry* 149(3):67-77.

13. Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, & Wigley DB (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* 97(1):75-84.
14. Korolev S, Hsieh J, Gauss GH, Lohman TM, & Waksman G (1997) Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell* 90(4):635-647.
15. Subramanya HS, Bird LE, Brannigan JA, & Wigley DB (1996) Crystal structure of a DExx box DNA helicase. *Nature* 384(6607):379-383.
16. Dillingham MS, Soultanas P, Wiley P, Webb MR, & Wigley DB (2001) Defining the roles of individual residues in the single-stranded DNA binding site of PcrA helicase. *Proc Natl Acad Sci U S A* 98(15):8381-8387.
17. Dillingham MS, Soultanas P, & Wigley DB (1999) Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Res* 27(16):3310-3317.
18. Soultanas P, Dillingham MS, Velankar SS, & Wigley DB (1999) DNA binding mediates conformational changes and metal ion coordination in the active site of PcrA helicase. *J Mol Biol* 290(1):137-148.
19. Soultanas P, Dillingham MS, Wiley P, Webb MR, & Wigley DB (2000) Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism. *Embo J* 19(14):3799-3810.
20. Brosh RM, Jr. & Matson SW (1996) A partially functional DNA helicase II mutant defective in forming stable binary complexes with ATP and DNA. A role for helicase motif III. *The Journal of biological chemistry* 271(41):25360-25368.
21. Brosh RM, Jr. & Matson SW (1997) A point mutation in Escherichia coli DNA helicase II renders the enzyme nonfunctional in two DNA repair pathways. Evidence for initiation of unwinding from a nick in vivo. *The Journal of biological chemistry* 272(1):572-579.
22. Zittel MC & Keck JL (2005) Coupling DNA-binding and ATP hydrolysis in *Escherichia coli* RecQ: role of a highly conserved aromatic-rich sequence. *Nucleic Acids Res* 33(22):6982-6991.
23. Hayward S & Berendsen HJ (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* 30(2):144-154.
24. Holm L, Kaariainen S, Rosenstrom P, & Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* 24(23):2780-2781.

25. Kim JL, Morgenstern KA, Griffith JP, Dwyer MD, Thomson JA, Murcko MA, Lin C, & Caron PR (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* 6(1):89-100.
26. Machius M, Henry L, Palnitkar M, & Deisenhofer J (1999) Crystal structure of the DNA nucleotide excision repair enzyme UvrB from *Thermus thermophilus*. *Proc Natl Acad Sci U S A* 96(21):11717-11722.
27. Nakagawa N, Sugahara M, Masui R, Kato R, Fukuyama K, & Kuramitsu S (1999) Crystal structure of *Thermus thermophilus* HB8 UvrB protein, a key enzyme of nucleotide excision repair. *Journal of biochemistry* 126(6):986-990.
28. Morrical SW, Lee J, & Cox MM (1986) Continuous association of *Escherichia coli* single-stranded DNA binding protein with stable complexes of recA protein and single-stranded DNA. *Biochemistry* 25(7):1482-1494.
29. Bernstein DA & Keck JL (2003) Domain mapping of *Escherichia coli* RecQ defines the roles of conserved N- and C-terminal regions in the RecQ family. *Nucleic Acids Res* 31(11):2778-2785.
30. Umezu K, Nakayama K, & Nakayama H (1990) *Escherichia coli* RecQ protein is a DNA helicase. *Proc Natl Acad Sci U S A* 87(14):5363-5367.
31. Svergun D, Barberato C, & Koch MHJ (1995) CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography* 28(6):768-773.
32. Yao N, Hesson T, Cable M, Hong Z, Kwong AD, Le HV, & Weber PC (1997) Structure of the hepatitis C virus RNA helicase domain. *Nature structural biology* 4(6):463-467.
33. Harmon FG & Kowalczykowski SC (2001) Biochemical characterization of the DNA helicase activity of the *Escherichia coli* RecQ helicase. *J Biol Chem* 276(1):232-243.
34. Xu HQ, Deprez E, Zhang AH, Tauc P, Ladjimi MM, Brochon JC, Auclair C, & Xi XG (2003) The *Escherichia coli* RecQ helicase functions as a monomer. *The Journal of biological chemistry* 278(37):34925-34933.
35. Li N, Henry E, Guiot E, Rigolet P, Brochon JC, Xi XG, & Deprez E (2010) Multiple *Escherichia coli* RecQ helicase monomers cooperate to unwind long DNA substrates: a fluorescence cross-correlation spectroscopy study. *J Biol Chem* 285(10):6922-6936.
36. Pan BY, Dou SX, Yang Y, Xu YN, Bugnard E, Ding XY, Zhang L, Wang PY, Li M, & Xi XG (2010) Mutual inhibition of RecQ molecules in DNA unwinding. *J Biol Chem* 285(21):15884-15893.

37. Otwinowski Z & Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology*, eds Carter CW, Jr. & Sweet RM (Academic Press, New York), Vol 276, pp 307-326.
38. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.
39. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
40. Winn MD, Isupov MN, & Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 1):122-133.
41. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, & Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* 66(Pt 2):213-221.
42. Putnam CD, Hammel M, Hura GL, & Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* 40(3):191-285.
43. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, & Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36:1277-1282.
44. Franke D & Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* 42:342-346.
45. Kozin MB & Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* 34:33-41.
46. Volkov VVaS, D.I. (2003) Uniqueness of ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* 36:860-864.
47. Schneidman-Duhovny D, Hammel M, & Sali A (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res* 38(Web Server issue):W540-544.
48. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, & Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* 7:539.

49. Ellis NA, Groden J, Ye TZ, Straughen J, Lennon DJ, Ciocchi S, Proytcheva M, & German J (1995) The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* 83(4):655-666.
50. German J, Sanz MM, Ciocchi S, Ye TZ, & Ellis NA (2007) Syndrome-causing mutations of the BLM gene in persons in the Bloom's Syndrome Registry. *Hum Mutat* 28(8):743-753.
51. Foucault F, Vaury C, Barakat A, Thibout D, Planchon P, Jaulin C, Praz F, & Amor-Gueret M (1997) Characterization of a new BLM mutation associated with a topoisomerase II alpha defect in a patient with Bloom's syndrome. *Hum Mol Genet* 6(9):1427-1434.
52. Barakat A, Ababou M, Onclercq R, Dutertre S, Chadli E, Hda N, Benslimane A, & Amor-Gueret M (2000) Identification of a novel BLM missense mutation (2706T>C) in a Moroccan patient with Bloom's syndrome. *Hum Mutat* 15(6):584-585.
53. Guo RB, Rigolet P, Ren H, Zhang B, Zhang XD, Dou SX, Wang PY, Amor-Gueret M, & Xi XG (2007) Structural and functional analyses of disease-causing missense mutations in Bloom syndrome protein. *Nucleic Acids Res* 35(18):6297-6310.

CHAPTER 5

Summary, Conclusions, and Future Directions

The importance of the family of RecQ DNA helicases in genome maintenance has been demonstrated from bacteria to humans. As they function in DNA replication, recombination, and repair processes, RecQs collaborate with other proteins, forming conserved protein complexes. In particular, *Escherichia coli* RecQ (EcRecQ) functionally interacts with Topo III, and this is conserved through humans, as one of five human RecQ proteins, BLM, interacts with Topo III α . Furthermore, BLM and Topo III α form a complex with the RMI (RecQ mediated genome instability) subcomplex (comprised of RMI1 and RMI2), which together is referred to as the BLM dissolvasome. This complex has been shown to be important in many aspects of genome maintenance, as is detailed in Chapter 1.

The importance of the RMI subcomplex. In Chapter 2, I described a structural approach to further understand the role of the RMI subcomplex. I solved the x-ray crystal structure of the RMI core complex, composed of RMI1 OB2 and RMI2. The interface between the two OB folds in the RMI core complex looks remarkably similar to two-thirds of the RPA (replication protein A) trimerization domain; however, the RMI proteins are unable to bind DNA. Through a collaboration with Weidong Wang at the NIH, co-immunoprecipitations were used to test the importance of critical residues at the interface and sister chromatid exchange (SCE) levels examined the functional importance of the interface in cells. These studies revealed the importance of RMI subcomplex complex formation in cells and showed that a functional complex is critical in genome maintenance.

Furthermore, we observed one intriguing inconsistency between the co-immunoprecipitation results and SCE assay. The RMI2 Lys121Ala variant co-immunoprecipitated with RMI1, Topo III α , and BLM, however, in the SCE assay, this variant led to high levels of SCEs, similar to *rmi2*^{-/-} cells. Due to the location of RMI2 Lys121 near the surface of the complex, we hypothesized that this residue may have an importance outside of RMI subcomplex formation. One possibility was that this residue may be important for another protein to interact with the RMI subcomplex, which we confirmed in Chapter 3.

The RMI subcomplex interacts with FANCM at the RMI1/RMI2 interface. Prior to my work on the RMI/FANCM complex, it was published that a 34 amino acid region of FANCM, termed MM2, was sufficient to interact with RMI1. In Chapter 3, I developed a method to purify the MM2 peptide from FANCM in order to study the interaction between FANCM and the RMI subcomplex. I then used this peptide in fluorescence polarization (FP) studies: the RMI core complex interacted with the MM2 peptide with low nanomolar affinity, while the core complex with RMI2 Lys121Ala was ≥ 80 -fold weaker in interacting with MM2. This confirmed the hypothesis that the RMI2 Lys121 was important in FANCM interaction, and showed that the RMI1/RMI2 interface was most probable interaction site for FANCM.

To better understand the how FANCM was interacting with the RMI subcomplex, I then crystallized the RMI core complex with the MM2 peptide. With this crystal structure, I was able to fully describe the interface between these three proteins and conclude that RMI2 Lys121 was important, as it moved back to form a hydrophobic pocket for FANCM to dock. I developed

a competition FP assay to show that mutating individual hydrophobic residues that dock on the RMI1/2 surface destabilized the interface to varying degrees. We continued our collaboration with Weidong Wang, and showed a similar pattern in co-immunoprecipitations and SCE assays. Through these studies, we showed that mutating residues at the RMI/FANCM interface lead to SCE levels similar to *rmi2*^{-/-} cells, indicating that the formation of this complex is important for SCE suppression in cells. Therefore, the role of the BLM dissolvasome in SCE suppression appears to be coordinated with FANCM, as described in Chapters 1 and 3.

Future directions: a small-molecule screen to identify RMI/FANCM inhibitors. Since the RMI/FANCM interface is critical in genomic stability, our next step is to indentify potential small-molecule inhibitors. Using the competition FP assay, I will use a high throughput screen to identify small molecules that are capable of competing with the MM2 peptide. We also plan to use the crystal structure of the RMI/FANCM complex to perform an *in silico* screen to identify other potential small molecules. Small molecules that are identified in our screen will be further tested for their ability to interact with the RMI core complex and tested for their ability to disrupt the complex *in vivo*. Furthermore, structural studies with identified small molecules will allow us to modify inhibitors in order to strengthen their interaction with the RMI core complex. Further collaborations will allow us to test if these inhibitors could be useful as chemotherapeutics, since cancer cells are often more sensitive to DNA damage.

The structure of a bacterial RecQ in complex with DNA. While the earlier work in my thesis was focused on the BLM dissolvasome, our laboratory has long been interested in the general

mechanisms of RecQ helicases. In 2003 the laboratory revealed the first structure of a RecQ DNA helicase by solving the x-ray crystal structure of the catalytic core of EcRecQ. However, many questions still remained on the mechanism of DNA unwinding, especially since certain elements like the aromatic rich loop (ARL) were not in position to contact the DNA, even though they had been shown to be important biochemically. Therefore, I focused on obtaining a structure of a RecQ helicase in complex with DNA, and was successful with the RecQ from *Chronobacter sakazakii* (CsRecQ), as described in Chapter 4. This structure revealed movement of the ARL so that it stacks within the ssDNA, which is likely important for all RecQ helicases. Furthermore, we observed movement of the winged-helix (WH) domain between the apo EcRecQ and DNA-bound CsRecQ protein. These studies have allowed us to develop a better understanding of the mechanism by which bacterial, and likely all RecQs unwind DNA.

Future studies with the bacterial RecQs will initially focus on biochemical experiments as described in Chapter 4 to further understand the crystallographic model we have obtained. Following those studies, the structure has lead to many new questions that will be addressed in future studies in our lab. One area of interest is the WH domain movement, and while SAXS has been used to confirm this movement (Chapter 4), we also plan to use SAXS to study potential WH domain movement in eukaryotic RecQs, such as RecQ1 and BLM. Other studies may focus on observing this DNA-induced domain movement in solution, such as through FRET studies, or by locking the domain in the apo orientation through cysteine crosslinking to see if unwinding is perturbed.

APPENDIX 1

Structural and Kinetic Isotope Effect Studies of Nicotinamidase (Pnc1) from *S. cerevisiae*

This work has been published:

Smith BC, Anderson, MA, Hoadley KA, Keck JL, Cleland WW, Denu JM. (2012) “Structural and Kinetic Isotope Effect Studies of Nicotinamidase (Pnc1) from *S. cerevisiae*”. *Biochemistry* **51**, 243-256.

Kelly A. Manthei (nee Hoadley) performed protein purification and structure determinations.

Abstract

Nicotinamidases catalyze the hydrolysis of nicotinamide to nicotinic acid and ammonia. Nicotinamidases are absent in mammals but function in NAD⁺ salvage in many bacteria, yeast, plants, protozoa, and metazoans. We have performed structural and kinetic investigations of the nicotinamidase from *S. cerevisiae* (Pnc1). Steady-state product inhibitor analysis revealed an irreversible reaction where ammonia is the first product released, followed by nicotinic acid. A series of nicotinamide analogs acting as inhibitors or substrates were examined revealing that the nicotinamide carbonyl oxygen and ring nitrogen are critical for binding and reactivity. X-ray structural analysis revealed a covalent adduct between nicotinaldehyde and Cys167 of Pnc1 and coordination of the nicotinamide ring nitrogen to the active-site zinc ion. Using this structure as a guide, the function of several residues was probed via mutagenesis and primary ¹⁵N and ¹³C kinetic isotope effects (KIE) on V/K for amide bond hydrolysis. The KIE values of almost all variants were increased indicating that C-N bond cleavage is at least partially rate limiting; however, a decreased KIE for D51N was observed indicative of a higher commitment to catalysis. In addition, KIE values using slower alternate substrates indicated that C-N bond cleavage is at least partially rate limiting with nicotinamide to highly rate limiting with thionicotinamide. A detailed mechanism is discussed involving nucleophilic attack of Cys167, followed by elimination of ammonia and then hydrolysis to liberate nicotinic acid. These results will aid design of mechanism-based inhibitors to target pathogens that rely on nicotinamidase activity.

Introduction

Nicotinamidases (EC 3.5.1.19) are amidohydrolases that catalyze the hydrolysis of nicotinamide to nicotinic acid and ammonia (Scheme A1.1). Nicotinamidases play a central role in the NAD⁺ salvage pathway (1) of multiple species of bacteria, yeast (2), protozoa (3), and plants (4) and are present in many metazoans such as *Drosophila melanogaster* (5) and *Caenorhabditis elegans* (6, 7). However, mammals do not encode a nicotinamidase but instead use nicotinamide phosphoribosyltransferase to convert nicotinamide directly to nicotinamide mononucleotide, which is then recycled to NAD⁺ (8).

The importance of nicotinamidase activity in the NAD⁺ salvage pathways of human pathogens, combined with the absence of a nicotinamidase in human NAD⁺ salvage pathways, suggests that nicotinamidases are potential drug targets. Indeed, nicotinamidase was shown to be an essential enzyme for the infectious phenotype of *Borrelia burgdorferi*, the bacteria that cause Lyme disease (9). In *Brucella abortus*, which causes abortion in domestic animals and undulant fever in humans, the *B. abortus* nicotinamidase is essential for replication (10). Furthermore, erythrocytes infected with *Plasmodium falciparum*, a parasite that causes malaria in humans, displayed increased nicotinamidase activity and NAD⁺ synthesis (3). It is likely that other human pathogens require nicotinamidase activity for viability, because many, including *P. falciparum*, do not possess the genes necessary for *de novo* NAD⁺ synthesis.

Current tuberculosis treatments target the nicotinamidase (PncA) of *Mycobacterium tuberculosis*, which hydrolyzes the prodrug pyrazinamide to the active form, pyrazinoic acid. Pyrazinamide,

when administered in combination with isoniazid and rifampin, forms the current short-course treatment recommended by the World Health Organization (11) and shortens tuberculosis treatment from 9- to 6-months. Pyrazinoic acid displays its toxicity by inhibiting *M. tuberculosis* trans-translation through binding the ribosomal protein S1 (12). Mutations in the *M. tuberculosis* *PNCA* gene are associated with clinical resistance to pyrazinamide (13-16).

Nicotinamidases have also been implicated in increasing the lifespan of *Saccharomyces cerevisiae* (17-20), *C. elegans* (7), and *D. melanogaster* (5). The observed lifespan extension in these organisms is potentially mediated through increasing the activity of sirtuin NAD⁺-dependent deacetylases by decreasing cellular nicotinamide levels and increasing NAD⁺ levels. In support of this hypothesis, sirtuin overexpression was reported to increase lifespan in *S. cerevisiae* (21, 22), *C. elegans* (23), and *D. melanogaster* (24, 25) and nicotinamide is a potent sirtuin product inhibitor (19, 26-28). However, a recent report revealed that the apparent lifespan extension through sirtuin overexpression in *C. elegans* and *D. melanogaster* was abolished when a more appropriate genetic background was used (29).

Despite the importance of nicotinamidases in diverse biological processes, their precise mechanism of catalysis has yet to be fully elucidated (30-33). Understanding the chemical mechanism and nature of the transition state in the reaction catalyzed by nicotinamidases would aid the design of inhibitors or prodrugs (such as pyrazinamide) that target nicotinamidases for anti-microbial applications. Here, we establish the overall kinetic mechanism of the eukaryotic nicotinamidase from *S. cerevisiae* (Pnc1) through product inhibition analysis. We then show that

ketone and aldehyde containing nicotinamide analogs are Pnc1 inhibitors. Using this knowledge we obtained the first crystal structure of a eukaryotic nicotinamidase with a nicotinamide analog bound in the active site. This structure suggested several residues potentially involved in catalysis, and the steady-state kinetic parameters of several Pnc1 mutants were determined. Steady-state kinetic parameters with alternate nicotinamide analog substrates were also measured. We then further delineated the mechanism by determining the primary ^{15}N and ^{13}C kinetic isotope effects (KIE) of the C-N bond breaking and the pH dependence of the reaction. To determine if C-N bond cleavage is partially or fully rate limiting, we compared the ^{15}N and ^{13}C KIE for nicotinamide hydrolysis by Pnc1 with those determined using slow substrate analogs. In addition, the KIE values for nicotinamide hydrolysis by several Pnc1 mutants were compared, allowing for the validation of the proposed catalytic function of active-site residues. These results suggest specific roles for several Pnc1 residues during catalysis as well as structure activity relationships for Pnc1 catalyzed hydrolysis of nicotinamide analogs.

Experimental Procedures

Expression and purification of Pnc1. Yeast Pnc1 cloned into pET-16b or pET-17b (20, 32) was transformed into *E. coli* strain BL21(DE3). A single transformant colony was grown in 1.0 L 2XYT medium at 37°C to OD_{600nm} ~0.6-0.8. IPTG was added to 0.5 mM, and expression was continued for 6-8 hr at 25°C. Cells were harvested by centrifugation and stored at -20 °C. Cellular pellets were resuspended in buffer A (50 mM NaH₂PO₄, pH 7.5, 300 mM NaCl, 1 mM β-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride) containing 30 mM imidazole and lysed by sonication. Recombinant Pnc1 was purified using immobilized metal affinity chromatography with a Ni²⁺-nitrilotriacetic column. The column was washed with buffer A containing 30 mM imidazole and bound proteins were eluted with a gradient of 30 to 500 mM imidazole in buffer A at pH 7.5. Pooled fractions containing Pnc1 were concentrated and protein concentrations were determined using the Bradford method (34) using bovine serum albumin as a standard.

Mutagenesis of Pnc1. Pnc1 mutagenesis was performed using the QuikChange Site-Directed Mutagenesis kit (Stratagene) according to the manufacturers protocol. The primers used for mutagenesis are included in Table A1.1.

Determination of kinetic parameters. Nicotinamidase activity was measured continuously using an enzyme-coupled assay with glutamate dehydrogenase using a Multiskan Ascent microplate reader (LabSystems; Franklin, MA, USA). This assay is slightly modified from that described by Su *et al.* (35). Typical assay mixtures contained 1.25 μM to 3.2 mM nicotinamide

or analog, 0.2 mM NADPH, 3.3 mM α -ketoglutarate, 50 nM to 10 μ M Pnc1 WT or mutant, and 3 units of glutamate dehydrogenase from bovine liver in 50 mM sodium phosphate at pH 7.5. Assays were carried out in a final volume of 300 μ L per well in a clear, flat-bottomed, 96-well plate. All assay components except Pnc1 were preincubated at 25 °C for 5 min or until absorbance at 340 nm stabilized, and the reaction was initiated by the addition of Pnc1. The rates were monitored continuously for NADPH consumption at 340 nm. Rates were determined from the slopes of the initial linear portion of each curve using an extinction coefficient for NADPH of $6.22 \text{ mM}^{-1}\text{cm}^{-1}$ and a pathlength of 0.9 cm for 300 μ L reactions. The background rates of reactions lacking Pnc1 resulting from the spontaneous formation of ammonia were subtracted from the initial velocities of the Pnc1-catalyzed reactions.

Determination of K_i values. Assay mixtures contained 6.7 μ M to 200 μ M nicotinamide, 0.2 mM NADPH, 1 or 3.3 mM α -ketoglutarate, 100 nM to 0.5 μ M Pnc1 WT, and 3 units of glutamate dehydrogenase from bovine liver in 50 mM sodium phosphate at pH 7.5. For nicotinic acid inhibition 200 μ M to 1.2 mM nicotinic acid was used. For nicotinaldehyde inhibition 1 μ M to 6 μ M was used. For 3-acetylpyridine inhibition 300 μ M to 1.8 mM was used. Benzaldehyde and pyrazinoic acid were initially dissolved in DMSO and 5 to 10 mM was used (10% final v/v DMSO). Assays were run and analyzed as detailed under determination of kinetic parameters. Initial velocity data were fitted in Kinetasyst (Intellikinetics, State College, PA) to competitive inhibition patterns (eq 1 or 2) based on the algorithms defined by Cleland (36). All data were displayed using Kaleidagraph (Synergy Software, Reading, PA).

$$v = \frac{V_{\max} [S]}{K_m \left(1 + \frac{[I]}{K_i} \right) + [S]} \quad (1)$$

$$\text{Log}(v) = \text{Log} \frac{V_{\max} [S]}{K_m \left(1 + \frac{[I]}{K_i} \right) + [S]} \quad (2)$$

Protein crystallization and structure determination. For crystallization trials, the protein was purified as described above, with the addition of a Sephacryl S-300 size-exclusion step. Protein was dialyzed against buffer containing 15 mM Tris pH 7.5, 50 mM NaCl, 4 mM MgCl₂, 10 mM Na citrate and 5% glycerol and the inhibitor nicotinaldehyde was added at a 4:1 molar ratio. The protein (5 mg/ml) was mixed with mother liquor (1.6 M NaOAc, 10% ethylene glycol, 0.1 M HEPES pH 7.4) at a 1:1 (vol) ratio. Crystals were formed by hanging drop vapor diffusion. Crystals were transferred to a cryoprotectant solution (1.5 M NaOAc, 20% ethylene glycol, 0.1 M HEPES pH 7.4) and flash-frozen in liquid nitrogen.

Diffraction data were indexed and scaled using HKL2000 (37). The structure of Pnc1p with the inhibitor nicotinaldehyde was solved by molecular replacement (Phaser) (38) using the apo Pnc1p structure (32) as a search model. The structure was improved by rounds of manual fitting using Coot (39) and refinement using REFMAC5 (40). Coordinate and structure factor files have been deposited in the Protein Data Bank (PDB ID 3V8E).

¹⁵N kinetic isotope effects. Reactions for isotope effect analysis were carried out in 7.5–10 mL of 20 mM potassium phosphate (pH 7.5) containing 10 mM nicotinamide or analog and 0.75–36

μ M Pnc1 WT or mutant. Reactions were quenched by addition of 8 M HCl to a final concentration of 40 mM. The enzyme was removed by Amicon filtration (10,000 MWCO) and the solution was diluted to 100 mL with distilled water. The pH was adjusted to 6 and the solution was loaded onto an AG1X8 resin column (Cl^- form, 2.5 cm x 30 cm) at 1 mL/min. Fraction collection (8.5 mL/fraction) was started at the time of loading. When the loading was complete, the products were eluted with water at 1 mL/min. NH_4Cl eluted between fractions 3-18 and was detected using 50 μ L samples of each fraction added to 50 μ L of Nessler's reagent in 96-well plates. A yellow color indicated the presence of NH_4Cl . The residual nicotinamide or analog generally eluted between fractions 25-50 and was detected by spotting on UV active TLC plates. The product, nicotinic acid, was eluted with 1 M HCl from the column and found by UV at 262 nm. Product NH_4Cl was reduced to 50 mL by rotary evaporation and purified by steam distillation with 12 mL of 13 M KOH. The NH_3 was trapped in 10 mL of 100 mM H_2SO_4 and collected until the total volume reached 50 mL.

The residual substrate was pooled and concentrated to 50 mL. It was then steam distilled at high pH, hydrolyzing the amide, and the resultant NH_3 was trapped in acid. The hydrolysis of nicotinamide (40-50 μ moles) required the addition of 12 mL of 13N NaOH to a 50 mL solution. The other analogs required 13N NaOH in the following amounts: 10 mL for thionicotinamide, 8 mL for 5-methylnicotinamide, and 3 mL for pyrazinamide. Considerable time was spent determining the correct amount of base to be used since too much caused the decomposition of the pyridine ring and production of non-amide NH_3 . This was especially true for pyrazinamide where the pyrazine is not as stable as pyridine.

The ammonia concentration for both the product and residual substrate samples was determined by UV-Vis spectrometry using Nessler's reagent to establish that no product NH_4Cl had been lost. The $(\text{NH}_4)_2\text{SO}_4$ was then reduced to ~ 1 mL by rotary evaporation and transferred to a flask with a side arm. The side arm was filled with ~ 4 mL of NaOBr . The system was sealed with a stopcock and the solution was frozen at -78°C and taken through three freeze-pump-thaw cycles to remove all gases. After the final thaw, the two solutions were mixed slowly and the liquid was again frozen at -78°C . The freshly produced N_2 was distilled through two -78°C traps, one -196°C trap, and then collected on molecular sieves at -196°C . The N_2 was analyzed by IRMS to give the $^{15}\text{N}/^{14}\text{N}$ ratio.

^{13}C kinetic isotope effects. This methodology is a modification of previous work by Scott and others (41-43) (Scheme A1.2). The solution of nicotinic acid product was reduced to dryness and dissolved in 2 mL of HPLC grade methanol. The methanol/nicotinic acid solution was added to a quartz glass tube (25cm, 9mm o.d., 7mm i.d.) and the methanol was removed under vacuum at slightly elevated temperature (40°C) for several hours. To this, 0.5 g of dry copper chromite was added and the tube was evacuated and flame sealed. The sample was placed in a furnace and heated at 250°C for nicotinic acid (232°C for pyrazinoic acid, 235°C for thionicotinic acid, 240°C for 5-methylnicotinic acid), for 3 hours and then cooled in a room temperature water bath. The tube was cracked and the CO_2 was distilled through two -130°C liquid N_2 /pentane traps and finally collected in a third trap at -196°C . Liquid N_2 /pentane traps were necessary because -78°C traps were inefficient and allowed pyridine, a product of the

decarboxylation of nicotinic acid, to contaminate the CO₂ sample. The collected CO₂ (~80% of the total sample) was isolated in the trap and was not removed from the line at this time.

The final 20% of the CO₂ was collected by acidification of the copper chromite powder. The copper chromite and a small stir bar were placed in a 30 mL flask equipped with a glass joint, side arm with a stopcock, and septum. The flask was fitted to a glass stopcock and placed on the high vacuum line and evacuated for at least 15 min. After closing the top stopcock, 2.5 mL of 1M H₂SO₄ was added via syringe through the septum. CO₂ was evolved immediately and the solution was stirred for 5 min. A -78 °C bath was used to freeze the sample and the CO₂ was collected in the same manner as above and trapped with the first portion of gas. The entire gas sample was isolated and the -196 °C trap was removed and replaced with a -78 °C trap. The entire CO₂ sample was collected in a gas sample tube at -196 °C and was then analyzed by IRMS. IRMS analysis of the two CO₂ portions individually showed a difference of 1-1.5 δ so collection of both was deemed necessary for the most accurate measurement.

When this technique was used to decarboxylate thionicotinic acid the delta values were not reproducible. It was assumed that the sulfur had not been completely removed from the compound, even though it was heavily acidified. Also these samples were not easily evacuated from the IRMS and a small residual impurity had to be frozen out of the system at -196 °C onto molecular sieves. Thus we were unable to reliably determine the ¹³C KIE for thionicotinamide.

Fraction of Reaction. The fraction of reaction (f value) was determined by using a known amount of substrate for the enzymatic reaction (50 μ moles) and then comparing that to the amount of product NH_3 recovered after the reaction. To double check the f value the amount of product NH_3 was also compared to the amount of NH_3 recovered in 0.1 N H_2SO_4 after steam distillation under basic conditions of the residual substrate. In all cases the amount of NH_3 was determined by comparison of the sample to a standardized curve of NH_4Cl by UV at 425 nm. The fraction of reaction determined for the NH_3 _{prod} vs the initial substrate concentration and the fraction of reaction for NH_3 _{prod} vs NH_3 _{res sub} were generally within 1-3% of one another.

Kinetic isotope effects. Isotope effects were determined from changes in the natural abundance ratio of $^{15}\text{N}/^{14}\text{N}$ (or $^{13}\text{C}/^{12}\text{C}$) in the compound during the reaction. The product and residual substrate are separated, purified, and converted to N_2 (or CO_2) gas. Each gas sample is analyzed individually by IRMS to determine its isotopic ratio compared to a known standard to give δ , defined as:

$$\delta = 1000 \left\{ \left(\frac{^{15}\text{N}_{\text{sample}} / ^{14}\text{N}_{\text{sample}}}{^{15}\text{N}_{\text{standard}} / ^{14}\text{N}_{\text{standard}}} \right) - 1 \right\} \quad (3)$$

To determine the isotope effect the samples are converted to an R value defined as:

$$\text{R} = \left(\frac{\delta_{\text{sample}}}{1000} \right) + 1 \quad (4)$$

R-values for the reaction product, R_p , and residual substrate, R_s , along with the isotopic ratio of the starting material, R_o , and f , the fraction of reaction, are used in the following equations to arrive at the KIE.

$$\text{KIE} = \frac{\ln(1-f)}{\ln(1-fR_p / R_o)} \quad (5)$$

$$\text{KIE} = \frac{\ln(1-f)}{\ln[(1-f)(R_s / R_o)]} \quad (6)$$

$$\text{KIE} = \frac{\ln(1-f)}{\ln\{(1-f)/(1-f + (f(R_p / R_s)))\}} \quad (7)$$

Dependence of activity on pH. Reactions contained 0.1–3 μM Pnc1 and varying concentrations of nicotinamide (2.5 μM to 1 mM) or pyrazinamide (20 μM to 2 mM) in 100 μL at 25 $^\circ\text{C}$. TBA buffer (50 mM Tris, 50 mM BisTris, 100 mM sodium acetate) was used for the pH range 4.0–8.5. ATE buffer (100 mM ACES, 52 mM Tris, 52 mM ethanolamine) was used for the pH range 8.0–10.5. These buffer mixtures are designed to give a constant ionic strength over a wide pH range (44). Reactions were quenched with 20 μL 6% v/v TFA before 10% of the substrate was converted to products at intervals of 60–120 seconds. Percent product conversion was determined spectrophotometrically by coupling to glutamate dehydrogenase (see previous paper) or using a previously published HPLC assay (45). Plots of k_{cat} versus pH were fitted to equation 8:

$$\log v = \log \{C/(1 + H/K_a + K_b/H)\} \quad (8)$$

using KinetAsyst (IntelliKinetics, State College, PA), where C is the pH-independent value, H is the proton concentration, and K_a and K_b are the ionization constants of the groups involved in the reaction.

Results

Product inhibition and irreversibility of reaction. To establish the kinetic mechanism for *S. cerevisiae* nicotinamidase (Pnc1), steady-state product inhibition analysis was performed. Double-reciprocal analysis of nicotinamidase inhibition by nicotinic acid and pyrazinoic acid displayed competitive inhibition patterns with K_i values of 120 μM and 6.7 mM, respectively (Table A1.2; Figure A1.1). The competitive inhibition displayed by nicotinic acid is consistent with that shown by Gadd *et al.* for a nicotinamidase from *Micrococcus lysodeikticus* (46) but in contrast to French *et al.* who observed no inhibition with nicotinic acid for *S. cerevisiae* Pnc1 (30). Double-reciprocal analysis of nicotinamidase inhibition by ammonium ion using a previously published HPLC assay (45) exhibited non-competitive inhibition patterns (data not shown). However, greater than 10 mM ammonium chloride was required for inhibition, consistent with several previous reports on nicotinamidases from *M. lysodeikticus* and yeast (46, 47). Therefore, it was unclear if the ammonium ion added was acting as a product inhibitor or as a denaturant as non-competitive inhibition patterns would be predicted in either case. To determine if the weak inhibition by ammonium ion could be explained by low reversibility of the reaction, 100 mM ammonium chloride and 5 mM nicotinic acid were incubated with 5 μM Pnc1 for 150 minutes at pH 7.5 and assayed for nicotinamide formation using an HPLC assay (45). No nicotinamide was detected under these conditions leading to an upper estimate of the reverse reaction rate at $< 10^{-5} \text{ s}^{-1}$, at least four orders of magnitude slower than the forward reaction rate with nicotinamide ($k_{\text{cat}} = 0.69 \text{ s}^{-1}$). Therefore, the Pnc1 catalyzed reaction is essentially irreversible.

Inhibition by nicotinamide analogs. The product inhibition analysis above suggested that nicotinic acid was the second product released and therefore that an enzyme intermediate might exist between Pnc1 and nicotinic acid. We hypothesized that non-hydrolyzable nicotinamide analogs might trap this intermediate and display potent nicotinamidase inhibition. All analogs tested displayed competitive inhibition with K_i values ranging from high nM to low mM (Table A1.2). The competitive inhibition observed by 3-acetylpyridine ($K_i = 316 \mu\text{M}$; Table A1.2) was consistent with several previous reports on nicotinamidases from *Mycobacterium phlei* (48), *M. lysodeikticus* ($K_i = 160 \mu\text{M}$) (49), *Tortula cremoris* ($K_i = 305 \mu\text{M}$) (50), Fleischmann's yeast ($K_i = 65 \mu\text{M}$) (47), *Flavobacterium peregrinum* (51), and *S. cerevisiae* ($K_i = 46 \mu\text{M}$) (30). Nicotinaldehyde displayed by far the most potent inhibition among the analogs tested with a K_i value of 940 nM (Table A1.2; Figure A1.1), 10-fold below the K_m value of 9.6 μM for nicotinamide. Potent inhibition by nicotinaldehyde was previously observed for nicotinamidases from *M. lysodeikticus* ($K_i = 18 \text{ nM}$) (49), *S. cerevisiae* ($K_i = 1.4 \mu\text{M}$) (30), and *Mycobacterium tuberculosis* ($K_i = 290 \text{ nM}$) (33). Benzaldehyde was the weakest inhibitor tested with a K_i value of 20.6 mM, similar to the previously observed benzaldehyde inhibition of *M. lysodeikticus* nicotinamidase ($K_i = 1.8 \text{ mM}$) (49).

Structure of nicotinaldehyde inhibited Pnc1. The potent inhibition displayed by nicotinaldehyde inspired us to co-crystallize Pnc1 with nicotinaldehyde. Pnc1 crystallized with the nicotinaldehyde inhibitor at a 4:1 (nicotinaldehyde:Pnc1) molar ratio in the same space group and unit cell parameters as the apo structure (Table A1.3) (32). The crystals diffracted to 2.7-Å resolution and the structure was determined by molecular replacement using the apo Pnc1

structure (32) as a search model (38). F_o-F_c electron density maps revealed the location of the inhibitor covalently bound to C167, which was modeled into the electron density as a modified amino acid (Figure A1.2A). Overall, the structure is similar to the apo structure with a root mean square deviation (rmsd) of 0.3 Å², and no major structural changes were induced by complex formation with the inhibitor.

Within the crystal structure, nicotinaldehyde is covalently bound within the Pnc1 active site through a thiohemiacetal linkage to C167. The residues D8, K122, and C167 are all within hydrogen bond distance from one another (Figure A1.2B) and form a putative catalytic triad that is conserved throughout all known nicotinamidases (Figure A1.3). In addition to the covalent linkage to C167, nicotinaldehyde is ligated to the active-site zinc through the pyridine ring nitrogen (Figure A1.2). This zinc ligation is consistent with recent structures of nicotinamidases from *Acinetobacter baumannii* and *Streptococcus pneumoniae* (31, 52). The active-site zinc is also ligated by D51, H53, and H94, consistent with the previous *S. cerevisiae* Pnc1 apo-structure (32). Furthermore, we observed evidence for two water molecules that also coordinate the zinc; one of these is within hydrogen bonding distance of one conformation of E129 (E129 is present in two conformations in the structure) (Figure A1.2C). The carbonyl oxygen of nicotinaldehyde forms hydrogen bonds to the backbone amide nitrogens of A163 and C167 consistent with recent structures of *S. pneumoniae* PncA (52). The pyridine ring of nicotinaldehyde is also bound within an aromatic cage consisting of F13, W91, Y131, and Y166 (Figure A1.2B).

Kinetic parameters of Pnc1 active-site mutants. The structure with nicotinaldehyde bound within the Pnc1 active site and a sequence alignment of prokaryotic and eukaryotic nicotinamidases (Figure A1.3) suggested several conserved residues that may be involved in catalysis. Therefore, we determined the steady-state kinetic parameters for Pnc1 mutants of D8, D51, H53, H94, K122, and C167 (Table A1.4). The putative base D8 was substituted with Ala, Asn, or Glu, yielding mutants that displayed 10^3 – 10^4 fold lower k_{cat} values compared to Pnc1 WT (Table A1.4). The D8E mutant exhibited slightly faster rates compared to either D8A or D8N. The equivalent D8E mutant in *M. tuberculosis* nicotinamidase was previously shown to harbor a 100-fold lower specific activity compared to wild type (53). When individually substituted with Ala, the zinc-binding residues D51, H53, and H94 yielded enzymes with 10-50 fold lower k_{cat} values. Similar losses of specific activity (10–3000 fold) were observed when the corresponding zinc-binding residues were mutated to Ala in the *M. tuberculosis* nicotinamidase (53, 54). The D51N mutant was slightly more active (3-fold) than the D51A mutant. The zinc-binding mutants displayed K_m values that were similar to or lower than those for Pnc1 WT (Table A1.4). The nicotinamidase active-site Zn^{2+} is tightly bound, as addition of 10 mM EDTA did not inhibit Pnc1 (data not shown) consistent with a previous report on nicotinamidases from *T. cremoris* (50), *M. lysodeikticus* (46), and *M. tuberculosis* (33). Addition of 2 mM ZnCl_2 failed to increase activity of wild type (consistent with a previous report (30)) or rescue activity of D51A, D51N, H53A, or H94A mutants (data not shown) further indicating that Zn^{2+} is tightly ligated by D51, H53, and H94 and that these residues are critical for catalysis. When K122 was replaced with Ala, the k_{cat} was decreased by 16-fold, whereas the K122R mutant displayed a dramatic 770-fold drop in k_{cat} compared to Pnc1 WT. When the proposed nucleophile C167 was

replaced with Ala, the observed rate was below the detection limit of the coupled assay (0.0005 s^{-1}), the lowest of any mutant. A complete loss of activity was also observed when the active-site Cys was mutated to Ala in the nicotinamidase from *M. tuberculosis* (53, 54).

Kinetic parameters of nicotinamide analogs as alternate substrates. Kinetic values for the reaction using substrate analogs produced a wide variation in $k_{\text{cat}}/K_{\text{m}}$ values (Table A1.5). The K_{m} value of $9.6\text{ }\mu\text{M}$ observed for nicotinamide was within the range of $6\text{--}10\text{ }\mu\text{M}$ obtained from previous reports on yeast nicotinamidases (30, 47). Substitution of the nicotinamide pyridine ring at the 4-position with nitrogen (pyrazinamide) or addition of a methyl group to the 5-position resulted in 3.7- and 2.5-fold faster k_{cat} values but lower $k_{\text{cat}}/K_{\text{m}}$ values due to 16- and 6-fold higher K_{m} values, respectively. The K_{m} value of $157\text{ }\mu\text{M}$ determined for pyrazinamide was similar to previously determined values for nicotinamidases from *S. cerevisiae* ($K_{\text{m}} = 200\text{ }\mu\text{M}$) (2) and *M. tuberculosis* ($K_{\text{m}} = 300\text{ }\mu\text{M}$) (33). The K_{m} value of $61\text{ }\mu\text{M}$ determined for 5-methylnicotinamide was 5-fold lower than a previously determined value of $360\text{ }\mu\text{M}$ for a nicotinamidase from *M. lysodeikticus* (46). Perturbation of the pyridine nitrogen of nicotinamide by methylation (1-methylnicotinamide) or substitution with carbon (benzamide) resulted in 8400- and 200-fold decreased $k_{\text{cat}}/K_{\text{m}}$ values compared to nicotinamide. The activity observed for 1-methylnicotinamide was in contrast to French *et al.* who observed no activity with *S. cerevisiae* Pnc1 (30). We were unable to accurately determine the k_{cat} value for 1-methylnicotinamide as saturation could not be obtained, but benzamide resulted in a 78-fold lower k_{cat} value compared to nicotinamide. Substitution of the amide oxygen of nicotinamide with a sulfur (thionicotinamide) resulted in a 2300-fold lower $k_{\text{cat}}/K_{\text{m}}$ value. Finally, no activity

could be detected above the detection limit of the HPLC assay (10^{-5} s^{-1}) for nicotinamide mononucleotide (NMN⁺) or NAD⁺.

¹⁵N and ¹³C kinetic isotope effects with substrate analogs. We next investigated the first portion of the nicotinamidase mechanism up to and including the first irreversible step (C-N bond cleavage) using kinetic isotope effects (KIEs). Using slow substrates with Pnc1 allowed us to determine if and when C-N bond cleavage is rate limiting for catalysis. The ¹⁵N and ¹³C KIE values were determined for Pnc1 with nicotinamide and three substrate analogs: 5-methylnicotinamide, pyrazinamide, and thionicotinamide (Table A1.5). With nicotinamide as the substrate the ¹⁵N and ¹³C isotope effects are ~1.2%. Since the upper theoretical limit for ¹⁵N and ¹³C effects is 3-4%, these primary effects indicate that C-N bond cleavage is at least partially rate-limiting. When 5-methylnicotinamide, pyrazinamide, or thionicotinamide were utilized as substrates, all of the KIE are more fully expressed than those with nicotinamide (Table A1.5). Also, in comparing the ¹⁵N effects versus the $k_{\text{cat}}/K_{\text{m}}$ data, there is good correlation between the kinetic data and the KIE showing that the slower substrates exhibit higher primary KIEs. The trend is not as well defined for the ¹³C data, but the KIE for 5-methylnicotinamide is roughly the same as pyrazinamide when the standard error is taken into consideration.

¹⁵N and ¹³C kinetic isotope effects of Pnc1 mutants. The same KIEs determined with the substrate analogs were also determined for four Pnc1 mutants: D8E, D51N, D51A, and K122R, with nicotinamide as the substrate (Table A1.4). These Pnc1 mutants were chosen because of the differences observed in the determined kinetic parameters and their ability to achieve

adequate product conversion to determine KIEs. These KIEs permitted us to investigate whether the mutated residues are responsible for the actions discussed here and by others (30, 31, 33, 52, 55). For the Zn^{2+} ligation mutant, D51N, the KIE values were both lower than for Pnc1 WT, whereas a higher ^{15}N KIE and a lower ^{13}C KIE was observed for D51A compared to Pnc1 WT. For the K122R mutant, the ^{15}N KIE is exactly the same as that determined for the wild-type enzyme while the ^{13}C effect is only ~0.4% higher. Finally, the D8E mutant displayed increased KIEs that were very close to the maximum observed with the substrate analogs.

pH dependence of kinetic parameters. Several residues within the Pnc1 active site may require a particular ionization state for catalysis. Within Pnc1, these include D8, K122, and C167. To provide evidence for the involvement of these residues during catalysis, the effect of pH on the k_{cat} values of Pnc1 was determined. First, the stability of Pnc1 at pH extremes was determined to ensure any changes in rate was due to the protonation states of active-site residues and not global protein unfolding. Determination of Pnc1 k_{cat} values with saturating nicotinamide over a pH range from 4.5 to 10.5 showed no critical ionizations but did show a slight pH-dependent rate change that varied only 4-fold between the pH extremes (0.45 s^{-1} at low pH and 1.87 s^{-1} at high pH) (Figure A1.4A). Similarly, the $k_{\text{cat}}/K_{\text{m}}$ profile did not reveal any critical ionizations (data not shown). The lack of a critical ionization using nicotinamide as the substrate was consistent with previous reports on nicotinamidases from *Torula cremoris* and *M. tuberculosis* (33, 50).

The higher KIE values and slower $k_{\text{cat}}/K_{\text{m}}$ values for pyrazinamide compared to nicotinamide (Table A1.5) indicated that C–N bond cleavage was only partially rate limiting for nicotinamide but nearly fully rate limiting for pyrazinamide. Therefore, we hypothesized that the pH profile for pyrazinamide would display critical ionizations involved in C–N bond cleavage. Indeed, Pnc1 k_{cat} values for pyrazinamide revealed two critical ionizations over a pH range from 4 to 10, one with an apparent $\text{p}K_{\text{a}}$ value of 5.1 ± 0.2 that must be unprotonated for activity and another with an apparent $\text{p}K_{\text{b}}$ value of 9.0 ± 0.2 that must be protonated for activity (Figure A1.4B). The residues responsible for these ionizations are discussed below.

Discussion

Nicotinamidase kinetic mechanism. Although several nicotinamidase chemical mechanisms have been proposed (30-33), the kinetic mechanism had not been investigated in detail. Here, we determined that Pnc1 follows an ordered Uni Bi kinetic mechanism (Figure A1.5) in which nicotinamide binding is followed by formation of a thioester intermediate between C167 and nicotinic acid. Evidence supporting this thioester intermediate is discussed below. Ammonia is released and then the thioester intermediate is hydrolyzed to form nicotinic acid, which is then released. The product inhibition pattern expected for an ordered Uni Bi mechanism is that the second product released acts as a competitive inhibitor versus nicotinamide and the first product released acts as a non-competitive inhibitor versus nicotinamide (56). Therefore, nicotinic acid is the second product released based on the competitive inhibition patterns displayed versus nicotinamide (Figure A1.1). The non-competitive inhibition displayed by ammonium chloride is also consistent with ammonia as the first product released, but we were unable to confirm that the inhibition was specific to the Pnc1 active site or that the inhibition was caused by NH_3 rather than NH_4^+ .

Nicotinamide binding may proceed through a hydrophobic tunnel. The Pnc1 active site is largely sequestered from solvent. However, a small tunnel exists from the active site to the protein surface that is constricted by L20, Y70, and I192 (Figure A1.6A). I192 appears to be especially important in the opening and closing of this tunnel as I192 resides on a dynamic C-terminal loop in Pnc1. The dynamic nature of the C-terminal loop and helix is exemplified by the high B factors of this region in the crystal structure (Figure A1.6B). Furthermore, the C-

terminal helix displayed greater variance than any other portion of the structure when aligned with previous nicotinamidase structures (31, 52, 53, 55) (Figure A1.6C). However, Pnc1 does not undergo significant conformational changes outside of this C-terminal loop and helix upon nicotinamide binding as shown by the low rmsd values (rmsd 0.23–0.38 Å) between the apo- (32) and nicotinaldehyde-bound Pnc1 structures. Once nicotinamide proceeds through this tunnel, binding is stabilized within the active site through edge-face aromatic interactions between the aromatic cage residues F13, W91, Y131, and Y166 and the pyridine ring of nicotinamide (Figure A1.2B).

Zinc is involved in both binding and activation of substrate. The Zn^{2+} ligated by D51, H53, and H94 also stabilizes nicotinamide binding within the active site. Our structure of nicotinaldehyde bound to Pnc1 revealed that the pyridyl nitrogen is ligated to the active-site Zn^{2+} at a distance of 2.6 Å. This Zn^{2+} ligation is important for binding and/or activating nicotinamide for catalysis as mutation of D51, H53, or H94 decreased the k_{cat} value 11- to 49-fold. The importance of the Zn^{2+} ligation is further shown by the alternate substrate benzamide where the pyridyl nitrogen is replaced with a carbon. Using benzamide as an alternate substrate reduced the $k_{\text{cat}}/K_{\text{m}}$ value 200-fold compared to nicotinamide (Table A1.5). This rate difference can partially be explained by the observed 7-fold slower pseudo-first-order rate of benzamide hydrolysis in basic solutions compared to nicotinamide (0.00408 versus 0.0301 hr^{-1} at 25 °C in 0.1 N NaOH) (57). In addition, benzaldehyde was a 22,000-fold less potent inhibitor compared to nicotinaldehyde (Table A1.2) indicating that the pyridyl nitrogen plays a significant role in substrate affinity through Zn^{2+} ligation. Besides nicotinamide binding, the Zn^{2+} also acts as a

Lewis acid to activate nicotinamide towards hydrolysis. Interestingly, this method of Lewis acid catalysis appears unique among zinc hydrolases, which typically facilitate hydrolysis through binding to the carbonyl oxygen of the substrate (58, 59). In nicotinamidases, the Zn^{2+} instead acts as a ‘vinylogous Lewis acid’ activating the amide carbon for nucleophilic attack by C167 through the electron-withdrawing character of Zn^{2+} , which acts through the conjugation of the pyridine ring instead of directly at the amide oxygen.

The Zn^{2+} ligand mutants also displayed the most interesting KIE results. For Pnc1 D51N, both the ^{15}N and ^{13}C KIE are $\sim 0.5\%$ compared to $\sim 1.25\%$ for those of Pnc1 WT (Table A1.4). We believe this is due to an increase in one of the back reaction off rates. A truncated derivation demonstrating how this is possible is included in the supporting information. Another intriguing KIE result is that for the Zn^{2+} ligand mutant, D51A, which has a ^{15}N KIE that is *larger* than the wild type, while the ^{13}C KIE is *smaller*. One feasible explanation is that the carbon is bonded more stiffly than the nitrogen at the transition state in the D51A mutant. This may be due to the inability of the alanine residue in the mutant enzyme to be liganded to the Zn^{2+} , which affects the position of the Zn^{2+} and nicotinamide substrate in the active site compared to the wild type. If true, the orientation of the substrate has changed before the nucleophilic attack by C167, possibly affecting the bond stiffness of carbon to nitrogen in the transition state (Scheme A1.3). One final possibility for the difference in KIE for the D51N and D51A mutants compared to wild type is the change of charge at the active site upon mutation of a negatively charged aspartate residue to the neutral asparagine or alanine. There is precedence for the requirement of certain charges in and around the active site for proper substrate binding and catalysis (60). With Pnc1 it is unclear

how loss of a negative charge may cause changes in the KIE and kinetic data but it may be due to the same factors discussed with the pH profiles.

C167 is activated by Pnc1 active-site residues. Our Pnc1 structure suggested D8, K122, and C167 form a catalytic triad, which resembles the catalytic triad of the nitrilase superfamily (61). These three residues are universally conserved among all nicotinamidases (Figure A1.3) and an active-site Cys residue (C167 in Pnc1) was previously proposed to act as a nucleophile during catalysis (30, 31, 46, 47, 50-52, 55) attacking the carbonyl carbon of nicotinamide to form a thioester intermediate. Here, we present several lines of evidence that are highly consistent with this nucleophilic role of C167. First, nicotinamide analog aldehydes and ketones are competitive inhibitors versus nicotinamide. Although these aldehydes and ketones are predicted to form covalent adducts with C167, classical competitive inhibition patterns and no time dependent loss of activity are expected as reaction with C167 is completely reversible (49, 50, 62) and infinite concentrations of nicotinamide would preclude inhibitor binding to Pnc1. Nicotinaldehyde is a particularly potent inhibitor (K_i value of 940 nM), and the weaker inhibition of 3-acetylpyridine ($K_i = 316 \mu\text{M}$) is consistent with the greater reactivity of aldehydes compared to ketones (63). We co-crystallized nicotinaldehyde with Pnc1 and continuous electron density was observed between C167 and nicotinaldehyde (Figure A1.2A) indicating a covalent thiohemiacetal adduct was present in the crystals. The importance of C167 was confirmed through mutagenesis as the Pnc1 C167A mutant lost all detectable nicotinamidase activity (Table A1.4).

The Pnc1 pH profile using pyrazinamide as a substrate revealed a critical ionization with an apparent pK_a value of 5.1 that must be unprotonated and a critical ionization with an apparent pK_a value of 9.0 that must be protonated for activity (Figure A1.5B). The group that must be unprotonated for activity might reflect C167, consistent with a previously determined pK_a of 6.6 for cysteine alkylation by iodoacetamide with the nicotinamidase from *M. tuberculosis* (33). The lowered pK_a of C167 is likely due to the positively charged K122 amino group within 3.1 Å of the C167 sulfur. The group that must be protonated ($pK_a \sim 9$) may reflect the requirement for protonated K122. The precise positioning of the K122 amine within the Pnc1 active site is critical as both the Pnc1 K122A and K122R mutants displayed significantly diminished k_{cat} values 16- and 767-fold below those of Pnc1 WT, respectively. For the K122R mutant, the ^{15}N KIE is exactly the same as that determined for Pnc1 WT while the ^{13}C effect is only $\sim 0.4\%$ higher. This may reflect a classic case of non-productive substrate binding, which affects the k_{cat} , but not the k_{cat}/K_m value. Arginine likely sterically interferes with nicotinamide binding causing a greater degree of non-productive binding. However, once the substrate achieves the proper near attack conformation the reaction proceeds normally resulting in a minimal increase in the KIE. Another possibility is that the positively charged K122 might lower the pK_a values of D8 and C167 (31, 33). If so, the KIEs of K122R suggest that the positive charge of arginine is sufficient to maintain the lowered pK_a values of D8 and C167 in the transition state. If not, the KIE would be expected to increase since an increased pK_a value of either residue would slow proton abstraction and thus the C-N bond cleavage.

Within the Pnc1 active site, D8 is likely the base that initiates catalysis by abstracting the proton from C167 leading to the attack at the nicotinamide amide carbon (30, 31, 33, 52). The Pnc1 D8A, D8N, and D8E mutants revealed the importance of both the charge and positioning of D8 within the Pnc1 active site. Removal of the charge (D8A and D8N) resulted in 767- and 1150-fold reductions in k_{cat} value compared to Pnc1 WT. Repositioning the negative charge (D8E) increased both the ^{15}N and the ^{13}C KIE to very close to the maximum ($\sim 2.7\%$ for ^{15}N ; Table A1.4) observed with the substrate analogs and resulted in a 100-fold reduction in k_{cat} value. Therefore, the extra methylene group of glutamate likely pushes the carboxyl group out of position to substantially slow proton abstraction, resulting in slower C-N bond cleavage and the fullest expression of the KIEs. Finally, nucleophilic attack by C167 is further favored by the presence of a putative oxyanion hole (Figure A1.2B) originally proposed by Fyfe *et al.* (31) consisting of the amide backbone nitrogens of the conserved A163 and invariant C167 (Figure A1.3). These nitrogens form hydrogen bonds to the oxygen of the trapped thiohemiacetal adduct in the structure of nicotinaldehyde inhibited Pnc1.

Structure-activity relationships and rate-limiting steps of nicotinamide analogs. The pH dependence of nicotinamide hydrolysis displayed no critical ionizations over the range 4.5-10.5 (Figure A1.4A), indicating a step other than C167 nucleophilic attack and C-N bond cleavage is rate-limiting for nicotinamide, such as tetrahedral intermediate breakdown and product release. In addition, the KIEs determined for nicotinamide were significantly less than the maximum observed for the substrate analogs indicating that C-N bond cleavage is only partially rate-limiting with respect to the $k_{\text{cat}}/K_{\text{m}}$ value. The two critical features of proper nicotinamide

binding and orientation are the ring nitrogen, which is liganded to the zinc atom and the nicotinamide carbonyl oxygen, which is held in place by backbone amide interactions. Replacing the oxygen with sulfur (*i.e.* thionicotinamide) leads to a $>10^3$ drop in $k_{\text{cat}}/K_{\text{m}}$ value and a large 1.4% increase in the ^{15}N KIE value compared to nicotinamide. Therefore, the overall rate limiting step likely shifts from product release for nicotinamide, to C–N bond cleavage for thionicotinamide. Oxygen is more electronegative than sulfur (3.44 vs 2.58, Pauling Scale) (64), which can affect the reaction in two possible ways. First, the hydrogen bonding strength is less for a thionyl group than a carbonyl leading to a more loosely bound substrate (*i.e.* the interaction with the backbone amides of A163 and C167). Secondly, the drop in electronegativity directly impacts the electrophilicity of the thionyl amide carbon thus diminishing its reactivity when attacked by C167. In addition, the larger size of sulfur and the longer bond length of C=S versus C=O bonds may also sterically clash with the enzyme and result in non-optimal substrate positioning. Any of these reasons would lead to a decreased forward commitment factor of the reaction and thus an increase in the observed ^{15}N and ^{13}C KIE.

Addition of a methyl group on the ring nitrogen (*i.e.* 1-methylnicotinamide) and replacement of N1 with carbon (*i.e.* benzamide) both block the interaction with the bound zinc atom resulting in a dramatic drop in k_{cat} and $k_{\text{cat}}/K_{\text{m}}$ values (see above). Methylation at the C5 position (*i.e.* 5-methylnicotinamide) is tolerated, yielding a substrate that displays modest effects on $k_{\text{cat}}/K_{\text{m}}$ values but not of k_{cat} values compared to nicotinamide. The higher KIE values for 5-methylnicotinamide (and all the analogs tested) compared to nicotinamide indicated that 5-methylnicotinamide is not as “sticky” as nicotinamide (*i.e.* the ratio of substrate dissociation to

forward reaction is greater for the analogs). Thus, the C-N bond cleavage is almost entirely rate-limiting in regard to the $k_{\text{cat}}/K_{\text{m}}$ value. Therefore, the extra methyl group of 5-methylnicotinamide is most likely sterically interacting with a nearby residue (possibly Y166) hindering achievement of the proper orientation for optimal catalysis. However, the steric hindrance of the methyl group of 5-methylnicotinamide also results in a more loosely bound substrate (higher K_{m}) and a higher k_{cat} value likely due to faster product release. Therefore, 5-methylnicotinic acid release, not C-N bond breaking, is the overall rate-limiting step reflected in the k_{cat} value. Consistent with the high specificity of nicotinamidases toward nicotinamide, NAD^+ salvage pathway metabolites NMN^+ and NAD^+ are not substrates.

Pyrazinamide displayed a 4.5-fold lower $k_{\text{cat}}/K_{\text{m}}$ value compared to nicotinamide, but a 3.7-fold higher k_{cat} value (Table A1.5). The size and geometry of the pyrazine ring are nearly identical to the pyridine ring of nicotinamide, so the additional nitrogen in the pyrazine ring slows the binding and chemical steps reflected in the $k_{\text{cat}}/K_{\text{m}}$ value (substrate binding, nucleophilic attack of C167, and release of ammonia) through inductive effects or unfavorable interaction of the unshared pair of electrons with nearby aromatic residues involved in substrate binding and positioning leading to a more loosely bound substrate. However, the higher k_{cat} value indicates that pyrazinamide reacts faster in the chemical steps after release of ammonia such as thioester hydrolysis and pyrazinoic acid release. Indeed, the 56-fold higher K_{i} value of pyrazinoic acid compared to nicotinic acid (Table A1.2) is consistent with a higher off rate constant for pyrazinoic acid and an overall rate enhancement due to an increased rate of rate-limiting product release. The 4.5-fold lower $k_{\text{cat}}/K_{\text{m}}$ value for pyrazinamide is due to slower and nearly fully rate-

limiting amide C–N bond as revealed by the higher ^{15}N and ^{13}C KIE values of 2.3% and 2.0% determined for pyrazinamide, compared to 1.2% and 1.3% determined for nicotinamide.

Nicotinamidase chemical mechanism. Previous nicotinamidase studies indicate that there are three residues that directly chelate the Zn^{2+} at the active site (D51, H53, and H94 in Pnc1) and three that are involved as a catalytic triad (C167, D8, and K122 in Pnc1) (31, 32, 55). Here, our data suggests specific roles for each of these residues and other active-site residues during catalysis (Scheme A1.4). In the first step of the mechanism, the substrate binds at the active site replacing one of two equatorial water molecules ligated to the Zn^{2+} and is stabilized by edge face interactions with the aromatic residues F13, W91, Y131 and Y166. The nitrogen of the pyridine ring is liganded to the Zn^{2+} at the active site and the proton is removed from C167 by D8, forming a thiolate that then attacks the amide carbonyl carbon creating a tetrahedral intermediate. The recently abstracted proton is transferred to the amino portion of the amide, which is then ejected as NH_3 as the intermediate collapses and the C–N bond is broken. Although this final step before C–N scission is shown as concerted in Scheme A1.4, proton transfer and C–N cleavage may be in fact stepwise.

At the next stage of the mechanism there is some disagreement on how the reaction moves forward. It has been hypothesized that the Zn^{2+} ligated water is in proper position to attack the thiol ester (52), but our crystal structure indicates the water is too far away ($\sim 5\text{--}7\text{ \AA}$) from the carbonyl carbon to be effective without a significant conformational rearrangement or dissociation from the Zn^{2+} . Another option is that the Zn^{2+} ligated water is deprotonated and can

then deprotonate a bulk solvent water molecule that attacks the thiol ester (30). The final possibility is that a bulk solvent water molecule is deprotonated by D8 and then acts as the nucleophile reforming a tetrahedral intermediate (31, 33). We favor the third option as D8 is in proper position to accomplish this deprotonation. Finally, the intermediate collapses to form nicotinic acid and the cysteine thiolate abstracts the proton from D8 and the catalytic cycle is complete.

It is possible that if the thiol pK_a in Pnc1 is lowered to the same level as in PncA (pK_a 6.6) (33), presumably by the interaction with the positive charge of K122, then under the conditions of our experiments (pH 7.5) ~90% of the cysteine would exist in its thiolate form and thus would not require deprotonation by D8. Accordingly, D8 would only be involved in the hydrolytic portion of the reaction, removing a proton from a water molecule and reforming the tetrahedral intermediate. The glutamate in the D8E mutant is out of position relative to the substrate for forming the tetrahedral intermediate, thus retarding the reaction so that the KIE is fully expressed in the D8E mutant because the forward commitment has become very small.

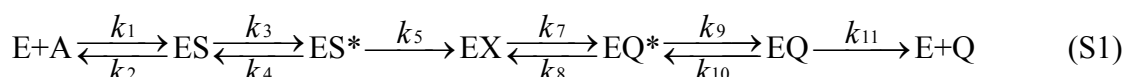
Conclusion

This detailed characterization of a eukaryotic nicotinamidase provides critical information on the structure and mechanism of nicotinamidases. In particular, this work provides fundamental understanding for the development of mechanism-based inhibitors and prodrugs that target nicotinamidases to treat fungal, bacterial, or parasite infections. In addition, results like those reported here, along with theoretical calculations of transition-state bond lengths, have been used

by others to design excellent transition-state inhibitors for other enzymes (65). Similarly, we anticipate that these findings can lead to the development of nicotinamidase inhibitors to treat infections in which nicotinamidase activity is critical for viability or the infectious phenotype.

Supporting Discussion

Derivation explaining observed kinetic isotope effects for Pnc1 D51N. For Pnc1 D51N, both the ^{15}N and ^{13}C KIE are $\sim 0.5\%$ compared to $\sim 1.25\%$ for those of Pnc1 WT (Table A1.4). We believe this is due to an increase in one of the back reaction off rates. A truncated derivation demonstrating how this is possible is shown below. The entire reaction is described by equation S1.



The KIE are on V/K which includes all steps up through and including the first irreversible step where k_1 is substrate binding, k_2 is substrate dissociation, k_3 is the formation of the tetrahedral intermediate, k_4 is the breakdown of the tetrahedral intermediate to reform nicotinamide, and k_5 is the isotope sensitive irreversible step (C–N bond cleavage and the loss of NH_3). V/K is expressed by:

$$\frac{V}{KE_t} = \frac{\frac{k_1 k_3 k_5}{k_2 k_4}}{1 + \frac{k_5}{k_4} + \frac{k_5 k_3}{k_4 k_2}} \quad (\text{S2})$$

However when the kinetic data are examined we see the k_{cat} and k_{cat}/K_m values have decreased compared to the wild-type enzyme. This is due to a decrease in binding of nicotinamide in the D51N mutant which leads to an increase in k_2 and a decrease in k_3 . As such k_3/k_2 would tend to be very small and can be ignored. Then using the ^{15}N KIE notation the equation becomes:

$$^{15}\left(\frac{V}{K}\right) = \frac{^{15}k_5 + \frac{k_5}{k_4}}{1 + \frac{k_5}{k_4}} \quad (\text{S3})$$

where $^{15}\left(\frac{V}{K}\right)$ is the measured KIE and $^{15}k_5$ is the intrinsic KIE. $^{15}\left(\frac{V}{K}\right)$ is defined as

$$\frac{^{15}k_5 + C_f + ^{15}K_{eq}}{1 + C_f + C_r} \text{ where } C_f \text{ and } C_r \text{ are forward and reverse commitments to catalysis and } ^{15}K_{eq} \text{ is}$$

the ^{15}N equilibrium isotope effect. In this reaction there is no reverse commitment since the C-N bond cleavage is irreversible and thus C_r and $^{15}K_{eq}$ can be disregarded and:

$$^{15}\left(\frac{V}{K}\right) = \frac{^{15}k_5 + C_f}{1 + C_f} \quad (\text{S4})$$

For this example, an intrinsic isotope effect of 3% will be assumed and then for wild type and D51N we have:

$$1.0122 \pm 0.0002 = \left(\frac{1.0300 + C_{fwt}}{1 + C_{fwt}} \right), \text{ and } C_{fwt} = 1.50 \pm 0.04 \quad (\text{S5})$$

$$1.0045 \pm 0.0006 = \left(\frac{1.0300 + C_{fd51N}}{1 + C_{fd51N}} \right), \text{ and } C_{fd51N} = 5.50 \pm 0.82 \quad (\text{S6})$$

Equation 12 shows that $C_f = \frac{k_5}{k_4}$ and so the C_f values are $\left(\frac{k_5}{k_4}\right)_{WT} \sim 1.5$ and $\left(\frac{k_5}{k_4}\right)_{Mut} \sim 5.5$

These values show that the reaction partitions forward faster in the D51N mutant, resulting in a smaller KIE. The mechanism of this partitioning is unknown but indicates that k_5 (loss of NH_3) has increased or that breakdown of the tetrahedral intermediate (k_4) has decreased.

Acknowledgements

We thank Advanced Photon Source staff (LS-CAT beamline) and Katrina Forest for assistance with data collection, and Ken Satyshur and Nick George for help with structure determination. We also thank Dr. Tonya Zeczycki for her helpful comments and critique of this manuscript. This work was supported by National Institutes of Health grants GM065386 (to J.M.D.) and GM18938 (to W.W.C.), by National Institutes of Health Biotechnology Training Grant NIH 5 T32 GM08349 (to B.C.S.), and by an NIH training grant in Molecular Biosciences GM07215 (to K.A.H.).

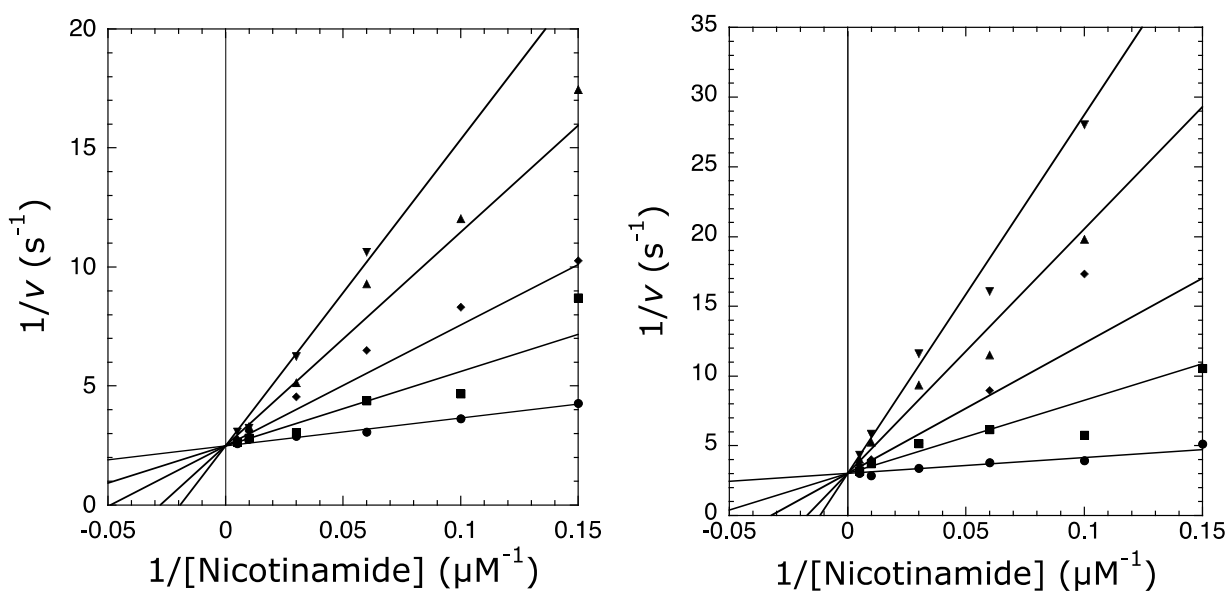
Figure A1.1.

Figure A1.1. Double-reciprocal inhibition plots. Initial rates were determined as described under Experimental Procedures. (left) Nicotinic acid exhibits competitive inhibition toward nicotinamide during the Pnc1-catalyzed reaction. The following nicotinic acid concentrations were used: 0 (l), 200 (n), 400 (u), 800 (s), and 1200 μM (t). (right) Nicotinaldehyde exhibits competitive inhibition toward nicotinamide during the Pnc1-catalyzed reaction. The following nicotinaldehyde concentrations were used: 0 (l), 1 (n), 2 (u), 4 (s), and 6 μM (t). Data were fit to competitive inhibition models using KinetAsyst as described in Experimental Procedures. All reactions were performed at pH 7.5.

Figure A1.2

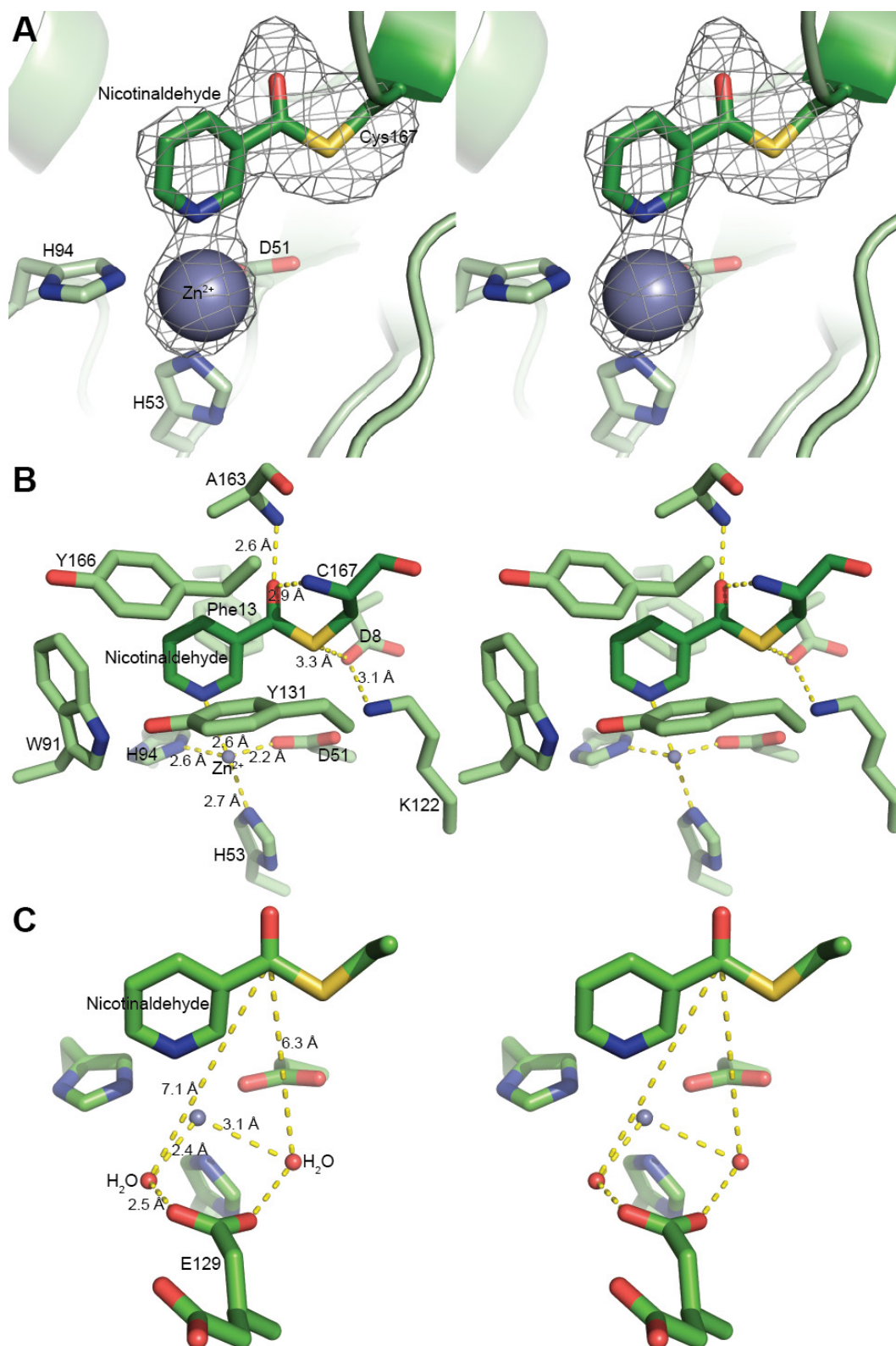


Figure A1.2. Structure of nicotinaldehyde covalently bound in the active site of yeast Pnc1.

(A) F_o-F_c electron density map (3.5σ) showing the inhibitor and Zn^{2+} in cross-eyed stereo view. The inhibitor (nicotinaldehyde) covalently attached to Cys167 was modeled into the density. Additional residues that coordinate the Zn^{2+} (Asp51, His53, and His94) are also shown. (B) A cross-eyed stereo view of the active site of Pnc1 with nicotinaldehyde covalently bound. (C) A cross-eyed stereo view of ordered water within the nicotinaldehyde-bound Pnc1 active site.

Figure A1.3. Multiple-sequence alignment of the nicotinamidases for which structures have been solved from *Pyrococcus horikoshii* (PDB entries 1ILW and 1IM5), *Mycobacterium tuberculosis* (PDB entry 3PL1), *Streptococcus pneumoniae* (PDB entries 3O90, 3O91, 3O92, 3O93, and 3O94), *Acinetobacter baumannii* (PDB entries 2WT9 and 2WTA), and *Saccharomyces cerevisiae* (PDB entry 2H0R and this work). Eukaryotic nicotinamidases from *Caenorhabditis elegans* and *Drosophila melanogaster* are also shown for comparison. The alignment was generated using ClustalW (66).

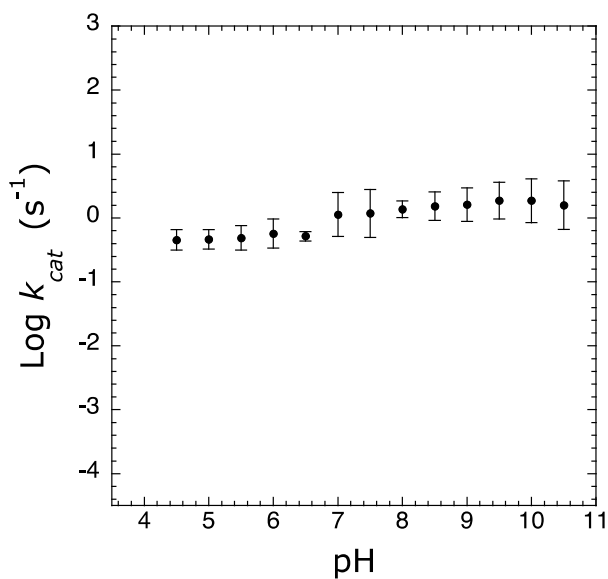
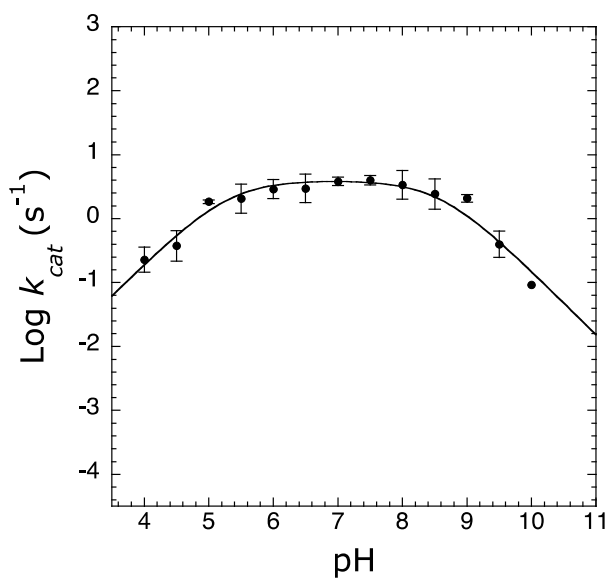
Figure A1.4**A****B**

Figure A1.4. pH rate profiles. Reaction rates were determined as described under Experimental Procedures. **(A)** Effect of pH on k_{cat} of Pnc1 catalyzed hydrolysis of nicotinamide. **(B)** Effect of pH on k_{cat} of Pnc1 catalyzed hydrolysis of pyrazinamide.

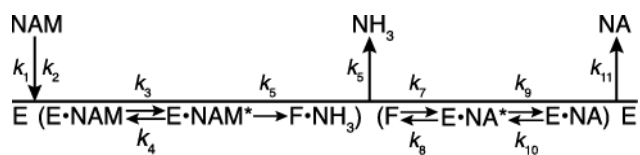
Figure A1.5

Figure A1.5. Proposed kinetic mechanism of the nicotinamidase reaction. E represents the unmodified Pnc1 enzyme, F represents the covalent thioester intermediate between nicotinic acid and C167, NAM represents nicotinamide, and NA represents nicotinic acid.

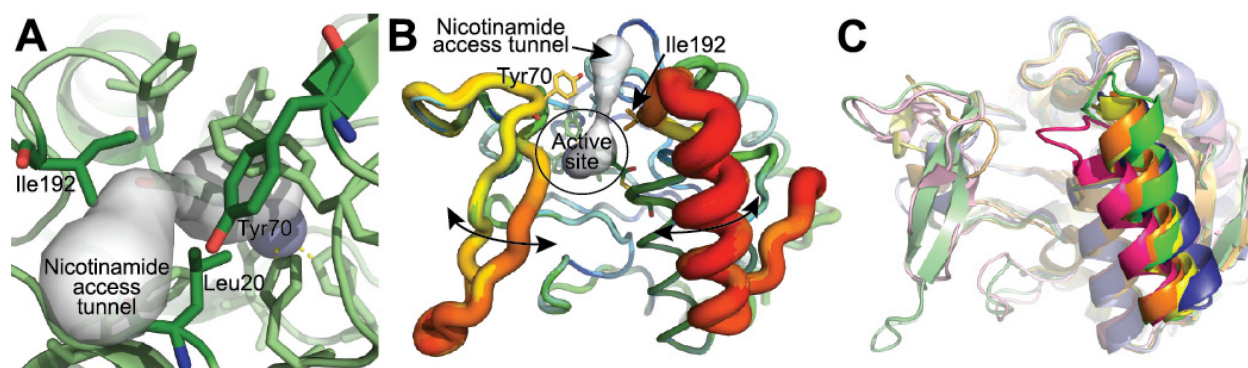
Figure A1.6

Figure A1.6. Putative route of nicotinamide access to the Pnc1 active site. (A) A narrow tunnel exists in the Pnc1 structure from the active site to the protein surface. This tunnel was discovered using Caver (67, 68). (B) The C-terminal loop and helix of the nicotinaldehyde-Pnc1 structure is the most dynamic portion of the protein. All seven Pnc1 monomers in the asymmetric unit were aligned and represented using the b factor putty preset of Pymol (69). The range of b factors in the structure are represented as thin (low b factor) to thick (high b factor) main chains and a color gradient from blue (low b factor) to red (high b factor). (C) Alignment of all published nicotinamidase structures reveals that the C-terminal loop and helix containing I192 is the most dynamic portion of the structures. The nicotinamidases shown are from *Pyrococcus horikoshii* (PDB 1IM5; yellow), *Mycobacterium tuberculosis* (PDB 3PL1; orange), *Streptococcus pneumoniae* (PDB 3O91; blue), *Acinetobacter baumannii* (PDB 2WT9; pink), and *Saccharomyces cerevisiae* (this work; green).

Table A1.1. Primers used in Pnc1 mutagenesis.

Pnc1 mutant	Primers
D8A	5'-CGAGATGAAGACTTTAATTGTTGTTGCTATGCAAAATGATTTTATTTTACC-3' 5'-GGTGAAATAAAATCATTTTGCATAGCAACAACAATTAAAGTCTTCATCTCG-3'
D8N	5'-CGAGATGAAGACTTTAATTGTTGTTAATATGCAAAATGATTTTATTTTACC-3' 5'-GGTGAAATAAAATCATTTTGCATATTAACAACAATTAAAGTCTTCATCTCG-3'
D8E	5'-CGAGATGAAGACTTTAATTGTTGTTGAGATGCAAAATGATTTTATTTTACC-3' 5'-GGTGAAATAAAATCATTTTGCATCTCAACAACAATTAAAGTCTTCATCTCG-3'
D51A	5'-GTGGTCACCAGAGCTTGGCACCCTTCC-3' 5'-GGAAGGGTGCCAAGCTCTGGTGACCAC-3'
D51N	5'-GTGGTCACCAGAAATTGGCACCCTTCC-3' 5'-GGAAGGGTGCCAATTTCTGGTGACCAC-3'
H53A	5'-GGTCACCAGAGATTGGGCCCCCTTCCAGAC-3' 5'-GTCTGGAAGGGGCCCAATCTCTGGTGACC-3'
H94A	5'-TGTGGCCCGTAGCCTGTGTGAAAAACACC-3' 5'-GGTGTTTTTTCACACAGGCTACGGGCCACA-3'
K122A	5'-GATTGTCGACGCGGGTTTCTTGACTGACC-3' 5'-GGTCAGTCAAGAAACCCGCGTCGACAATC-3'
K122R	5'-GATTGTCGACAGGGTTTCTTGACTGACC-3' 5'-GGTCAGTCAAGAAACCCCTGTGCGACAATC-3'
C167A	5'-AGCTTTGGAGTATGCTGTCAAAGCCACCG-3' 5'-CGGTGGCTTTGACAGCATACTCCAAAGCT-3'

Table A1.2. K_i values of Pnc1 competitive inhibition by nicotinamide analogs. K_i values were determined as described under Experimental Procedures. All reactions were performed at pH 7.5.

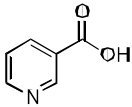
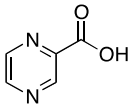
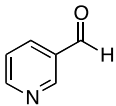
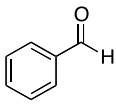
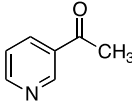
Inhibitor	Structure	K_i value (μM)
nicotinic acid		120 ± 19
pyrazinoic acid		6700 ± 1700
nicotinaldehyde		0.94 ± 0.35
benzaldehyde		20600 ± 5700
3-acetylpyridine		316 ± 101

Table A1.3. X-ray data collection and structure determination statistics

Data Collection	
Wavelength, Å	0.97872
Resolution Range (high resolution bin), Å	30-2.70 (2.75-2.70)
Space Group	R ₃
Unit Cell (a, b, c (Å))	298.72, 298.72, 112.65
(α , β , γ (°))	90, 90, 120
Completeness, %	100.0 (100.0)
Total/Unique Reflections	578,608/101,147
Redundancy	5.7 (5.5)
$\langle I/\sigma I \rangle$	19.8 (5.3)
R _{sym} ^a , %	11.3 (34.8)
Refinement	
Resolution, Å	30-2.70
R _{work} /R _{free} ^b , %	18.9/21.4
Rms deviations	
Bonds, Å	0.009
Angles, °	1.23
Ramachandran statistics, %	
Most favored	89.8
Allowed	9.9
Generously allowed	0.3
Disallowed	0.0
# atoms	
Protein	12,439
Water	296
Ligand ^c	14
$\langle B \text{ factor} \rangle$, Å ²	
Protein	27.5
Water	23.9
Ligand ^c	82.7

^a $R_{\text{sym}} = \frac{\sum \sum_j |I_j - \langle I \rangle|}{\sum I_j}$, where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

^b $R_{\text{work}}/R_{\text{free}} = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$, where the working and free R factors are calculated by using the working and free reflection sets, respectively. The free R reflections (5% of the total) were held aside throughout refinement.

^c Ligand refers to 7 zinc and 7 magnesium ions (one zinc and one magnesium for each of 7 monomers in the asymmetric unit).

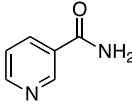
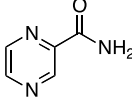
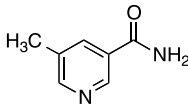
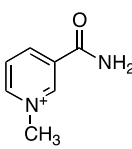
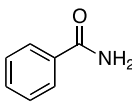
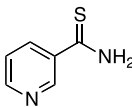
Table A1.4. Steady-state kinetic parameters of Pnc1 mutants. Assays were performed as described under Experimental Procedures. All reactions were performed at pH 7.5.

Pnc1 mutant	k_{cat} (s^{-1}) ^a	K_{m} (μM) ^a	$k_{\text{cat}}/K_{\text{m}}$ ($\text{M}^{-1}\text{s}^{-1}$)	¹⁵ N KIE	¹³ C KIE
wild type	0.69 ± 0.04	9.6 ± 2.1	$(7.2 \pm 1.2) \times 10^4$	1.0122 ± 0.0002	1.0125 ± 0.0004
D8A	0.0009 ± 0.0001	NA ^b	NA ^b	ND ^c	ND ^c
D8N	0.0006 ± 0.0001	NA ^b	NA ^b	ND ^c	ND ^c
D8E	0.0070 ± 0.0001	NA ^b	NA ^b	1.0218 ± 0.0015	1.0258 ± 0.0012
D51A	0.020 ± 0.005	< 2.4	$> 8.3 \times 10^3$	1.0166 ± 0.0018	1.0075 ± 0.0015
D51N	0.064 ± 0.012	< 1.8	$> 3.5 \times 10^4$	1.0047 ± 0.0006	1.0058 ± 0.0006
H53A	0.014 ± 0.002	6.5 ± 3.6	$(2.2 \pm 0.9) \times 10^3$	ND ^c	ND ^c
H94A	0.020 ± 0.002	< 2.7	$> 7.4 \times 10^3$	ND ^c	ND ^c
K122A	0.044 ± 0.023	< 6.5	$> 6.7 \times 10^3$	ND ^c	ND ^c
K122R	0.0009 ± 0.0002	NA ^b	NA ^b	1.0127 ± 0.0003	1.0172 ± 0.0001
C167A	< 0.0005	NA ^b	NA ^b	ND ^c	ND ^c

^a k_{cat} and K_{m} values were determined from fitting the data to the Michaelis-Menten equation using Kaleidagraph (Synergy Software, Reading, PA). Errors represent the error of the fit to the data.

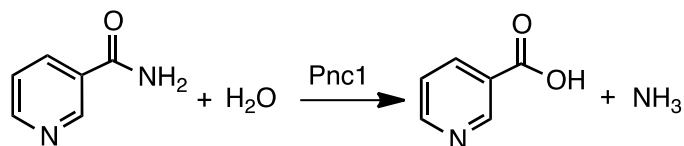
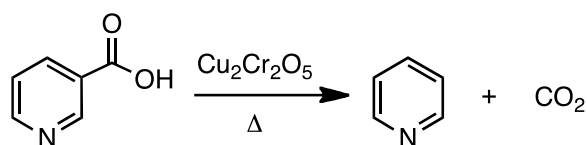
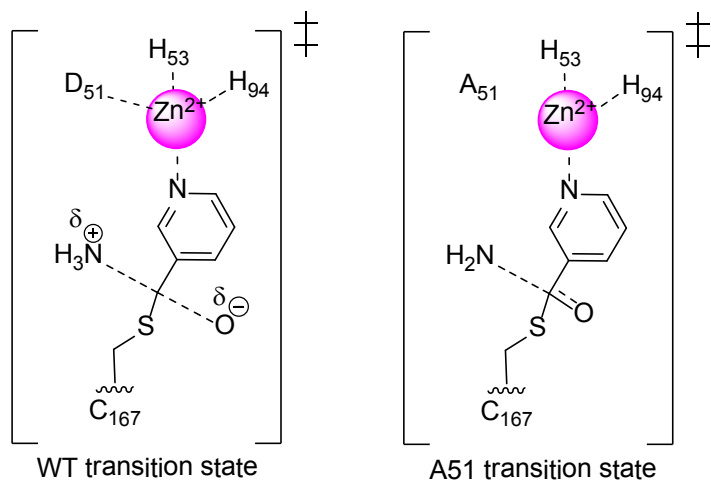
^bNot available. Measured rates were too slow for accurate K_{m} values to be determined. ^cNot determined.

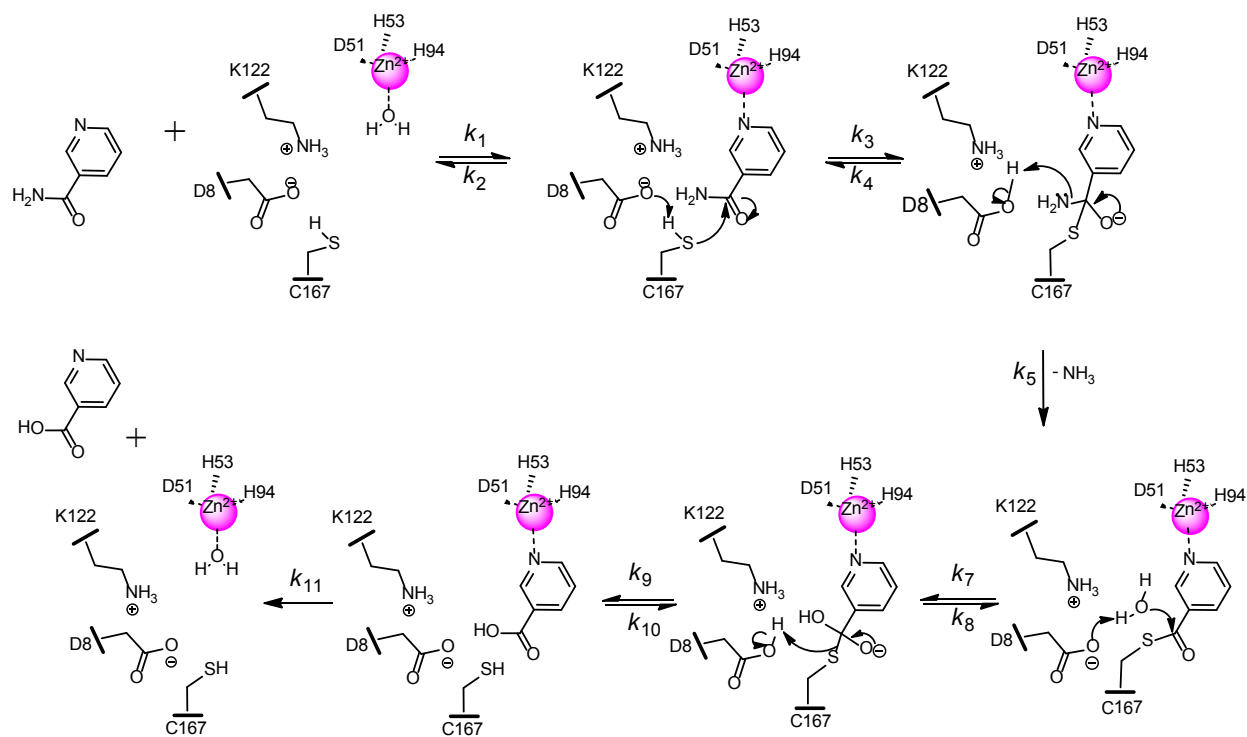
Table A1.5. Steady-state kinetic parameters and ^{15}N and ^{13}C KIE for the reaction catalyzed by Pnc1 with alternate substrates. Assays were performed as described under Experimental Procedures at pH 7.5.

Substrate	Structure	k_{cat} (s^{-1}) ^a	K_{m} (μM) ^a	$k_{\text{cat}}/K_{\text{m}}$ ($\text{M}^{-1}\text{s}^{-1}$)	^{15}N KIE	^{13}C KIE
nicotinamide		0.69 ± 0.04	9.6 ± 2.1	$(7.2 \pm 1.2) \times 10^4$	1.0122 ± 0.0002	1.0125 ± 0.0004
pyrazinamide		2.56 ± 0.10	157 ± 16	$(1.6 \pm 0.1) \times 10^4$	1.0231 ± 0.0007	1.0196 ± 0.0023
5-methylnicotinamide		1.75 ± 0.05	61 ± 6	$(2.9 \pm 0.2) \times 10^4$	1.0151 ± 0.0009	1.0269 ± 0.0037
1-methylnicotinamide		NA ^b	NA ^b	$(8.6 \pm 1.7) \times 10^0$	ND ^d	ND ^d
benzamide		0.0088 ± 0.0004	25 ± 7	$(3.5 \pm 0.9) \times 10^2$	ND ^d	ND ^d
thionicotinamide		NA ^b	NA ^b	$(5.0 \pm 1.7) \times 10^1$	1.0262 ± 0.0019	ND ^d
nicotinamide mononucleotide (NMN ⁺)		$< 10^{-5}$ ^c				
nicotinamide adenine dinucleotide (NAD ⁺)		$< 10^{-5}$ ^c				

^a k_{cat} and K_{m} values were determined from fitting the data to the Michaelis-Menten equation using Kaleidagraph (Synergy Software, Reading, PA). Errors represent the error of the fit to the data.

^bSaturation was not obtained, therefore only $k_{\text{cat}}/K_{\text{m}}$ values are reported. ^cNo activity was observed above the detection limit of the assay of 10^{-5} s^{-1} . ^dNot determined.

Scheme A1.1. General reaction catalyzed by nicotinamidases.**Scheme A1.2. The decarboxylation of nicotinic acid by copper chromite.****Scheme A1.3. Possible transition states depicting a stiffer bonded carbon (lower KIE) and a less stiffly bond nitrogen (higher KIE) in the D51A mutant compared to WT.**

Scheme A1.4. Proposed Pnc1 chemical mechanism.

References

1. Denu JM (2007) Vitamins and aging: pathways to NAD⁺ synthesis. *Cell* 129(3):453-454.
2. Ghislain M, Talla E, & Francois JM (2002) Identification and functional analysis of the *Saccharomyces cerevisiae* nicotinamidase gene, PNC1. *Yeast* 19(3):215-224.
3. Zerez CR, Roth EF, Jr., Schulman S, & Tanaka KR (1990) Increased nicotinamide adenine dinucleotide content and synthesis in *Plasmodium falciparum*-infected human erythrocytes. *Blood* 75(8):1705-1710.
4. Wang G & Pichersky E (2007) Nicotinamidase participates in the salvage pathway of NAD biosynthesis in *Arabidopsis*. *Plant J* 49(6):1020-1029.
5. Balan V, Miller GS, Kaplun L, Balan K, Chong ZZ, Li F, Kaplun A, VanBerkum MF, Arking R, Freeman DC, Maiese K, & Tzivion G (2008) Life span extension and neuronal cell protection by *Drosophila* nicotinamidase. *J Biol Chem* 283(41):27810-27819.
6. Vrablik TL, Huang L, Lange SE, & Hanna-Rose W (2009) Nicotinamidase modulation of NAD⁺ biosynthesis and nicotinamide levels separately affect reproductive development and cell survival in *C. elegans*. *Development* 136(21):3637-3646.
7. van der Horst A, Schavemaker JM, Pellis-van Berkel W, & Burgering BM (2007) The *Caenorhabditis elegans* nicotinamidase PNC-1 enhances survival. *Mech Ageing Dev* 128(4):346-349.
8. Belenky P, Bogan KL, & Brenner C (2007) NAD⁺ metabolism in health and disease. *Trends Biochem Sci* 32(1):12-19.
9. Purser JE, Lawrenz MB, Caimano MJ, Howell JK, Radolf JD, & Norris SJ (2003) A plasmid-encoded nicotinamidase (PncA) is essential for infectivity of *Borrelia burgdorferi* in a mammalian host. *Mol Microbiol* 48(3):753-764.
10. Kim S, Kurokawa D, Watanabe K, Makino S, Shirahata T, & Watarai M (2004) *Brucella abortus* nicotinamidase (PncA) contributes to its intracellular replication and infectivity in mice. *FEMS Microbiol Lett* 234(2):289-295.
11. Zhang Y & Mitchison D (2003) The curious characteristics of pyrazinamide: a review. *Int J Tuberc Lung Dis* 7(1):6-21.
12. Shi W, Zhang X, Jiang X, Yuan H, Lee JS, Barry CE, 3rd, Wang H, Zhang W, & Zhang Y (2011) Pyrazinamide inhibits trans-translation in *Mycobacterium tuberculosis*. *Science* 333(6049):1630-1632.
13. Konno K, Feldmann FM, & McDermott W (1967) Pyrazinamide susceptibility and amidase activity of tubercle bacilli. *Am Rev Respir Dis* 95(3):461-469.

14. Scorpio A & Zhang Y (1996) Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med* 2(6):662-667.
15. Hirano K, Takahashi M, Kazumi Y, Fukasawa Y, & Abe C (1997) Mutation in *pncA* is a major mechanism of pyrazinamide resistance in *Mycobacterium tuberculosis*. *Tuber Lung Dis* 78(2):117-122.
16. Sreevatsan S, Pan X, Zhang Y, Kreiswirth BN, & Musser JM (1997) Mutations associated with pyrazinamide resistance in *pncA* of *Mycobacterium tuberculosis* complex organisms. *Antimicrob Agents Chemother* 41(3):636-640.
17. Sandmeier JJ, Celic I, Boeke JD, & Smith JS (2002) Telomeric and rDNA silencing in *Saccharomyces cerevisiae* are dependent on a nuclear NAD(+) salvage pathway. *Genetics* 160(3):877-889.
18. Anderson RM, Bitterman KJ, Wood JG, Medvedik O, & Sinclair DA (2003) Nicotinamide and PNC1 govern lifespan extension by calorie restriction in *Saccharomyces cerevisiae*. *Nature* 423(6936):181-185.
19. Bitterman KJ, Anderson RM, Cohen HY, Latorre-Esteves M, & Sinclair DA (2002) Inhibition of silencing and accelerated aging by nicotinamide, a putative negative regulator of yeast sir2 and human SIRT1. *J Biol Chem* 277(47):45099-45107.
20. Gallo CM, Smith DL, Jr., & Smith JS (2004) Nicotinamide clearance by Pnc1 directly regulates Sir2-mediated silencing and longevity. *Mol Cell Biol* 24(3):1301-1312.
21. Kaeberlein M, McVey M, & Guarente L (1999) The SIR2/3/4 complex and SIR2 alone promote longevity in *Saccharomyces cerevisiae* by two different mechanisms. *Genes Dev* 13(19):2570-2580.
22. Lin SJ, Defossez PA, & Guarente L (2000) Requirement of NAD and SIR2 for life-span extension by calorie restriction in *Saccharomyces cerevisiae*. *Science* 289(5487):2126-2128.
23. Tissenbaum HA & Guarente L (2001) Increased dosage of a sir-2 gene extends lifespan in *Caenorhabditis elegans*. *Nature* 410(6825):227-230.
24. Rogina B & Helfand SL (2004) Sir2 mediates longevity in the fly through a pathway related to calorie restriction. *Proc Natl Acad Sci U S A* 101(45):15998-16003.
25. Wood JG, Rogina B, Lavu S, Howitz K, Helfand SL, Tatar M, & Sinclair D (2004) Sirtuin activators mimic caloric restriction and delay ageing in metazoans. *Nature* 430(7000):686-689.

26. Jackson MD, Schmidt MT, Oppenheimer NJ, & Denu JM (2003) Mechanism of nicotinamide inhibition and transglycosidation by Sir2 histone/protein deacetylases. *J Biol Chem* 278(51):50985-50998.
27. Landry J, Slama JT, & Sternglanz R (2000) Role of NAD(+) in the deacetylase activity of the SIR2-like proteins. *Biochem Biophys Res Commun* 278(3):685-690.
28. Landry J, Sutton A, Tafrov ST, Heller RC, Stebbins J, Pillus L, & Sternglanz R (2000) The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc Natl Acad Sci U S A* 97(11):5807-5811.
29. Burnett C, Valentini S, Cabreiro F, Goss M, Somogyvari M, Piper MD, Hoddinott M, Sutphin GL, Leko V, McElwee JJ, Vazquez-Manrique RP, Orfila AM, Ackerman D, Au C, Vinti G, Riesen M, Howard K, Neri C, Bedalov A, Kaeberlein M, Soti C, Partridge L, & Gems D (2011) Absence of effects of Sir2 overexpression on lifespan in *C. elegans* and *Drosophila*. *Nature* 477(7365):482-485.
30. French JB, Cen Y, Vrablik TL, Xu P, Allen E, Hanna-Rose W, & Sauve AA (2010) Characterization of nicotinamidases: steady state kinetic parameters, classwide inhibition by nicotinaldehydes, and catalytic mechanism. *Biochemistry* 49(49):10421-10439.
31. Fyfe PK, Rao VA, Zemla A, Cameron S, & Hunter WN (2009) Specificity and mechanism of *Acinetobacter baumannii* nicotinamidase: implications for activation of the front-line tuberculosis drug pyrazinamide. *Angew Chem Int Ed Engl* 48(48):9176-9179.
32. Hu G, Taylor AB, McAlister-Henn L, & Hart PJ (2007) Crystal structure of the yeast nicotinamidase Pnc1p. *Arch Biochem Biophys* 461(1):66-75.
33. Seiner DR, Hegde SS, & Blanchard JS (2010) Kinetics and inhibition of nicotinamidase from *Mycobacterium tuberculosis*. *Biochemistry* 49(44):9613-9619.
34. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248-254.
35. Su S & Chaykin S (1971) Nicotinamide deamidase. *Methods Enzymol.* 18(Pt. 2):185-192.
36. Cleland WW (1977) Determining the chemical mechanisms of enzyme-catalyzed reactions by kinetic studies. *Adv Enzymol Relat Areas Mol Biol* 45:273-387.
37. Otwinowski Z & Minor W (1997) *Processing of X-ray Diffraction Data Collected in Oscillation Mode* (Academic Press, New York).
38. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.

39. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
40. Winn MD, Isupov MN, & Murshudov GN (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* 57(Pt 1):122-133.
41. Hanks LV & Henderson LM (1957) The metabolism of carboxyl-labeled 3-hydroxyanthranilic acid in the rat. *J Biol Chem* 225(1):349-354.
42. Scott TA (1967) A method for the degradation of radioactive nicotinic acid. *Biochem J* 102(1):87-93.
43. Watson GK & Cain RB (1975) Microbial metabolism of the pyridine ring. Metabolic pathways of pyridine biodegradation by soil bacteria. *Biochem J* 146(1):157-172.
44. Ellis KJ & Morrison JF (1982) Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol* 87:405-426.
45. Oishi M, Ogasawara Y, Ishii K, & Tanabe S (1998) Assay of nicotinamide deamidase activity using high-performance liquid chromatography. *J Chromatogr B Biomed Sci Appl* 720(1-2):59-64.
46. Gadd REA & Johnson WJ (1974) Kinetic studies of nicotinamide deamidase from *Micrococcus lysodeikticus*. *Int. J. Biochem.* 5(4):397-407.
47. Calbreath DF & Joshi JG (1971) Inhibition of nicotinamidase by nicotinamide adenine dinucleotide. *J Biol Chem* 246(13):4334-4339.
48. Grossowicz N & Halpern YS (1956) Inhibition of nicotinamidase activity in cell-free extracts of *Mycobacterium phlei* by 3-acetylpyridine. *Biochim Biophys Acta* 20(3):576-577.
49. Johnson WJ & Gadd REA (1974) Inhibition of nicotinamide deamidase from *Micrococcus lysodeikticus* by analogs of nicotinamide. *Int. J. Biochem.* 5(7-8):633-641.
50. Joshi JG & Handler P (1962) Purification and properties of nicotinamidase from *Torula cremoris*. *J Biol Chem* 237:929-935.
51. Tanigawa Y, Shimoyama M, Dohi K, & Ueda I (1972) Purification and properties of nicotinamide deamidase from *Flavobacterium peregrinum*. *J Biol Chem* 247(24):8036-8042.
52. French JB, Cen Y, Sauve AA, & Ealick SE (2010) High-resolution crystal structures of *Streptococcus pneumoniae* nicotinamidase with trapped intermediates provide insights

- into the catalytic mechanism and inhibition by aldehydes. *Biochemistry* 49(40):8803-8812.
53. Petrella S, Gelus-Ziental N, Maudry A, Laurans C, Boudjelloul R, & Sougakoff W (2011) Crystal structure of the pyrazinamidase of *Mycobacterium tuberculosis*: insights into natural and acquired resistance to pyrazinamide. *PLoS One* 6(1):e15785.
 54. Zhang H, Deng JY, Bi LJ, Zhou YF, Zhang ZP, Zhang CG, Zhang Y, & Zhang XE (2008) Characterization of *Mycobacterium tuberculosis* nicotinamidase/pyrazinamidase. *FEBS J* 275(4):753-762.
 55. Du X, Wang W, Kim R, Yakota H, Nguyen H, & Kim SH (2001) Crystal structure and mechanism of catalysis of a pyrazinamidase from *Pyrococcus horikoshii*. *Biochemistry* 40(47):14166-14172.
 56. Cleland WW (1963) The kinetics of enzyme-catalyzed reactions with two or more substrates or products. II. Inhibition: nomenclature and theory. *Biochim Biophys Acta* 67:173-187.
 57. Kakemi K, Sezaki H, Nakano M, Ohsuga K, & Mitsunaga T (1969) Hydrolytic and associative behavior of aromatic amides in aqueous solution. *Chem. Pharm. Bull. (Tokyo)* 17(5):901-905.
 58. Hernick M & Fierke CA (2005) Zinc hydrolases: the mechanisms of zinc-dependent deacetylases. *Arch Biochem Biophys* 433(1):71-84.
 59. Lombardi PM, Cole KE, Dowling DP, & Christianson DW (2011) Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. *Curr Opin Struct Biol.*
 60. Lodi PJ, Chang LC, Knowles JR, & Komives EA (1994) Triosephosphate isomerase requires a positively charged active site: the role of lysine-12. *Biochemistry* 33(10):2809-2814.
 61. Thuku RN, Brady D, Benedik MJ, & Sewell BT (2009) Microbial nitrilases: versatile, spiral forming, industrial enzymes. *J Appl Microbiol* 106(3):703-727.
 62. Yan C & Sloan DL (1987) Purification and characterization of nicotinamide deamidase from yeast. *J Biol Chem* 262(19):9082-9087.
 63. Wolfenden R (1976) Transition state analog inhibitors and enzyme catalysis. *Annu Rev Biophys Bioeng* 5:271-306.
 64. Huheey JE (1983) *Inorganic Chemistry: Principles of Structure and Reactivity* (Harper & Row, New York); 3rd Ed.

65. Schramm VL (2007) Enzymatic transition state theory and transition state analogue design. *J Biol Chem* 282(39):28297-28300.
66. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, & Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-2948.
67. Damborsky J, Petrek M, Banas P, & Otyepka M (2007) Identification of tunnels in proteins, nucleic acids, inorganic materials and molecular ensembles. *Biotechnol J* 2(1):62-67.
68. Petrek M, Otyepka M, Banas P, Kosinova P, Koca J, & Damborsky J (2006) CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* 7:316.
69. DeLano WL (2002) The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, CA, USA).