THE TREX1 3' EXONUCLEASE AND AUTOIMMUNE DISEASE:
STRUCTURAL AND BIOCHEMICAL ANALYSIS OF DISEASE
MUTANTS INVOLVED IN AUTOIMMUNE DYSFUNCTION

By

Suzanna L. Bailey

A Dissertation Submitted to the Graduate Faculty of
WAKE FOREST UNIVERSITY GRADUATE SCHOOL OF ARTS AND
SCIENCES
in Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY
in the Department of Biochemistry and Molecular Biology
of the Wake Forest University School of Medicine
May 2010
Winston-Salem, North Carolina

Approved By:

Thomas Hollis, Ph.D., Advisor

Examing Committee:

Rebecca Alexander, Ph.D., Chairperson

David A. Horita, Ph.D.

Fred W. Perrino, Ph.D.

Leslie B. Poole, Ph.D.
Acknowledgements

First I would like to thank my advisor, Dr. Thomas Hollis, for his guidance and support throughout my time at Wake Forest. He has emphasized the importance of creativity and has encouraged me to think about how to approach a question from multiple directions. He has spent countless hours of his time reading and revising my research proposals and papers, and in the process, has helped me to improve my scientific writing. His excitement about science and new ideas has been an exceptional example that I will remember long after I leave his laboratory.

I would like to thank both current and former members of the Hollis laboratory for their friendship, support, and helpful discussion over the last five years. Special thanks goes to Dr. Udesh de Silva, who taught me about working with proteins from the ground up. He was encouraging and always very patient with the plethora of questions that I shot his way.

I would like to thank my advisory committee Drs Rebecca Alexander, David Horita, Fred Perrino, and Leslie Poole for the contribution of many constructive comments and ideas on my project. I would additionally like to thank Fred Perrino for providing me with purified TREX1 protein and the use of his laboratory to run DNA sequencing gels.

Special thanks goes to Lynnette Johnson who has become a close friend over the last few years. She has been a phenomenal resource for questions both in the laboratory and in life. I am going to miss having her around.
I would like to thank my undergraduate research advisor, Dr. Moses Lee, as well as Dr. Karen Buchmueller, who were instrumental in my decision to go to graduate school. Working with them, I discovered how much I enjoyed research. Without that experience, I never would have considered attending graduate school.

My parents, Ken and Ann Bailey, deserve special mention. Throughout my life they have provided me with many opportunities and challenged me to take advantage of those opportunities to the best of my ability. I cannot thank them enough for all of the encouragement and support they have provided me with over the years. I know that I would not be where I am today without them.

Finally, I would like to thank my husband, Dr. Michael Murray. There is no way for me to express on paper all of the things that he has taught me or the ways that he has encouraged me that have helped me through the last five years. This work is dedicated to him.
# Table of Contents

List of Abbreviations ....................................................... vi

List of Figures ....................................................................... ix

List of Tables ......................................................................... x

Chapter 1  Introduction ....................................................... 1
  1.1 Innate and Adaptive Immunity in Autoimmune Disease ........ 2
  1.2 Recognition of Nucleic Acids by PRRs ............................. 3
  1.3 Nucleases in Autoimmune Disease ................................. 7
    1.3.1 Ribonuclease H2 ...................................................... 8
    1.3.2 SAMHD1 ............................................................... 13
    1.3.3 TREX1 ................................................................. 14
  1.4 TREX1-mediated Autoimmune Disease ............................ 23

Chapter 2  Insight into DnaQ family DEDDh catalysis from the crystal structures of TREX1 3’ exonuclease autoimmune disease mutants .......................... 27
  2.1 Introduction ................................................................. 30
  2.2 Experimental Procedures ............................................. 33
    2.2.1 Protein Expression and Purification ......................... 33
    2.2.2 Protein Crystallization and X-ray Data Collection ........ 34
    2.2.3 Phasing and Refinement ........................................ 35
  2.3 Results and Discussion .................................................. 36
    2.3.1 Active site rearrangements in the TREX1 apoprotein and mutant substrate complex structures .................. 37
    2.3.2 Two Metal Ion-Mediated Nucleic Acid Hydrolysis and the TREX1 Mechanism ........................................ 46
    2.3.3 Non-processivity and Substrate Trapping in TREX1 ....... 53
    2.3.4 Why Two Catalytic Mechanisms in the DnaQ Family of Exonucleases? ............................................... 55
  2.4 Acknowledgements ........................................................ 56
**Chapter 3**  Dominant TREX1 Autoimmune Disease Mutants Inhibit Wt TREX1 Exonuclease Activity on Double-stranded DNA by Competitively Binding and Protecting DNA 3' Termini

3.1 Introduction .......................... 59
3.2 Experimental Procedures .................. 62
  3.2.1 Enzyme Preparation .................. 63
  3.2.2 Exonuclease Assays .................. 65
  3.2.3 Electrophoretic Mobility Shift Assays (EMSAs) ............. 66
3.3 Results .................................. 66
  3.3.1 Exonuclease activities of TREX1 AGS and FCL mutants using a 30mer oligonucleotide ssDNA ..................... 66
  3.3.2 The TREX1 V201D homodimer mutant exhibits an approximately 4-fold reduction in activity on models of all putative biological substrates ..................... 68
  3.3.3 D18N, D200H, and D200N active site mutant homodimers and heterodimers are inactive on nicked or blunt-ended dsDNA substrates and inhibit wt TREX1 activity on either substrate .......... 70
  3.3.4 Dominant TREX1 D18N, D200H, and D200N homo- and heterodimer activities on single-stranded phagemid DNA .............. 75
  3.3.5 Dominant TREX1 active site disease mutants do not inhibit wt TREX1 activity on single-stranded phagemid DNA .............. 77
  3.3.6 The dominant TREX1 autoimmune disease mutants inhibit wt TREX1 activity on dsDNA by protecting the DNA 3' termini 78
3.4 Discussion ................................ 80
  3.4.1 Dominant TREX1 autoimmune disease is likely triggered by the cells inability to dispose of dsDNA .................. 80
  3.4.2 EMSA assays reveal a mechanism where dominant TREX1 mutants inhibit wt TREX1 activity by protecting the DNA 3' termini of dsDNA substrates .................. 82
  3.4.3 The recessive TREX1 V201D mutant may follow a distinct pathway to autoimmunity .................. 84

**Chapter 4**  Conclusions ................................ 87
  4.1 New developments in the understanding of the role of TREX1 in autoimmune disease .................. 89

Scholastic Vitae ................................. 128
List of Abbreviations

AGS Aicardi-Goutières syndrome
AIM2 absent in melanoma 2
APC antigen presenting cell
ASC apoptosis-associated speck-like protein containing a caspase recruitment domain
CSF cerebrospinal fluid lymphocytosis
DAI DNA-dependent activator of interferon-regulatory factors
DC dendritic cell
DEDDh histidine-containing subgroup of the DnaQ family of exonucleases
DEDDy tyrosine-containing subgroup of the DnaQ family of exonucleases
DKO double knockout
dsDNA double-stranded DNA
dsRNA double-stranded RNA
EDTA ethylenediaminetetraacetic acid
EMSA electrophoretic mobility shift assay
ER endoplasmic reticulum
exo I exonuclease I
6-FAM 6-carboxyfluorescein
FCL familial chilblain lupus
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GzmA</td>
<td>granzyme A</td>
</tr>
<tr>
<td>IFN</td>
<td>interferon</td>
</tr>
<tr>
<td>IFNaR1</td>
<td>type 1 interferon receptor</td>
</tr>
<tr>
<td>IPS-1</td>
<td>interferon-β promoter stimulator 1</td>
</tr>
<tr>
<td>IRF3</td>
<td>interferon regulatory factor 3</td>
</tr>
<tr>
<td>IRF7</td>
<td>interferon regulatory factor 7</td>
</tr>
<tr>
<td>ISD</td>
<td>interferon stimulatory DNA</td>
</tr>
<tr>
<td>MBP</td>
<td>maltose binding protein</td>
</tr>
<tr>
<td>MDA5</td>
<td>melanoma differentiation-associated gene 5</td>
</tr>
<tr>
<td>MEF</td>
<td>mouse embryonic fibroblast</td>
</tr>
<tr>
<td>MR</td>
<td>molecular replacement</td>
</tr>
<tr>
<td>mTREX1</td>
<td>mouse TREX1</td>
</tr>
<tr>
<td>ORN</td>
<td>oligoribonuclease</td>
</tr>
<tr>
<td>PAMP</td>
<td>pathogen associated molecular pattern</td>
</tr>
<tr>
<td>PCNA</td>
<td>proliferating cell nuclear antigen</td>
</tr>
<tr>
<td>PPII</td>
<td>polyproline type II</td>
</tr>
<tr>
<td>PRR</td>
<td>pattern recognition receptor</td>
</tr>
<tr>
<td>RAG2</td>
<td>recombination activating protein 2</td>
</tr>
<tr>
<td>RNase H</td>
<td>ribonuclease h</td>
</tr>
<tr>
<td>RNase H2</td>
<td>ribonuclease h2</td>
</tr>
<tr>
<td>RIG-I</td>
<td>retinoic acid-inducible gene I</td>
</tr>
</tbody>
</table>
RLR    RIG-I-like receptors
RMSD  root mean square deviation
RVCL  retinal vasculopathy and cerebral leukodystrophy
SAM   sterile alpha motif
SLE   systemic lupus erythematosus
ssDNA single-stranded DNA
ssRNA single-stranded RNA
TBK1  tank binding kinase 1
TLR   toll-like receptor
TLS   translation/libration/screw
TNFα  tumor necrosis factor α
TREX1 three prime repair exonuclease 1
TREX2 three prime repair exonuclease 2
wt    wild type
List of Figures

1.1 Signaling of the proinflammatory response by cytosolic nucleic acid PRRs 5
1.2 The structure of RNase H2 ............................................ 11
1.3 The human SAMHD1 protein has tandem SAM and HD motifs ... 13
1.4 The structure of the TREX1 3’ exonuclease ............................. 17

2.1 Superposition of the active sites from the wt TREX1 substrate complex reveals an active and a resting conformation within the dimer .... 37
2.2 The structure of wt TREX1 ............................................... 39
2.3 Active site morphology of TREX1 mutants ............................. 42
2.4 Stabilization of the inward and flipped out conformations of H195 ... 43
2.5 Positioning of the active site loop in DEDDh versus DEDDy exonucleases 49
2.6 Proposed mechanism of the TREX1 homodimer ........................... 52

3.1 The wt TREX1 dimer in complex with ssDNA and calcium .......... 67
3.2 Activities of TREX1 wt and D200H variants on a 30mer ssDNA .... 69
3.3 Activities of the wt and V201D mutant homodimers on dsDNA and ssDNA substrates ......................................................... 71
3.4 The TREX1 D200H/D200H homodimer and wt/D200H heterodimer inhibit wt TREX1 activity on a nicked dsDNA plasmid substrate ... 73
3.5 The wt/D18N, wt/D200H, and wt/D200N TREX1 heterodimers are inactive on a linear, blunt-ended dsDNA substrate ..................... 74
3.6 The D18N, D200H, and D200N TREX1 variants inhibit wt TREX1 activity on a linear, blunt-ended dsDNA ........................................ 76
3.7 TREX1 D18N, D200H, and D200N homodimer, but not heterodimer, variants are inactive on a large ssDNA ....................................... 77
3.8 The D18N, D200H, and D200N TREX1 variants do not inhibit wt TREX1 activity on a linear ssDNA ........................................... 79
3.9 The D18N, D200H, and D200N dominant TREX1 variants inhibit TREX1 wt activity by binding and protecting the 3’ termini of a dsDNA 81
3.10 The wt TREX1 and recessive V201D TREX1 mutant protein bind dsDNA in the absence of metal .............................................. 83
# List of Tables

1.1 Putative TREX1 Biological Substrates ............................................. 18

2.1 Crystallographic Data and Refinement Statistics ............................ 35

3.1 Relative Activities of TREX1 Wt and Disease Mutants on a 30mer ssDNA ................................................................. 70
Abstract

The homodimeric TREX1 protein is a member of the DnaQ family of exonucleases and catalyzes the major 3′ exonuclease activity detected in mammalian cell extracts. Mutations within the Trex1 gene are the underlying cause of multiple autoimmune diseases including Aicardi-Goutières syndrome (AGS), familial chilblain lupus (FCL), systemic lupus erythematosus (SLE), and retinal vasculopathy and cerebral leukodystrophy (RVCL). Several cellular roles have been proposed for TREX1 including degradation of nicked genomic DNA during the granzyme A cell death pathway, disposing of single-stranded DNA aberrant replication intermediates, and degrading DNA derived from endogenous retroelements. Four conserved acidic residues (D18, E20, D130, D200) coordinate two divalent cations within the TREX1 active site. Metal ion A directs nucleophilic attack on the scissile phosphate and metal ion B stabilizes the pentacovalent intermediate. A conserved histidine residue in the TREX1 active site activates a water molecule coordinated to metal A to initiate catalysis. Specific mutations within the vicinity of the TREX1 active site have been linked to dominant AGS (D200H, D200N), recessive AGS (V201D), and FCL (D18N). Recent evidence suggests that dominant TREX1 active site mutants are defective in their double-stranded DNA degradation activity.

This work addresses several important questions pertaining to TREX1 catalysis and how TREX1 mutation within the active site may lead to autoimmunity. What role does the binding of metal and DNA play in arranging the TREX1 active site
for catalysis? Are there global or local structural changes in the TREX1 protein resulting from mutation that might inhibit TREX1 cellular function? Does the context of the DNA 3’ terminus affect mutant TREX1 exonucleolytic activity? Can the dominant TREX1 mutants inhibit wild type TREX1 activity on both single- and double-stranded DNA substrates?

The X-ray crystal structures of the wt apoprotein as well as TREX1 active site mutants were determined in order to identify structural changes that may be important in TREX1-mediated autoimmunity. The catalytic histidine likely acts as a switch between an active and a resting state for a TREX1 protomer. The apparent loss of mobility in this residue observed in the TREX1 mutant structures may contribute to diminished catalytic activity. Mutation within the active site vicinity alters (D18N and V201D) or precludes (D200H and D200N) the coordination of metal ion A and likely contributes to the observed decreased catalytic function of the mutant enzymes.

The activities of the TREX1 active site mutants were determined on model substrates to evaluate the possible contributions of different proposed TREX1 substrates to the development of an aberrant immune response. The dominant active site mutants are inactive on dsDNA regardless of the context of the 3’ terminus. Furthermore, these dominant mutants are able to inhibit wt TREX1 activity on double-, but not single-stranded DNA substrates by competitively binding and protecting the DNA 3’ termini. The accumulation of double-stranded DNA in the cell due to a loss of TREX1 function may result in the aberrant recognition of self DNA that could stimulate an autoimmune response.
Chapter 1: Introduction
More than 80 chronic autoimmune diseases afflict 5-8% of the United States population, and the incidence of autoimmune disease is on the rise. Autoimmune diseases can be organ-specific such as type 1 diabetes, which affects the pancreas, or multiple sclerosis, which affects the central nervous system. Non-organ-specific diseases, such as systemic lupus erythematosus (SLE), attack organs and tissues throughout the body and can result in widespread injury. Pathogenic autoantibodies are common in autoimmune disease patients and can directly injure tissue or amplify an autoimmune response. Systemic autoimmune diseases similar to SLE commonly exhibit pathogenic antibodies to nucleic acids and increased levels of type 1 interferons (1). This highlights the importance of understanding the underlying mechanisms of nucleic acid-associated systemic autoimmune disease. Recent advances in this field of research have identified a role for certain nucleases in the elimination of nucleic acids to avoid systemic autoimmune diseases including SLE.

1.1 Innate and Adaptive Immunity in Autoimmune Disease

The innate and adaptive components of the immune system work in concert to recognize an invading pathogen, initiate a pro-inflammatory response, eliminate the pathogen, and generate an immunologic memory. The innate immune response provides the first line of host defense. Specific cell types including antigen-presenting cells (APCs), granulocytes, macrophages, and dendritic cells (DCs) participate in the coordination of the innate immune response (2, 3). The recognition of an invader (including viruses, bacteria, fungi or parasites) via evolutionarily conserved structures called pathogen associated molecular patterns (PAMPS) is important for the
initiation of an innate immune response (3, 4). PAMPs are distinctive features of an invading pathogen that the cell recognizes using pattern recognition receptors (PRRs) (3, 5). PRRs can be categorized into secreted, transmembrane, and cytosolic classes (2, 6). Motifs that are recognized by PRRs include bacterial cellular membrane components, flagella, single-stranded RNA (ssRNA), double-stranded RNA (dsRNA), and DNA (2, 6). The adaptive immune response is stimulated over time and enables immunologic memory of a particular pathogen via antibody production. The signaling pathways that are induced by the binding of PRRs by PAMPs control certain DC functions that are critical for the activation of the adaptive immune response (3, 7). The aberrant activation of innate and adaptive immunity causes autoimmune disease and autoantibody production that is the basis for immunologic memory of host nucleic acids. The accumulation of self nucleic acids in the cytosol may cause an autoimmune response in the absence of an invader due to the lack of specificity of nucleic acid recognition by PRRs for pathogenic nucleic acids.

1.2 Recognition of Nucleic Acids by PRRs

Self/nonself discrimination is straightforward for receptors that bind unique bacterial products, but there is little structural basis for discrimination between host and pathogen-derived nucleic acids. The compartmentalization of nucleic acid receptors from the extracellular matrix probably plays a role in the process of nucleic acid discrimination (8–10). PRRs that recognize nucleic acids are predominantly localized to intracellular compartments (e.g. endosomes) and the cytoplasm. The toll-like receptors (TLRs) are the most extensively studied class of PRRs in mammalian species.
TLRs are transmembrane proteins (4, 13) that are predominantly expressed in sentinel immune cells including DCs and macrophages (8). Nucleic acid sensing TLRs are almost exclusively found in intracellular compartments where they sample endosomal cargo for nucleic acids (3, 14). TLRs 3, 7, 8, and 9 all recognize nucleic acids (12, 15). TLR3 recognizes dsRNA, TLR7/8 recognize ss-RNA, and TLR9 recognizes unmethylated CpG repeats in DNA and may be capable of sequence-independent recognition of DNA (7, 11, 12, 16–20). TLR9 stimulates a MyD88 and interferon regulatory factor 7 (IRF7)-dependent type 1 interferon (IFN) response upon DNA binding (21–24).

TLR-independent mechanisms for the recognition of nucleic acids involve cytosolic receptors. Retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs) include the interferon-inducible helicases RIG-I and melanoma differentiation-associated gene 5 (MDA-5) that are predominantly involved in the recognition of RNA in the cytosol (25–27). The recognition of dsRNA or 5′-triphosphate RNA by RLRs triggers an antiviral response through the adaptor protein IPS-1 (interferon-β promoter stimulator 1; also known as MAVS/VISA/Cardiff) that activates IRF3 (interferon regulatory factor 3), IRF7, and NF-κB to stimulate type I IFN production (Figure 1.1) (3, 4, 25, 28, 29).

A putative mechanism of cytosolic DNA recognition is via RNA polymerase III and RIG-I (30–32). Cytosolic RNA polymerase III may bind AT-rich DNA and transcribe it into a dsRNA substrate with a 5′-triphosphate; a potent inducer of type 1 IFN via interaction with RIG-I. RNA polymerase III sensing of cytosolic DNA describes
Figure 1.1: Signaling of the proinflammatory response by cytosolic nucleic acid PRRs. DsRNA binds both the MDA-5 and the RIG-I receptors to stimulate type 1 IFNs and a proinflammatory response. The RIG-I induction of type 1 IFNs can also be stimulated by 5′-triphosphate-dsRNA (5′-3P-dsRNA) and 5′-3P-ssRNA. DsDNA is recognized by the cytosolic RNA polymerase III, which transcribes it into 5′-3P-dsRNA, which can then stimulate type 1 IFN induction via RIG-I. DsDNA and possibly ssDNA stimulate the ISD response via DAI or another, unidentified receptor. AIM2 recognizes dsDNA in the cytosol, assembles the AIM2 inflammasome, and stimulates a type 1 IFN-independent proinflammatory response.

a probable mechanism of host defense against bacteria and DNA viruses (30, 31).

The protein AIM2 (absent in melanoma 2) preferentially recognizes double-stranded DNA (dsDNA) in the cytosol and is a member of the AIM2 inflammasome (33–37). In addition to AIM2, the inflammasome is composed of an adaptor protein ASC
(apoptosis-associated speck-like protein containing a caspase recruitment domain) and pro-caspase-1. The inflammasome activates caspase-1, which can process pro-IL-1β into the proinflammatory cytokine IL-1β \( (33, 37) \). Prolonged activation of caspase-1 will eventually lead to pyroptotic cell death, a cell death pathway resulting from an immune response \( (38-40) \). The AIM2 inflammasome may also trigger apoptosis via the activation of caspase-3 \( (33, 36) \), but is not involved in the induction of type I interferon.

The interferon stimulatory DNA (ISD) response is a cytosolic pathway that stimulates a proinflammatory response to DNA analogous to the RIG-I and MDA5 pathway for RNA. However, ISD signaling of type 1 IFN induction via IRF3 is not dependent on IPS-1 (Figure 1.1) \( (1, 26, 41, 42) \). It is proposed that the pathway is initiated by the DNA-dependent activator of IFN-regulatory factors (DAI, also known as DLM-1/ZBP1) that likely recognizes cytosolic dsDNA and is a candidate ISD sensor \( (43, 44) \). DAI activates a TBK1 (tank binding kinase 1)/ IRF3-dependent type I IFN response (Figure 1.1) \( (43) \). However, IRF3 activation after exposure to dsDNA is not DAI-dependent, suggesting that there are DAI-independent cytosolic dsDNA sensing mechanisms that also stimulate a type I IFN response \( (44) \).

The localization of nucleic acid-binding PRRs in the cytosol underscores the importance of keeping levels of self nucleic acids in the cytosol low. Toll-like receptors and cytosolic nucleic acid sensors, together, account for virtually all IFN-mediated antiviral immunity \( (45) \). The activation of the transcription factors IRF3 and NF-κB by most PRRs is sufficient to stimulate both T and B cell responses \( (2, 5) \). Thus,
keeping levels of nucleic acids in the cytosol low may be critical to avoid both autoim-
mune activation and the stimulation of autoantibody production that would provide 
an immunologic memory of host DNA.

1.3 Nucleases in Autoimmune Disease

Certain nucleases may play important roles in the disposal of DNA and RNA to 
avoid an aberrant immune response. Cytosolic and extracellular nucleases may play 
a major part in the disposal of nucleic acids during normal cellular metabolism or 
during programmed cell death that keeps self DNA and RNA from binding PRRs and 
stimulating an autoimmune response. Nucleases that have putative roles in nucleic 
acid elimination to avoid autoimmune activation include DNase I, DNase II, TREX1  
(DNase III), RNase H2, and the putative phosphohydrolase SAMHD1.

DNase I is an extracellular endonuclease that is the primary enzyme involved in 
clearance of extracellular DNA. DNase I-deficient mice exhibit a phenotype that re-
sembles systemic lupus erythematosus including the presence of antinuclear antibodies 
and the deposition of immune complexes in glomeruli (9). There is evidence that some 
SLE patients have low serum levels of DNase I activity (46). Furthermore, mutations 
within the *DNase I* gene have been identified in patients with SLE, confirming the 
connection between DNase I and systemic autoimmune disease (47, 48). A common 
feature of SLE patients with DNase I mutations are anti-dsDNA autoantibodies that 
likely result from the accumulation of extracellular DNA (47, 48).

DNase II is a lysosomal endonuclease with prominent expression in macrophages 
(49, 50). The cellular role of DNase II is thought to be in degradation of DNA from
endocytosed apoptotic cells and degradation of nuclear DNA ejected from erythroid precursor cells during erythropoiesis (49, 51–53). Mice deficient in DNase II exhibit a phenotype resembling rheumatoid arthritis (53). The deletion of DNase II from mice causes an accumulation of DNA expelled from erythroid precursor cells and subsequent IFN-\(\beta\) induction, resulting in lethal anemia in the embryos (49, 50, 52, 54). Double knockout mice that are deficient in both DNase II and the IFN receptor as well as mice that have the *DNase II* gene inactivated as adults develop chronic polyarthritis (53).

The TREX1 (DNase III) 3′ exonuclease is ubiquitously expressed with a perinuclear localization. TREX1 knockout mice develop inflammatory myocarditis and die prematurely from cardiomyopathy and circulatory failure (55). Mutations in the *Trex1* gene, the three non-allelic components of Ribonuclease H2 (RNase H2), and the *Samhd1* gene are the five loci that have been identified in patients with Aicardi-Goutières syndrome (AGS) (56–58). Furthermore, mutations in the *Trex1* gene sequence have been identified in patients with other autoimmune disorders including familial chilblain lupus (FCL), systemic lupus erythematosus, and retinal vasculopathy and cerebral leukodystrophy (RVCL) (56, 59–64). The three nucleases implicated in AGS – RNase H2, SAMHD1, and TREX1 – will be the focus of the rest of this introduction.

### 1.3.1 Ribonuclease H2

Ribonuclease H2 is an endoribonuclease with putative roles in removing ribonucleotides from Okazaki fragments during lagging strand DNA synthesis, removing
single ribonucleotides from DNA:DNA hybrids that have been misinserted by DNA polymerases, and disposing of RNA:DNA hybrids during cell death. There are two types of ribonuclease H (RNase H) in the cell – type 1 (or I) and type 2 (or II) (for a review see (65)). Briefly, type 1 RNase H recognizes and removes a minimum of four consecutive ribonucleotides within an RNA:DNA hybrid with a 5′-3′ polarity. RNase H2 enzymes can recognize a single ribonucleotide within a DNA:DNA hybrid and will cleave the RNA-containing strand one nucleotide away from the RNA-DNA junction.

Ribonucleotides may be the most common noncanonical nucleotides that are introduced into the eukaryotic genome by polymerases during DNA replication and repair (66). RNase H2 activity may be critical for the repair of these DNA replication and repair errors, which can alter the geometry of the DNA helix and affect the efficient and accurate replication of the genome (66, 67). In cell extracts, RNase H2 removes RNA primers from an RNA:DNA duplex in cooperation with FEN-1 (human) or RAD27 (yeast). This suggests a role in processing Okazaki fragments during DNA replication (68–70). An interaction between RNase H2 and proliferating cell nuclear antigen (PCNA) indicates that RNase H2 may participate during DNA replication and repair by excising misincorporated ribonucleotides (66, 71–73).

Eukaryotic RNase H2 enzymes are encoded by three subunits (RNase H2A, B, and C) to form a fully active ribonuclease activity (57, 65, 74, 75). RNase H2 subunit A is the catalytic subunit of the enzyme and has high sequence homology with the prokaryotic RNase H2 proteins that require a single subunit for activity (75). Subunits B and C have putative roles in scaffolding and protein-protein interactions important
for RNase H2 activity in the cell. Mutations within all three RNase H2 subunits have been identified in patients with AGS (57). The genes encoding these three subunits are three of the five AGS loci – AGS2 (RNASEH2B), AGS3 (RNASEH2C), and AGS4 (RNASEH2A).

AGS is a recessive encephalopathy that mimics the sequelae of congenital infection (57, 63, 76–79). Cases of dominant AGS have also been reported (57, 63, 80). AGS is an early onset disorder, presenting at birth or within the first few years of life. Generally patients with TREX1 mutations exhibit disease onset within the first few months after birth, while patients with mutations in the B subunit of RNase H2 generally have a later disease onset (76, 77, 79). Disease onset most often occurs over the course of a few months with extensive neurological damage including cerebral calcification, cerebral atrophy and leukodystrophy (76, 78, 81). After the first few months, the condition stabilizes and there is no further neurological damage or regression (63, 76, 79). Other common symptoms of AGS include seizures, cerebrospinal fluid (CSF) lymphocytosis, and high levels of IFNα in the CSF (63, 76, 79). Approximately 40% of AGS patients die in early childhood (78, 82).

The X-ray crystal structure of murine RNase H2 reveals an architecture where subunit A interacts directly with subunit C. The C subunit functions as a connector between subunits A and B (Figure 1.2b). The catalytic subunit, subunit A, adopts an RNase H-like fold and exhibits an active site geometry that suggests a two metal ion mechanism for catalysis (75). Based on the architectural arrangement of the RNase H2 complex, the central β-sheet of subunit A extends into solution, suggesting a
Figure 1.2: The structure of RNase H2. (A) Structural features within each of the three RNase H2 subunits are highlighted. The 1, 2, and 3 highlighted in orange within subunit A mark the positions of the metal-coordinating residues within the active site. The C-terminal exposed β-strand in subunit A, as well as the disulfide bond are indicated. The triple-barrel motif in subunit A mediates the interaction with subunit C. Selected AGS mutations within each subunit are indicated. (b) The crystal structure of RNase H2 at 3.1 Å. Subunit A is represented in light orange, subunit C in blue, and subunit B in green. The location of the active site within subunit A is indicated with an arrow.
potential site for an interaction with an unknown protein (75).

The RNase H2 structure provides insight into possible changes in RNase H2 cellular function due to certain AGS mutations. A mutation near the putative RNase H2 active site in the A subunit has been identified in patients (57). The G37S mutation in subunit A likely disrupts the active site morphology to cause the observed decrease in activity and altered substrate specificity (57, 75, 83). The K162T mutation in subunit B is near the protein surface, in the middle of a long helix. K162 may participate in stabilizing the packing of this helix to the core of the protein, an interaction that might be disrupted when the lysine is mutated to threonine. The mutation of R69W in the C subunit is near the interface with the A subunit and is exposed to solution (75). This arginine most likely participates in a potential protein-protein interaction at the nearby, exposed β-strand on subunit A (75). Many of the other known mutations are located within disordered loops in the structure and are likely involved in important protein-protein interactions within the cell (75). While a loss of catalytic function does not seem to be the direct cause of AGS resulting from RNase H2 mutation (83), it is likely that disruption of normal cellular protein-protein interactions that are mediated by flexible loops and the solvent-exposed β-strand of subunit A may contribute to an effective loss in RNase H2 cellular function. A deficiency in RNase H2 function may result in the accumulation of immunogenic RNA:DNA hybrids that stimulate an aberrant immune response.
1.3.2 SAMHD1

The Samhd1 gene encodes a 626 amino acid protein that is exclusively localized to the nucleus (58). SAMHD1 has a putative role in the immune response based on results showing upregulation of SAMHD1 expression in response to viral infection (84–86) and a function in tumor necrosis factor α (TNFα) regulation during a proinflammatory response (87). Furthermore, when macrophages derived from MyD88 knockout mice were transfected with immunostimulatory DNA, Samhd1 expression was upregulated suggesting that SAMHD1 is involved in the ISD response (58). SAMHD1 is the only known human protein with a tandem SAM (sterile alpha motif) domain and HD domain (Figure 1.3.2) (58). SAM domains have known roles in protein-protein and protein-nucleic acid interactions (88). HD domain proteins compose a diverse superfamily with conserved metal-coordinating histidine and aspartic acid residues that have a putative role in catalysis (58). Members of this superfamily have predicted or known phosphohydrolase activity (58, 89, 90). Mutations within the gene encod-
ing SAMHD1 have been linked to AGS in patients (58), making the Samhd1 gene the AGS5 locus. SAMHD1 protein from a patient with the Q149X mutation (Figure 1.3.2) maintains nuclear localization suggesting that the N-terminal 149 residues are sufficient to confer nuclear localization (58). The catalytic function of the human SAMHD1 protein has not been established, therefore, how SAMHD1 mutations implicated in AGS affect its catalytic activity is unknown.

1.3.3 TREX1

The TREX1 protein catalyzes the major 3′ exonuclease activity detected in mammalian cell extracts. TREX1 has a close homolog, TREX2, that shares 40% amino acid sequence identity, but is missing a unique C-terminal extension that is present in TREX1 (91). Although TREX1 and TREX2 are structurally similar, evidence supports different biological roles for these two proteins. TREX1 is involved in processing DNA to avoid immune dysfunction. TREX2 null mice suggest a role for TREX2 in preventing skin carcinogenesis (92). The TREX enzymes are homodimeric exonucleases with activity on both single-stranded DNA (ssDNA) and dsDNA (91, 93). Both TREX enzymes are members of the DnaQ family of exonucleases. Members of the DnaQ family have four conserved acidic residues within three exonuclease motifs that participate in the coordination of two divalent cations within the enzyme active site (93–95). Metal ion \( \alpha \) directs nucleophilic attack on the scissile phosphate of the substrate (94, 96–98). Metal ion \( \beta \) is important for stabilizing the pentacovalent intermediate. Other known members of the DnaQ family include the WRN exonuclease (99), the klenow fragment of DNA polymerase I (96), and the bacterial epsilon
subunit of DNA polymerase III (100). A fifth conserved residue divides the DnaQ family into two subgroups based on the presence of a conserved histidine (DEDDh) or tyrosine (DEDDy) in the third exonuclease motif. While in some cases two metal ions alone may be sufficient to generate the nucleophile and direct nucleophilic attack during two metal ion catalysis (97, 101), the conserved histidine and tyrosine have proposed roles in contributing to the catalytic activity of the DnaQ family exonucleases (96). Both TREX1 and TREX2 are members of the DEDDh subgroup in which the conserved histidine is important for generation of the nucleophile (95, 102–104).

The X-ray crystal structures of the TREX1 and TREX2 exonucleases have been determined (94, 95, 105, 106). Symmetry in the TREX dimer positions the active sites at opposite outer edges of the same face of the molecule (Figure 1.4). Each TREX monomer adopts an RNase H-like fold, containing a central, antiparallel β-sheet surrounded by nine α-helices. Both TREX1 and TREX2 have an extensive dimer interface that buries over 1500 Å² of surface area (94, 95). Attempts to generate a monomeric TREX2 molecule have been unsuccessful, suggesting that dimerization may be important for TREX function (103). A polyproline type II (PPII) helix is positioned adjacent to the dimer interface in TREX1 that is absent in TREX2 (94, 106). This PPII helix has a putative role in mediating interactions with other proteins in the cell. This same region in TREX2 adopts a β-hairpin conformation and may serve as a distinguishing element in recruiting TREX1 and TREX2 to their respective biological pathways (95). TREX1 has an extended C-terminal domain that is important for subcellular localization on the cytosolic side of the endoplasmic retic-
ulum (ER) (62, 64, 107). TREX2, which lacks this C-terminal extension, is present in both the nucleus and cytoplasm (108, 109). These unique structural features in TREX1 suggest distinct cellular roles for these two exonucleases. TREX1 has various proposed cellular roles (which will be discussed below), which may be important to avoid an autoimmune response.

*Putative TREX1 biological roles*

Over the last five years several biological roles have been proposed for the TREX1 3′ exonuclease including granzyme A cell death, disposal of cytosolic DNA-derived from endogenous retroelements, and elimination of aberrant ssDNA replication intermediates in the cytosol (Table 1.1). The TREX1 protein is a member of the ER-associated SET complex (110) that functions in the caspase-independent granzyme A (GzmA) cell death pathway (110), and possibly in the oxidative stress response (111). As a member of this complex, TREX1 has a putative interaction with the nucleosome assembly protein, SET (110). The SET complex ranges in size from 270-420 kD. Known members of the SET complex include the endonucleases Ape1 and NM23H1, the tumor suppressor protein pp32, and the DNA bending protein HMG-2 (110). TREX1 has been identified as the exonuclease required for the elimination of nicked genomic DNA during the GzmA cell death pathway (110). GzmA-mediated cell death can kill cells that are resistant to caspase-mediated apoptosis (110, 112, 113), suggesting a possible role for this pathway in immune defense against cancers and viruses that are able to evade caspase-mediated apoptosis (110).
Figure 1.4: The structure of the TREX1 3′ exonuclease. (A) The primary sequence of TREX1 is shown with highlighted secondary structure elements, including the three exo motifs that contain the acidic residues involved in metal-coordination in the TREX1 active site. Selected AGS, FCL, and SLE mutations are indicated. Exo: exonuclease motifs (1, 2, and 3); DBL: DNA binding loop (B) The X-ray crystal structure of the homodimeric TREX1 protein, where one monomer is shown in blue and the other in gray. The active site metals are shown as yellow spheres. The active site of TREX1 is located in the boxed area and is expanded to the right to show important active site residues.

GzmA is a serine protease that is stored in the cytolytic granules of natural killer cells and cytotoxic T lymphocytes and activates cell death when it is delivered into the cytoplasm of target cells by perforin. Once inside the cell, GzmA induces a rapid increase in reactive oxygen species and a loss of mitochondrial transmembrane potential (113). Oxidative stress triggers the release of the SET complex from the endoplasmic reticulum and its translocation to the nucleus (110, 113, 114). In the nucleus, NM23H1 generates single-stranded nicks in the genomic DNA (111, 115),
Table 1.1: Putative TREX1 Biological Substrates

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Size</th>
<th>Derived from</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicked dsDNA</td>
<td>Genomic DNA</td>
<td>GzmA cell death; other cell death pathways?</td>
<td>(110)</td>
</tr>
<tr>
<td>ssDNA</td>
<td>60-65mer</td>
<td>Aberrant replication intermediates</td>
<td>(117)</td>
</tr>
<tr>
<td>ssDNA or dsDNA?</td>
<td>100-10,000-bp</td>
<td>Endogenous retroelements</td>
<td>(107)</td>
</tr>
</tbody>
</table>

providing a suitable substrate for TREX1 to degrade, preventing DNA repair and ensuring cell death. Similarly, a study investigating a role for TREX1 in DNA repair in drug-treated tumor cells suggested a role for TREX1 in degrading DNA from dying cells before phagocytosis occurs (116). The absence of TREX1 activity may result in the accumulation of nicked double-stranded DNA degradation intermediates from dying cells that could stimulate an aberrant immune response.

TREX1 has also been implicated in the degradation of 60-65 nucleotide ssDNA aberrant replication intermediates to avoid chronic ATM checkpoint signaling (117, 118). TREX1 may be recruited to stalled replication forks after DNA damage where it degrades ssDNA to resolve the fork and enable replication to resume (117). In mouse embryonic fibroblast (MEF) cells lacking TREX1, ssDNA of a discrete size (60-65 nucleotides) that originated during S phase accumulated around the ER in the absence of DNA damage. Single-stranded DNA also accumulates in the cytosol of fibroblast cells derived from AGS patients (117). The accumulation of ssDNA in the cytosol in the absence of TREX1 activity suggests that under normal cellular conditions TREX1 does not degrade DNA at the replication fork, but may instead primarily function in the cytosol (117). Furthermore, TREX1 null MEFs or fibroblast cells derived from AGS patients with inactivating TREX1 mutations accumulated in G0/G1 with a reduced S phase population and exhibited chronic activation of checkpoint proteins.
Together these results suggest a role for TREX1 in degrading ssDNA aberrant replication intermediates in the cytosol to avoid chronic ATM checkpoint signaling and prevent the accumulation of ssDNA that may trigger an autoimmune response. However, others have not observed cell-cycle abnormalities or chronic DNA damage checkpoint activation in TREX1-deficient primary murine embryonic fibroblasts (107), leaving this an area of contention.

TREX1 may be a negative regulator of the interferon-stimulatory DNA response and DNA accumulating in the absence of TREX1 activity may stimulate type 1 IFN production via a TLR-independent mechanism (Figure 1.1) (107). TREX1 is an ISD binding protein and Trex1 gene transcription is robustly induced by ISD stimulation. TREX1 knockout mice die from lethal autoimmunity at a median age of 9 weeks (107). SsDNA fragments derived from endogenous retroelements that accumulated in the hearts of the null mice were identified as potential TREX1 substrates, although, it is possible that the complementary strands were missed during sequencing and that the biological TREX1 substrate is dsDNA (107). The mortality of TREX1 null mice is rescued in TREX1/IRF3 double knockout (DKO) mice, suggesting that type 1 interferon induction resulting from the loss of TREX1 is IRF3 dependent. Similarly, TREX1/IFNαR1 (type 1 interferon receptor) DKO mice were fully protected from premature mortality, further supporting the role for IFNα in TREX1-mediated autoimmunity. This is consistent with the increased levels of IFNα observed in patients with AGS and SLE. Furthermore, the deletion of both TREX1 and RAG2 (recombination activating protein 2) rescues mice from mortality. RAG2 is involved in the
initiation of V(D)J recombination during B and T cell development (119). Thus, the rescue of TREX1-deficient mice by the deletion of RAG2 demonstrates that lymphocytes are essential for TREX1-mediated disease. This implicates the IFN-dependent mobilization of lymphocytes as the critical step in triggering autoimmunity since IFNα induction is not sufficient to cause immunity in the TREX1/RAG2 DKO (107). The accumulation of single- or double-stranded DNA derived from endogenous retroelements in the absence of TREX1 stimulates the ISD response and may be a major contributor to autoimmunity in the absence of TREX1 activity.

TREX1 mutations associated with autoimmune disease

Mutations within the *Trex1* gene were identified in 2006 at the AGS1 locus over 20 years after the disease was first described (56, 79). Since then, mutations throughout the TREX1 sequence have been identified in patients with Aicardi-Goutières syndrome, familial chilblain lupus, systemic lupus erythematosus, and retinal vasculopathy and cerebral leukodystrophy (60–64, 78, 80, 120). Abnormally high levels of IFNα are observed in patients with AGS and SLE (1, 78–80, 82) and antinuclear antibodies to dsDNA are observed in patients with AGS, FCL, and SLE (60–62, 79, 80, 82). Systemic lupus erythematosus is a chronic inflammatory disease that is influenced by both genetic and environmental factors. SLE is nine times more common in women than in men. Dominant TREX1 mutations account for about 2% of SLE cases, but are responsible for the majority of cases of monogenic SLE (82, 121). Familial chilblain lupus is an autosomal dominant disorder that results in painful bluish-red lesions in
acral locations (62, 122). Similar lesions are observed in about 40% of AGS patients (79, 82, 122). Both AGS and FCL manifest in early childhood, whereas SLE has a later onset, most commonly in the mid-twenties.

RVCL is distinct from other known TREX1-associated diseases in its late disease onset and clinical manifestations. RVCL exhibits autosomal dominant inheritance and patients typically present at middle age with a progressive retinal vasculopathy that causes loss in visual acuity and variable neurological changes including white matter abnormalities (64). Patients with RVCL may experience microaneurysms, stroke, and dementia (64). Death occurs in most families 5 to 10 years after disease onset.

The TREX1 mutations that have been implicated in AGS are located throughout the Trex1 gene sequence including the active site (eg. D200H, D200N, V201D), dimer interface (R114H), and C-terminal tail (56, 59, 61–64, 80, 120). All mutations seem to result in an equally severe AGS phenotype in patients regardless of the localization of the mutation. A single mutation within the TREX1 active site has been implicated in FCL (D18N) (59, 61) and TREX1 mutations causing SLE are predominantly located in the C-terminal tail region of TREX1 (62, 80). Other SLE mutations are located at the TREX1 dimer interface (eg. R114H, E198K) and near the TREX1 active site (R128H) (62, 120). To date, mutations that have been implicated in RVCL have been localized exclusively to the TREX1 C-terminal tail (64). The RVCL mutation closest to the C-terminus results in the loss of the perinuclear localization and instead leads to diffuse distribution of TREX1 throughout the cell (64), emphasizing the importance
of the subcellular localization of TREX1 for normal cellular function.

Dominant TREX1 mutations have been identified in patients with *de novo* AGS, FCL, SLE, and RVCL. Additionally, recessive mutations in the *Trex1* gene have been implicated in inherited AGS. The pattern of TREX1 allele expression probably determines the composition of TREX1 dimers in the cells of patients with dominant TREX1 mutations. Simultaneous expression of both the wild type (wt) and mutant alleles could result in wt/wt homodimer, mutant/mutant homodimer, and wt/mutant heterodimer TREX1 protein. The activities of both the mutant/mutant homodimer and the wt/mutant heterodimer and how they behave in the presence of the wt/wt TREX1 protein may contribute to the autoimmune phenotype.

TREX1 variants that are associated with an autoimmune phenotype have a broad spectrum of TREX1 catalytic activity. Mutations within the C-terminal domain of TREX1 that cause RVCL in patients exhibit no reduction in catalytic activity, but likely result in the functional loss of TREX1 activity due to improper subcellular localization (64). The R114H mutation at the TREX1 dimer interface causes SLE when a patient is heterozygous for the mutation and is the most common homozygous mutation observed in AGS (56, 62). The homodimer R114H/R114H TREX1 mutant exhibits a 34-fold reduction in activity on a single-stranded 30mer DNA and a more dramatic, 300-fold reduction in exonuclease activity on a nicked dsDNA substrate (94, 123). The wt/R114H TREX1 heterodimer protein has an approximately 3-fold reduction in activity on a ssDNA 30mer, but has an activity on a nicked dsDNA substrate that is indistinguishable from wt TREX1 activity (123). These results are
consistent with the fact that the homozygous inheritance of the R114H allele causes AGS and the heterozygous inheritance of R114H causes the less severe SLE phenotype.

The ability of the R114H mutation to diminish TREX1 catalytic activity when it is at a distance from the active site suggests that communication across the dimer interface may be important for TREX1 catalysis. Furthermore, R114 may participate in transmitting a signal across the dimer interface during the catalytic cycle. The decrease in R114H homodimer catalytic activity on a ssDNA and its severely retarded rate of dsDNA degradation probably results in the accumulation of DNA in the cell that could stimulate aberrant autoimmunity in AGS. The presence of undetermined levels of the R114H/R114H homodimer in the cells of SLE patients may also explain this autoimmune phenotype (123).

The dominant mutation of the active site metal coordinating residues, D18N and D200N, cause FCL and de novo AGS, respectively. Homodimeric D18N/D18N and D200N/D200N are inactive on both ssDNA and dsDNA substrates (94, 123). The wt/D18N and wt/D200N heterodimer mutants retain approximately 50% of wt TREX1 activity on ssDNA, but have no activity on a nicked dsDNA substrate (123). Furthermore, both the D18N and D200N homodimer and heterodimer TREX1 mutants inhibit wt TREX1 activity on the nicked dsDNA substrate. This suggests a primary role for TREX1 in the degradation of dsDNA in vivo.

### 1.4 TREX1-mediated Autoimmune Disease

Mutations throughout the Trex1 gene sequence have been identified in patients with a spectrum of autoimmune disorders including AGS, FCL, SLE, and RVCL. Common
features of AGS, FCL, and SLE include autoantibodies to dsDNA. To our knowledge, this feature has not been reported in patients with RVCL. This suggests a common pathogenic mechanism in AGS, FCL, and SLE. Patients with RVCL, who exhibit a much later disease onset, may follow a different pathway to autoimmunity. The TREX1 3’ exonuclease efficiently degrades ssDNA as well as dsDNA substrates with available 3’ termini. TREX1 is most active in vitro on a dsDNA with mispairs at the 3’ terminus, (93).

Dominant TREX1 autoimmune disease mutant homodimers are inactive on both ssDNA and nicked plasmid dsDNA substrates, however the wt/mutant TREX1 heterodimers retain approximately 50% of wt activity on a ssDNA, but are inactive on the nicked plasmid substrate (61, 123). Furthermore, the heterodimeric TREX1 mutants are able to inhibit the activity of wt TREX1 on the nicked dsDNA substrate. The dominant mutant activities suggest that TREX1 is important for dsDNA degradation in vivo, in contrast to two of the putative TREX1 roles that suggest TREX1 is primarily a single-stranded DNase (107, 117).

There is evidence that TREX1 deficiency results in the accumulation of DNA in the cytosol that in turn activates the cytosolic ISD response (107). This immune response results in the induction of type 1 IFN via the adaptor protein IRF3 (107). This is in contrast to both TLR9-mediated induction of type 1 IFN, which is IPS-1-dependent, and the AIM2 inflammasome, which does not have a known role in the induction of type 1 IFNs (34, 35, 37). It is possible that the autoimmune response in AGS that results from either TREX1 or RNase H2 mutation may follow the RNA
polymerase III/RIG-I pathway to autoimmunity, which requires the transcription of cytosolic DNA into RNA that is recognized by RIG-I. RIG-I-mediated induction of type 1 IFNs, however, requires the adaptor protein IPS-1, which is not required for the induction of type 1 IFN in TREX1 null mice (107). Thus, RIG-I-mediated signaling due to TREX1 mutation is unlikely. While the ISD response is the primary candidate for the detection of cytosolic DNA due to loss of TREX1 cellular function, it is yet to be seen if this response is present in TREX1 autoimmune disease patients as well as TREX1 null mice.

Much remains to be discovered about how mutations within the TREX1 sequence lead to the autoimmune disease phenotype. Structural changes within the TREX1 protein due to disease mutations may affect TREX1 function in vivo. Furthermore, the biological substrate of TREX1 remains unclear. Whether or not the context of the 3′ end of a dsDNA substrate (eg. blunt ended dsDNA instead of a nick) may affect the ability of a mutant to degrade the DNA remains an unanswered question. Whether TREX1 homodimers and heterodimers are inactive on a large ssDNA, similar to the nicked plasmid is also unknown. The dominant TREX1 homodimer mutants are inactive on the 30mer ssDNA, but whether or not these mutants can inhibit wt activity on ssDNA has not been shown. The recessive AGS V201D mutant retains 25% of wt TREX1 activity on a 30mer ssDNA, but whether this mutant exhibits a further reduction in enzymatic activity on dsDNA has not been addressed. These questions are important to better understand what substrates may be relevant for TREX1 in the cell. My work has sought to answer several important questions about
TREX1 catalysis and how the mutations within the active site vicinity of TREX1 contribute to autoimmune disease:

- How do mutations in the TREX1 active site vicinity alter the active site architecture to diminish enzyme activity? Do these changes provide further insight into catalysis?

- Does the context of the 3′ terminus affect the ability of a TREX1 mutant to degrade the substrate?

- Are the dominant mutant heterodimers capable of degrading a large ssDNA? Are the dominant homodimer mutants, which are inactive on ssDNA, able to inhibit wt activity on ssDNA?

- What is the mechanism by which dominant TREX1 mutants inhibit wt TREX1 activity?

- Does the recessive AGS V201D mutant have altered substrate specificity?

I have employed an integrated approach of structural biology and biochemistry to answer these questions. I have determined the X-ray crystal structures of TREX1 active site disease mutants (D18N, D200H, D200N, and V201D). This will address structural changes within the TREX1 mutants that may play an important role in the ability of TREX1 to function in the cell. Furthermore, I have evaluated the activities of TREX1 active site mutants on various ssDNA and dsDNA substrates, modeled after putative biological substrates. This will identify possible changes in TREX1 substrate specificity and provide further insight into in vivo substrates.
Chapter 2: Insight into DnaQ family DEDDh catalysis from the crystal structures of TREX1 3’ exonuclease autoimmune disease mutants

Suzanna L. Bailey, Scott Harvey, Fred W. Perrino, Thomas Hollis†

Center for Structural Biology and Department of Biochemistry, Department of Biochemistry, Wake Forest University Health Sciences, Medical Center Boulevard, Winston-Salem, North Carolina 27157

†To whom correspondence should be addressed: E-mail: thollis@wfubmc.edu. Tel: 336-716-7230. Fax: 336-777-3242.

S. Bailey performed the X-ray crystallization experiments, structure refinement, and prepared the manuscript. S. Harvey prepared the mouse TREX1 protein for crystallization. Drs. F. W. Perrino and T. Hollis acted in an advisory and editorial capacity.
Abstract

Active site mutations within the TREX1 3′ exonuclease have been identified in patients with Aicardi-Goutières Syndrome and familial chilblain lupus. Both AGS and FCL are autoimmune diseases that result in circulating antibodies to DNA. TREX1 is a member of the endoplasmic reticulum-associated SET complex and participates in granzyme A-mediated cell death to degrade nicked genomic DNA. The loss of TREX1 DNase activity may result in the accumulation of double-stranded DNA degradation intermediates that trigger autoimmune activation. Homodimer and heterodimer dominant AGS and FCL mutants inhibit wt TREX1 activity on a nicked double-stranded plasmid DNA in vitro. TREX1 is a member of the DnaQ family of exonucleases and has four conserved acidic residues in the active site (D18, E20, D130, and D200) that coordinate two divalent metals for catalysis. The structure of the wt TREX1 protein in complex with substrate revealed that the catalytic histidine adopts two conformations in a single dimer. The X-ray crystal structures of the wt apoprotein, the dominant D200H, D200N and D18N homodimer mutants, as well as the recessive V201D homodimer mutant have been determined. The D200H and D200N mutants no longer coordinate one of the active site metals. The TREX1 D18N and V201D mutants bind both metals in the active site, but with inter-metal distances that are larger than optimal for catalysis. Additionally, the catalytic histidine adopts the same conformation in both protomers of each mutant homodimer, suggesting a loss of mobility in this residue that may hinder catalysis. Together with the previously determined wt substrate and product structures, these data reveal a unique mecha-
nism for DEDDh members of the DnaQ family of exonucleases, where the motion of the catalytic histidine acts as a switch between an active and a resting state.
2.1 Introduction

The TREX1 3′ exonuclease is a member of the DnaQ family of exonucleases and catalyzes the major 3′ exonuclease activity in mammalian cells (91). TREX1 interacts with the SET protein as a member of the ER-associated SET complex (110). The SET complex ranges in size from 270-440 kDa and contains such proteins as the endonuclease NM23H1, the apurinic endonuclease APE1, and the DNA bending protein HMG-2. The SET complex participates in the Gzm A cell death pathway. As a part of this pathway, TREX1 degrades nicked genomic DNA to ensure cell death (110). This TREX1 function may be important to avoid the accumulation of host DNA that could trigger an autoimmune response.

The TREX1 protein forms a homodimer in solution, and the crystal structure reveals that the active sites are positioned at opposite outer edges of the same face of the dimer (94, 106). The TREX1 protein contains the four highly conserved acidic residues (DEDD) within the active site that are characteristic of the DnaQ family of exonucleases. These residues are required for coordinating two divalent metal ions that are necessary for catalysis. Furthermore, a single protomer of the TREX1 protein adopts a fold very similar to that of the subgroup members of the DnaQ family of exonucleases characterized by the presence of a catalytic histidine rather than a tyrosine. The catalytic histidine has a proposed role in deprotonating a water molecule and promoting nucleophilic attack on the scissile phosphate during catalysis (91, 94, 102, 103, 106). The protein has several distinct structural features that likely contribute to its cellular function. Adjacent to each active site is a flexible
loop that is involved in DNA binding (103). A polyproline type II helix is positioned adjacent to the active site and dimer interface in each TREX1 protomer. The PPII helix in TREX1 may mediate contacts with other proteins in the cell, and may be important for its interaction with the SET complex. The TREX1 protein also has an extended C-terminal domain with a predicted transmembrane helix that is involved in the subcellular localization of TREX1 on the cytosolic side of the ER (62, 110).

Trex1 deficient mice develop inflammatory myocarditis, suggesting a role for TREX1 activity in avoiding an autoimmune phenotype (55). Recent genetic studies have linked mutations in the Trex1 gene with a spectrum of autoimmune diseases including Aicardi-Goutières syndrome, familial chilblain lupus, systemic lupus erythematosus, and retinal vasculopathy and cerebral leukodystrophy (56, 59–64, 120). The sites of mutations causing autoimmune diseases in humans are located throughout the Trex1 gene sequence; affecting areas of the protein including the active site, the dimer interface, and the C-terminal domain. Mutations within the TREX1 active site have been linked to both AGS and FCL. AGS is a genetically determined encephalopathy exhibiting recessive inheritance, although isolated cases of dominant AGS have been identified. The symptoms of AGS mimic congenital viral infection without evidence of an infectious agent (76, 77, 79). These patients exhibit a severe phenotype involving the calcification of basal ganglia and white matter in the brain that negatively impacts motor and social development (78, 79, 82). FCL is a dominant form of lupus erythematosus that causes painful bluish-red skin lesions called chilblains that predominantly develop in acral locations such as the hands and feet (59, 61, 124).
AGS and FCL have significant phenotypic and clinical overlap including circulating antinuclear antibodies to dsDNA or ssDNA, suggesting a common pathogenesis.

The exact cellular role of TREX1 remains unclear, but potential functions in the degradation of nicked genomic DNA during Granzyme A cell death (110), elimination of DNA derived from endogenous retroelements as a part of the interferon-stimulatory DNA response (107), and the disposal of aberrant ssDNA replication intermediates (117, 118) have been proposed. Complicating the issue further, dominant TREX1 mutations might enable the formation of three different species of the TREX1 dimer in the cell depending on allele expression. Wt/wt homodimers, mutant/mutant homodimers, or wt/mutant heterodimers are all feasible and might be present simultaneously in a cell. Dominant AGS and FCL mutants are defective in their dsDNA degradation activity and both homo- and heterodimers are able to inhibit the dsDNA degradation activity of the wt enzyme (123), supporting the proposed role for TREX1 in degrading nicked dsDNA during granzyme A cell death. The mechanism for inhibition has been proposed to involve competitive binding of DNA nicks by the catalytically inactive dominant mutants.

In order to better understand the effects of disease causing mutations on the structure and function of the TREX1 enzyme, we have determined the X-ray crystal structures of the wt TREX1 apoprotein as well as the structures of four AGS and FCL mutants. In order to establish a starting point for TREX1 catalysis we determined the TREX1 apoprotein structure, which was previously unknown. This structure may provide further insight into the TREX1 mechanism and the contribution of active site
mutations to autoimmune disease. The FCL D18N, the dominant AGS D200N and D200H, and the recessive AGS V201D mutants are all in or directly adjacent to the TREX1 active site. These structures indicate an important role for flipping of the catalytic histidine during catalysis. Furthermore, these structures reveal unique characteristics of two metal ion catalysis in the context of the DEDDh subgroup of the DnaQ family of exonucleases that may explain substrate trapping in FCL and dominant AGS.

2.2 Experimental Procedures

2.2.1 Protein Expression and Purification

The mouse TREX1 (mTREX1) enzymes used for crystallization contain the catalytic core residues (1-242) of the protein missing the C-terminal amino acids. The TREX1 proteins were expressed in a modified pMAL-C2 vector, where a 6x polyhistidine sequence was added at the maltose binding protein (MBP) N-terminus and the Rhinovirus 3C recognition sequence is located between the MBP and the Trex1 genes. The vector was transformed into Escherichia coli Bl21*(DE3) Rosetta II cells (Novagen) for overexpression. The cells were grown to an A600=0.5 and induced with 1 mM isopropyl-β-D-thiogalactopyranoside and quickly cooled on ice to 16°C. Cells were allowed to grow for 18 hours at 16°C. The MBP-TREX1 fusion was purified by affinity chromatography using amylose resin followed by PreScission Protease treatment to remove the MBP fusion protein. The TREX1 protein was then dialyzed against 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10% glycerol and 1 mM EDTA and purified to homogeneity using phosphocellulose chromatography. Purified TREX1
proteins were flash frozen and stored at -80°C.

2.2.2 Protein Crystallization and X-ray Data Collection

The mTREX1 mutant proteins and the wt TREX1 apoprotein were crystallized using the sitting drop vapor diffusion technique. All TREX1 proteins were dialyzed into 20 mM MES (pH 6.5), 50 mM NaCl. Substrate complex was formed by incubating the protein with ssDNA in a molar ratio of 1:2 and 5 mM calcium chloride (5 mM magnesium chloride in the case of the D200N mutant). TCEP-HCl (pH 8.0) was added to a final concentration of 1 mM to each protein solution prior to placing in the crystallization tray. 2 µl protein complex at 5 mg/ml mTREX1 was mixed with an equal volume of reservoir solution and placed on a bridge above 500 µl of the reservoir solution. Optimized crystals of the wt TREX1 apoprotein grew in 0.25 M tri-sodium citrate dihydrate and 20% PEG 3350. Optimized crystals of the D18N complex were obtained by microseeding into 0.15 M MES (pH 6.5), 19% PEG 4000 and 10% ethylene glycol and grown at 15°C. Crystals of the D200N mutant complex were obtained by microseeding into 0.1 M MES (pH 5.5), 0.075 M NaCl, 12% PEG 3350 and 5% 1,4-butanediol and grown at 15°C. The D200H complex was obtained by microseeding into 0.1 M MES (pH 6.0), 16% PEG 4000, 2% 1,4-butanediol and grown at 30°C. Crystals of the V201D homodimer in complex with ssDNA were grown at 30°C in 16% PEG 3350, 0.1 M NaI and 5% 1,4-butanediol. All crystals grew within one week. Prior to data collection all crystals were dipped into reservoir solution containing 20% 1,4-butanediol or (for D18N) 20% glycerol in preparation for cryo-cooling. Crystals were mounted on a nylon loop and flash cooled to 100 K in a stream
Table 2.1: Crystallographic Data and Refinement Statistics

<table>
<thead>
<tr>
<th>TREX1</th>
<th>Apoprotein</th>
<th>D18N</th>
<th>D200H</th>
<th>D200N</th>
<th>V201D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2_1</td>
<td>P2_1</td>
<td>P2_1</td>
<td>P2_1</td>
<td>P2_1</td>
</tr>
<tr>
<td>Unit cell</td>
<td>a=64.2, b=85.7, c=100.3</td>
<td>a=65.2, b=58.8, c=66.5</td>
<td>a=65.2, b=57.9, c=68.3</td>
<td>a=64.9, b=56.6, c=68.1</td>
<td>a=70.5, b=86, c=92.8</td>
</tr>
<tr>
<td>α,β,γ</td>
<td>=90°</td>
<td>=90°</td>
<td>=90°</td>
<td>=90°</td>
<td>=90°</td>
</tr>
<tr>
<td>Molecules/asymmetric unit</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Wavelength, Å</td>
<td>1.54</td>
<td>1.54</td>
<td>1.54</td>
<td>1.54</td>
<td>1.54</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>29-1.95</td>
<td>43-2.55</td>
<td>29-2.35</td>
<td>43-2.2</td>
<td>47-1.75</td>
</tr>
<tr>
<td>Completeness, %</td>
<td>99.7 (97.1)</td>
<td>96.6 (95.1)</td>
<td>99.5 (99.5)</td>
<td>99.7 (99.6)</td>
<td>99 (91.2)</td>
</tr>
<tr>
<td>R_merge, %</td>
<td>5.7 (27.6)</td>
<td>10 (23)</td>
<td>8.4 (24.8)</td>
<td>6.4 (28.1)</td>
<td>6.8 (24.0)</td>
</tr>
<tr>
<td>Mean I/σ</td>
<td>16 (4.1)</td>
<td>10.9 (4.5)</td>
<td>8.8 (3.7)</td>
<td>17.4 (4.4)</td>
<td>11.9 (3.4)</td>
</tr>
<tr>
<td>Average redundancy</td>
<td>6.7</td>
<td>6.9</td>
<td>3.5</td>
<td>7</td>
<td>5.7</td>
</tr>
<tr>
<td>R-factor, %</td>
<td>20.4</td>
<td>23.1</td>
<td>24.9</td>
<td>25.1</td>
<td>19.5</td>
</tr>
<tr>
<td>R_free, %</td>
<td>23.7</td>
<td>28.4</td>
<td>29.5</td>
<td>32.5</td>
<td>22.6</td>
</tr>
<tr>
<td>Average Wilson B-factor, Å²</td>
<td>24.7</td>
<td>26.9</td>
<td>33.8</td>
<td>36.5</td>
<td>19</td>
</tr>
<tr>
<td>Root mean square deviation bond lengths, Å</td>
<td>0.007</td>
<td>0.008</td>
<td>0.016</td>
<td>0.008</td>
<td>0.010</td>
</tr>
<tr>
<td>Root mean square deviation bond angles, °</td>
<td>1.069</td>
<td>1.291</td>
<td>1.842</td>
<td>1.227</td>
<td>1.279</td>
</tr>
</tbody>
</table>

*a*\(R_{merge} = \sum |I - \langle I \rangle|/ \sum I\), where I is the observed intensity and \(\langle I \rangle\) is the average intensity.

*b*\(R\)-factor = \(\sum |F_o| - |F_c|/ \sum |F_o|\), \(R_{free}\) is the same as \(R\), but calculated with 5% of the reflections that were never used in crystallographic refinement.

of N\(_2\) (l). The D18N, D200N, and D200H mutants in complex with ssDNA belong to the P2\(_1\) spacegroup. The V201D homodimer ssDNA complex and the wt apoprotein belong to the P2\(_1\)2\(_1\)2\(_1\) spacegroup (Table 2.1).

### 2.2.3 Phasing and Refinement

The X-ray data were collected using CuK\(_\alpha\) radiation on a MicroMax 007 generator and a Saturn 92 CCD detector (Rigaku). Intensity data were processed using the programs d*TREK or HKL2000 (125, 126). Phases for the data were obtained by maximum likelihood molecular replacement using the program PHASER (127) and the mTREX1 wt dimer (PDB:2OA8) as the search model. The mTREX1 apoprotein and homodimer mutant models were built in the program COOT (128) and the structures refined using the programs CNS (129), Refmac5 (130), and Phenix.refine (131). Translation/libration/screw (TLS) refinement was utilized to independently
define subgroups within the apoprotein structure and to further refine their directions of movement as individual rigid bodies (132, 133). The inspection of clashes and stereochemical parameters was carried out using the program MolProbity (134). All structure figures were generated in the program Pymol (135).

2.3 Results and Discussion

The TREX1 FCL D18N and the AGS D200H and V201D mutants were co-crystallized with ssDNA and Ca\(^{2+}\). Calcium enables DNA binding without supporting catalysis. The D200N mutant was co-crystallized with ssDNA and Mg\(^{2+}\). Magnesium was utilized as the cofactor for D200N crystallization because diffraction quality crystals of this mutant were not obtained with calcium and this mutant enzyme is catalytically inactive. Data were collected to 1.95 Å, 2.55 Å, 2.35 Å, 2.2 Å, and 1.75 Å resolution for the wt apoprotein, D18N, D200H, D200N, and V201D structures, respectively.

All structures were determined by molecular replacement (MR), using the wt TREX1 dimer (PDB: 2OA8) as the search model. After MR and composite omit map analysis to remove model bias, good electron density was observed for most residues in the TREX1 dimer in each mutant structure and in the wt apoprotein structure. In addition, electron density for the ssDNA was observed in the active site of all of the mutant structures indicating that the observed decrease in activity in these mutants is not due to an inability to bind substrate. Final R/R\(_{free}\) at the conclusion of refinement for the X-ray crystallographic models of the homodimeric TREX1 apoprotein and mutants are 20.4%/23.7% for the wt apoprotein, 23.1%/28.4% for D18N, 24.9%/29.5% for D200H, 25.1%/32.5% for D200N, and 19.5%/22.6% for V201D. Further refinement
Figure 2.1: Superposition of the active sites from the wt TREX1 substrate complex reveals an active and a resting conformation within the dimer. The active conformation is shown in green. The resting conformation, observed in the other protomer of the homodimer, is shown in blue.

statistics are listed in Table 2.1.

2.3.1 Active site rearrangements in the TREX1 apoprotein and mutant substrate complex structures

In the TREX1 active site, both D18 and D200 coordinate metal A and D18 additionally coordinates metal B (Figure 2.1). V201 is on the back side of the α-helix containing the adjacent metal-coordinating D200 residue. In the two metal ion mechanism for catalysis, metal ion B is important for DNA binding and stabilization of
the pentacovalent transition state and metal ion A is required for directing the inline attack of the scissile phosphate by an activated water molecule (96, 97) (Figure 2.1). The catalytic histidine (H195) has a putative role in the deprotonation of a water molecule coordinated by metal A to generate the nucleophile (100, 136). Mutagenesis of the catalytic histidine in the TREX2 exonuclease and RNase T abolishes exonuclease activity (95, 103, 104). Quantum mechanical/molecular mechanical studies of the catalytic mechanism of the epsilon subunit of *E. coli* DNA polymerase III indicate that a proton is transferred from the nucleophilic water to the catalytic histidine, further supporting a role for the histidine in the catalytic mechanism (102). H195 is observed in two different conformations within a single homodimer in the X-ray crystal structure of the wt TREX1 substrate complex (Figure 2.1). A similar dual histidine conformation is observed in the X-ray crystal structure of the homodimeric, bacterial RNase T in complex with metal but not substrate (137). In contrast, a single conformation is observed for the equivalent histidine in the homodimeric oligoribonuclease (ORN) structures (138–140). Alternating activity between protomers within the ORN dimer may not be beneficial for oligoribonuclease, which degrades only very short, 2-5 nucleotide substrates, and may explain the single catalytic histidine conformation observed in its crystal structures. The presence of two conformations of the catalytic histidine in both the TREX1 and RNase T structures suggests a role for communication across the dimer interface, and that this may coordinate alternating catalytic activity between protomers of the dimer.
Figure 2.2: The structure of wt TREX1. (A) Representation of structural features on the primary sequence. All residues within the gray box (1-234) are located within the catalytic domain of TREX1. The approximate positions of the D18, D200, and V201 residues are indicated. Exo: exonuclease motifs (1, 2, and 3); DBL: DNA binding loop (B) Wt TREX1 apoprotein structure with one monomer shown in green and the other in blue. The boxed area has been enlarged to show the active site residues, that are involved in metal coordination. The sidechain of the catalytic histidine is disordered in this structure. (C) Structural superposition of the wt TREX1 substrate complex (blue) and the apoprotein structure (green). (RMSD 0.66 Å)
TREX1 apoprotein structure

Mobility in the loop containing the catalytic histidine (H195) in the TREX1 apoprotein structure suggests a role for metal and/or DNA binding in structuring this loop for catalysis. The wt apoprotein structure maintains the same overall fold as the wt structure in complex with DNA (Figure 2.2), indicating that there are no major structural rearrangements upon DNA binding (RMSD 0.66 Å). The TREX1 active site retains a similar morphology in the apoprotein structure to that observed in the wt substrate complex with the exception of the flexible loop containing histidine 195 (94). There is electron density for the backbone of this loop in the TREX1 apoprotein structure, but it lacks density for the sidechains of residues 194-195. Histidine 195 has a putative role in the deprotonation of a water molecule coordinated to one of the active site metals to generate the nucleophile for attack on the scissile phosphate (100, 136). The ordering of this residue upon binding of metal cofactor or substrate may be important for catalysis. The loop extending behind the PPII helix also has higher flexibility in the apoprotein structure, which may preclude the formation of protein-protein contacts mediated by the PPII helix motif in the absence of bound metal and/or DNA.

TREX1 autoimmune disease mutant structures

Each of the TREX1 mutant proteins adopts the same overall α/β fold as the wt TREX1 enzyme (RMSD 0.37-0.62 Å). The structures of the mutant TREX1 proteins fit into two distinct groups based on the number of metals coordinated within
the active site and the position of H195 in both active sites. In both structures of the D200H and D200N AGS mutants, H195 is flipped out of the active site in each monomer and there is no electron density for metal ion A (Figure 2.3). In the structures of the D18N and V201D TREX1 mutants, H195 is flipped into the active site in both protomers and electron density is present for both metal ions (Figure 2.3). The fact that a single H195 conformation is observed in each of the TREX1 mutant structures suggests a contribution of the decrease in histidine mobility to the reduction in TREX1 exonuclease activity that has been detected for these mutants (61, 94, 123).

In the D200 mutant structures, H195 is positioned out of the active site and only metal ion B is bound (Figure 2.3). The flipped out conformation of H195 is stabilized by hydrogen bonds from D193 to the backbone amide of H195 and the sidechain hydroxyl of T196, as observed in the wt structure (Figure 2.4). The mutation of D200 to histidine or asparagine may preclude the formation of the hydrogen bond to the backbone amide of H195, stabilizing the flipped in conformation of H195. This may explain the observed position of H195 out of the active site. It is also possible that the binding of metal A helps stabilize the flip of H195 into the active site and that the loss of this metal in the D200N and D200H structures explains the observed flipped out H195. The loss of both the motion of H195 to activate the nucleophile and metal A to coordinate the nucleophile and direct nucleophilic attack clearly explains the loss of activity on both ssDNA and dsDNA that is observed for the D200 mutants.

In both protomers of the D18N and V201D structures metal ions A and B are bound and H195 is flipped into the active site in what we have termed the inward
Figure 2.3: Active site morphology of TREX1 mutants. Bound metal cofactor is shown as yellow spheres. The mutated residues in the AGS and FCL mutant active sites are highlighted in blue. (a) TREX1 dominant AGS D200H homodimer (b) TREX1 dominant AGS D200N homodimer (c) TREX1 FCL D18N homodimer (d) TREX1 recessive AGS V201D homodimer
conformation (Figure 2.3). In the D18N structure, the calcium ions in the two protomers are separated by 4.18 and 4.45 Å (Figure 2.3). This distance is consistent with the 4.5 Å inter-magnesium distance that Nowotny et al. (98) observed for their D132N-Mg$^{2+}$ RNase H structure (D132 is a bridging aspartate in RNase H), which is a catalytically inactive enzyme – similar to the TREX1 D18N homodimer. The metal ion spacing is an important factor in two metal ion catalysis, and it has been proposed that metal A moves toward metal ion B during catalysis to bring the nucleophile close enough to the scissile phosphate to promote catalysis (101). In both the substrate

Figure 2.4: Stabilization of the inward and flipped out conformations of H195. (A) When H195 is out of the active site, the loop is stabilized by two hydrogen bonds from D193. (B) The inward conformation of H195 is stabilized by a single hydrogen bond from D200 to the backbone amide of H195.
and product complex structures of the wt TREX1 protein, electron density for two metal ions is present in the enzyme active site and the metal ion spacing is similar to the inter-metal distance of 3.5 Å that has been proposed as the metal spacing for an active nuclease (98, 141, 142). In the wt TREX1 substrate complex (PDB: 2OA8), the distance between metal ions is 3.6 Å in one monomer and 3.1 Å in the other monomer (94). When the metal ions are separated by 4 Å or more, as observed in the D18N structure, metal A may be unable to help the nucleophile bridge the distance to the phosphorous. The metal ion spacing is between 3.7 and 3.8 Å in both V201D protomers (Figure 2.3d), but this mutant protein still enables degradation of a ssDNA substrate at 25% of the rate of the wt enzyme (94). As observed in both wt structures, the inward conformation of H195 is stabilized by a hydrogen bond from D200 to the backbone amide of H195 (Figure 2.4). If metal A plays a role in stabilizing the H195 in the inward conformation, then it is also possible that an increase in the spacing between the metal ions further stabilizes this conformation. This could explain why H195 is in the inward conformation in both the D18 and V201 mutant structures. The increased metal distance and the loss of H195 mobility likely explain the reduced activity of these mutant enzymes.

**V201D and recessive AGS**

Although the V201D structure supports the observed decrease in exonucleolytic activity, this enzyme still retains 25% of the wt exonuclease activity (94). Thus, the contribution of this recessive mutant to AGS remains unclear. Cell extracts derived
from a Turkish patient with the V201D mutation showed no detectable TREX1 activity (56). These results in combination with the fact that the recombinant V201D mutant retains activity in vitro suggests that there may be another mechanism for shutting down TREX1 activity to cause recessive disease. It is possible that the V201D protein is readily degraded after it is translated. Another possibility is that the TREX1 exonuclease activity is turned off by post-translational modification or by modified or novel protein-protein interactions. Based on the crystal structure, there are no apparent structural changes within the N-terminal 242 residues that might target the protein for degradation or alter protein-protein interactions. Perhaps there are structural changes within the C-terminal linker region between the TREX1 catalytic domain and the putative transmembrane helix (Figure 2.2A) that are not present in the crystal structure that affect post-translational modification to predicted serine phosphorylation sites (62). Phosphorylation of these serines under conditions of oxidative stress has been proposed as the trigger for TREX1 translocation to the nucleus (62). The inability to phosphorylate at these sites could cause TREX1 to be sequestered at the ER and result in an effective loss of TREX1 activity. Further analysis of the TREX1 sequence reveals three predicted sites for ubiquitination (143), two within the TREX1 linker region and one within the DNA binding loop adjacent to the active site. Ubiquitination at any of these sites could target the catalytically active V201D mutant to the proteasome for degradation or simply turn off enzyme activity, both of which would deprive the cell of TREX1 activity. Furthermore, ubiquitination can be a signaling event for translocation to the nucleus. If ubiquitination signals the
translocation of the TREX1 protein to the nucleus, the loss of this capability could lead to DNA accumulation and an aberrant immune response.

### 2.3.2 Two Metal Ion-Mediated Nucleic Acid Hydrolysis and the TREX1 Mechanism

Many enzymes utilize a two metal ion mechanism to catalyze the hydrolysis of a phosphodiester bond and there is significant structural diversity among this group of enzymes. A large number of these nucleases require only the two metal ions to generate the nucleophile and promote attack on the scissile phosphate (96, 97, 101). This mechanism includes the RNase H family of enzymes that coordinate two divalent cations via a DEDD motif within the active site that alone promote and stabilize the removal of a nucleotide from an RNA:DNA hybrid (98, 101, 144). The current interpretation of this two metal ion mediated catalytic mechanism is that metal ion A reduces the pKₐ of a water molecule located directly above it and assists nucleophilic attack via the formation of a hydroxide ion. Metal ion B stabilizes the pentacovalent intermediate and facilitates product leaving (97). Members of the DnaQ family of exonucleases require a protein residue (His or Tyr) in addition to the two metals during catalysis. As in the RNase H family, the DnaQ family of exonucleases has a highly conserved DEDD motif for the binding of the two divalent cations within the enzyme active site (94, 96, 99, 100, 102, 137, 145, 146). A close examination of the active site architecture within the DnaQ exonuclease family reveals two distinct structural subclasses, divided into the DEDDy and DEDDh subgroups that have significant catalytic implications. The DEDDy-containing nucleases are most commonly found
within a subdomain of a multifunctional protein, such as the exonuclease domain of WRN or bacterial DNA polymerase I. The DEDDh nucleases are structurally separate proteins that contain only exonucleolytic activity and may associate with other proteins to modulate that activity. DEDDh exonucleases include enzymes such as the ε subunit of bacterial DNA polymerase III, exonuclease I (exo I), oligoribonuclease, RNase T, TREX1, and TREX2.

The proposal that the DEDDy and DEDDh subgroups utilize different catalytic mechanisms is substantiated by in vitro biochemical evidence, as well as structural data. Tyrosine has a high pK\textsubscript{a} and in vitro biochemical experiments show that when the conserved tyrosine residue is mutated to phenylalanine exonuclease activity is reduced, but not abolished. These data support the idea that the conserved tyrosine is not involved in generation of the nucleophilic hydroxyl ion, but rather as a hydrogen bond donor that helps direct nucleophilic attack (99). This would suggest a role for metal ion A in generation of the nucleophile as has been proposed for two metal ion enzymes such as those in the RNase H family.

The presence of histidine in the DEDDh family, with a lower pK\textsubscript{a} than tyrosine, hints at the possibility of a different mechanism. In fact, mutation of the conserved histidine within the DEDDh exonucleases has dramatic consequences for catalysis. For example, mutation of His188 to alanine in TREX2 shows a 3 x 10\textsuperscript{4}-fold reduction in relative exonuclease activity (95, 103). Quantum mechanic/molecular mechanic simulations on the ε subunit further support the role of the histidine in generation of the nucleophile for catalysis (102). These data are consistent with a mechanism where
the conserved histidine directly participates in the generation of the nucleophile.

The differences between catalysis by the DEDDy and DEDDh exonucleases further diverge based on structural observations. A superposition of the structure of the TREX1 monomer with those of the WRN exonuclease (pdbid 2FBT), ε subunit (pdbid 1J53), exo I (pdbid 1FXX), orn (pdbid 1YTA) and the exonuclease domains of DNA polymerase I (Klenow fragment, pdbid 1KSP), RB69 gp43 DNA polymerase (pdbid 2P5O), and phage T7 DNA polymerase (pdbid 1T7P) show a general conservation of the DnaQ family fold (99, 100, 136, 139, 147–149). In all of the analyzed structures the acidic metal coordinating residues superimpose onto one another nearly perfectly, as do the substrate or product nucleotides when present in any of the structures. However, the structures differ significantly in the region containing the conserved tyrosine or histidine residue that divides this family into the two subgroups. The DEDDy structures show the conserved tyrosine to be within a long helix in the back of the active site. This helix is preceded by another short helix and a long loop that stretches across and over the active site in close proximity to the tyrosine (Figure 2.5). Structures of several DEDDy 3′ exonucleases have been determined, both in the presence and absence of DNA, that show the tyrosine is capable of adopting conformations pointing both towards and away from the active site (99, 147–149). The position of the active site loop relative to the tyrosine indicates that it poses a steric block to the tyrosine switching between the two conformations (Figure 2.5), and suggests the loop may play a role in coordinating movement of this residue. The requirement for the rearrangement of the active site loop to allow the two conforma-
Figure 2.5: Positioning of the active site loop in DEDDh versus DEDDy exonucleases. In DEDDh exonucleases the catalytic histidine is positioned in a flexible loop with room around the residue for it to easily alternate between an inward and flipped out conformation. In DEDDy exonucleases the active site loop acts as a steric gate requiring a rearrangement of the loop for the tyrosine to move in and out of the active site.
tions of the conserved tyrosine indicates that alternation between these two tyrosine conformations would not happen readily during catalysis. This rearrangement most likely would occur upon substrate binding in the active site as previously proposed (99). In contrast, structures of the DEDDh exonucleases show these enzymes to have a conserved histidine in the loop preceding a shorter helix (Figure 2.5). However, the loop originates above and away from the active site and is relatively removed from the nucleotide-binding pocket. Therefore, the loop does not hinder motion of the histidine in and out of the active site. This histidine-assisted catalysis is a modification of the metal-ion dependent mechanism of the DEDDy exonucleases that may allow for more efficient exonucleolytic activity.

A proposed mechanism for DEDDh exonucleases

Together, the wt and mutant structure data provide us with structures for most steps in the TREX1 catalytic process. This enables us to propose a catalytic mechanism for the TREX1 exonuclease that is specific to DEDDh exonucleases. The cumulative structural data suggest a mechanism where the two protomers switch back and forth between an active and resting state. This is demonstrated by the dual conformation of H195 in the wt-substrate complex and the shift in the metal ion spacing from 3.1 Å when H195 is out of the active site (resting) to 3.6 Å when H195 moves into the active site (active) (Figure 2.1). This is consistent with the active and resting states that were previously proposed for two metal ion catalysis based solely on the inter-metal ion distance (101). An activated water molecule is thought
to be the nucleophile that attacks the scissile phosphate during catalysis. The fact that a single water molecule is coordinated by metal A when H195 is out of the active site, but not when H195 is flipped into the active site in the structure of the TREX1 substrate complex further supports this model (94).

Due to the possibility of cooperation in DNA catalysis across the dimer interface, it is important to address the significance of homodimerization on the catalytic mechanism of TREX1. We propose that the TREX1 3' exonuclease and other similar homodimeric nucleases such as TREX2 catalyze the nucleotide excision reaction via a mechanism where the catalytic histidine (H195 in TREX1) acts as a switch between the active and resting states of the enzyme and that the protomers alternate between these two states as they degrade a substrate. A similar switching between a flipped in and flipped out conformation for the conserved tyrosine during catalysis in the DEDDy exonucleases is unlikely due to the steric block of the active site loop. We propose that for TREX1 the reaction begins with the binding of metal and DNA in the enzyme active site (Figure 2.6a-b), which results in the ordering of the H195 loop. In one protomer, H195 flips into the active site to abstract a proton from the lone H$_2$O molecule coordinating metal A, displacing it towards the scissile phosphate (Figure 2.6c). The motion of the activated water molecule (the sixth ligand of metal A) towards the substrate may result in the shift in metal A diagonally away from metal B and towards the substrate, directing the nucleophilic attack on the scissile phosphate by the hydroxide anion. This motion of metal A is supported by the X-ray crystal structure of the wt TREX1-substrate complex, where metal A shifts from being 3.1
Figure 2.6: Proposed mechanism of the TREX1 homodimer. The scissile phosphate is highlighted in red. The divalent cation cofactors are shown as green spheres and water molecules involved in coordination of the metals are shown as small, dark blue spheres.

Å from metal B when H195 is out of the active site, to 3.6 Å from metal B and closer to the scissile phosphate when H195 is in the inward conformation (PDBID: 2OA8). This motion of metal A is different from that proposed for metal A towards metal B to promote catalysis in RNase H, but may be explained by the fact that metal A is penta-coordinated in TREX1 and hexa-coordinated in RNase H (98, 101). Metal B participates in the stabilization of the pentacovalent transition state (Figure 2.6d). Once the scissile bond is broken, H195 flips back out of the active site, releasing the product and allowing the alignment of the new 3′-terminal nucleotide for excision.
Simultaneously, H195 in the opposite monomer flips into the active site initiating a new reaction. TREX1 is a non-processive enzyme; therefore, the substrate and/or metals will be lost from the active site at the end of most cycles of catalysis (Figure 2.6e-a). More metal and another DNA 3’ end must bind in the active site for catalysis to continue. Other DEDDh subgroup exonucleases likely follow a similar mechanism involving an active and resting state that are defined by both the inter-metal distance and the position of the catalytic histidine.

### 2.3.3 Non-processivity and Substrate Trapping in TREX1

TREX1 proteins containing the dominant active site mutations inhibit the activity of wt TREX1 on a dsDNA substrate, suggesting that the mutant TREX1 protein binds, but does not release the substrate. The mutation of either D18 to asparagine or D200 to asparagine or histidine in TREX1 renders the homodimer mutant enzyme essentially inactive on either ssDNA or dsDNA (61, 94, 123, chapter 3). The D18 and D200 heterodimers, containing one wild-type and one mutant protomer in the dimer, are about 50% active on ssDNA and have no activity on a nicked dsDNA plasmid substrate. The mutation of either of these residues causes dominant disease in humans and in vitro both the homo- and heterodimer D18 and D200 mutants inhibit the activity of wt TREX1 on a dsDNA substrate (123, chapter 3).

The R114H mutation at the TREX1 dimer interface causes AGS as a homozygous mutation and SLE as a heterozygous mutation (56, 62). However, neither the R114H/R114H homodimer nor the wt/R114H heterodimer can inhibit wt TREX1 activity on the nicked dsDNA plasmid substrate (123). This suggests that the domi-
nant TREX1 active site mutants are binding the substrate and not readily releasing it, and that this stabilized interaction is unique to the active site mutants. D18 and D200 participate in metal A coordination. Metal A is absent from the D200 homodimer mutant structures and is coordinated with a significantly larger inter-metal ion distance in the D18 homodimer mutant structure than is observed in the wt TREX1 protein, suggesting that a significant contribution to the dominance of these mutants may be attributed to the altered binding or absence of metal A. A stabilized interaction between the DNA substrate and the flexible DNA binding loop adjacent to the TREX1 active site may also contribute to the ability of these active site mutants to inhibit wt TREX1 activity.

We propose that the release of metal A from the enzyme active site at the end of the catalytic cycle triggers the release of substrate, resulting in the non-processive nature of TREX1. Because the D200 mutants do not have metal A to lose from the active site, they may release bound oligonucleotide much more slowly. The D18 mutant binds metal A, but with a 0.6-0.8 Å greater inter-metal distance than is observed in the wt structure. There is likely a delicate balance between the stability of the coordination of the metals in the active site and the charge repulsion between the two active site metals. If that balance gets tipped, metal A may get bumped out of the active site. The larger inter-metal distance in the D18 structure likely stabilizes metal A in the active site, subsequently stabilizing the interaction between TREX1 and the substrate. The D18 mutant may still lose metal A and release the substrate explaining the FCL phenotype. The D200 mutants result in the severe AGS phenotype and do
not coordinate metal A in the active site. This may explain the proposed substrate trapping by these dominant autoimmune disease mutants that results in the observed inhibition of wt TREX1 activity.

2.3.4 Why Two Catalytic Mechanisms in the DnaQ Family of Exonucleases?

The structural and biochemical information gleaned from the DnaQ enzymes points to a distinction between the DEDDy and DEDDh enzymes. The obvious question is why are there two different mechanisms to achieve the same result? One possibility is that different DNA metabolic functions require different levels of exonuclease activity. DEDDy exonucleases are part of larger multifunctional enzymes that have other DNA metabolic functions such as polymerase or helicase activity. Theoretical considerations show DNA polymerases have a cost of exonuclease activity when balanced against polymerization (150, 151). The cost is the proportion of correctly matched to incorrect nucleotides removed and is an important limiting factor in efficiency of editing. Intuitively, multifunctional enzymes that pay a cost for 3′ exonuclease activity must have evolved to a threshold of activity that provides the most effective balance between removal of nucleotides and their other enzymatic functions. Having the most efficient or highest exonuclease activity could tip the balance of function towards nucleotide removal and impose a prohibitive cost to these enzymes. DEDDh exonucleases, on the other hand, can afford to have more efficient activity either because they are not balancing it against another DNA metabolic function or higher activity can be modulated through protein-protein interactions since they are usually
structurally separate enzymes.

2.4 Acknowledgements

This work was supported by an American Heart Association predoctoral fellowship to SLB, American Cancer Society Grant RSG-04-187-01-GMC (TH) and National Institutes of Health Grant R01-GM069962 (FWP).
Chapter 3:  Dominant TREX1 Autoimmune Disease Mutants Inhibit Wt TREX1 Exonuclease Activity on Double-stranded DNA by Competitively Binding and Protecting DNA 3' Termini

Suzanna L. Bailey, Scott Harvey, Fred W. Perrino, Thomas Hollis†

Center for Structural Biology and Department of Biochemistry, Department of Biochemistry, Wake Forest University Health Sciences, Medical Center Boulevard, Winston-Salem, North Carolina 27157

†To whom correspondence should be addressed: E-mail: thollis@wfubmc.edu. Tel: 336-716-7230. Fax: 336-777-3242.

S. Bailey performed the experiments and prepared the manuscript. S. Harvey prepared the human TREX1 protein for the activity assays. Drs. F. W. Perrino and T. Hollis acted in an advisory and editorial capacity.
Abstract

The TREX1 enzyme catalyzes the major 3’ exonuclease activity detected in mammalian cell extracts and is the underlying cause of a spectrum of autoimmune diseases. Mutations throughout the Trex1 gene sequence have been identified in patients with Aicardi-Goutières syndrome, familial chilblain lupus, systemic lupus erythematosus, and retinal vasculopathy and cerebral leukodystrophy. TREX1 is a dimeric exonuclease with proposed cellular roles in degradation of nicked genomic DNA during the granzyme A cell death pathway, disposal of ssDNA or dsDNA derived from endogenous retroviral elements, and the elimination of aberrant ssDNA replication intermediates to avoid chronic checkpoint activation. Dominant autoimmune disease mutants may result in the formation of wt/wt homodimer, mutant/mutant homodimer, or wt/mutant heterodimer TREX1 in the cell. The activities of the recessive AGS V201D, dominant AGS D200H and D200N, as well as the FCL D18N mutant were determined on ssDNA and dsDNA that model putative biological substrates. The V201D mutant, which causes recessive AGS, was active on all substrates tested. Dominant TREX1 homodimer and heterodimer active site mutants inhibit wt TREX1 activity on dsDNA substrates modeling putative dsDNA cellular substrates by binding and protecting the 3′ termini. However, the TREX1 mutants are unable to inhibit wt activity on ssDNA substrates, suggesting a primary role for TREX1 in dsDNA degradation in the cell. A recessive AGS mutant retains activity on ssDNA and dsDNA substrates that mimic putative in vivo TREX1 substrates, suggesting a unique mechanism for recessive AGS.
3.1 Introduction

Mammalian cells have mechanisms for detecting components of invading bacteria or viruses as part of the innate immune response. However, there is limited structural basis for the discrimination between self and nonself DNA and RNA, making it crucial to maintain low cytosolic levels of self nucleic acids. Certain nucleases may play a major role in avoiding autoimmune activation based on their genetic linkage to a spectrum of autoimmune diseases. DNase I, DNase II, and RNase H2 endonuclease deficiencies have been implicated in the induction of an autoimmune response in mice and humans (9, 53, 57). The TREX1 3′ exonuclease has been implicated in multiple autoimmune diseases in human patients including Aicardi-Goutières syndrome, monogenic familial chilblain lupus, systemic lupus erythematosus, and retinal vasculopathy and cerebral leukodystrophy. AGS is a recessive encephalopathy that causes calcification of the basal ganglia and white matter in the brain (56, 76, 78, 79, 82). This causes severe retardation or regression in motor and social development without further regression after an initial acute stage of the disease. FCL exhibits autosomal dominant inheritance and causes painful bluish-red skin lesions that develop in acral locations (59, 61, 124). A common feature of both AGS and FCL are autoantibodies to dsDNA, which suggests a common disease etiology.

The TREX1 enzyme functions as a dimer and catalyzes the major 3′ exonuclease activity detected in mammalian cells. The X-ray crystal structure of TREX1 reveals the architectural arrangement of the dimer and the position of each active site at opposite outer edges of the same face of the protein (94, 106). While TREX1 is
active on both single- and dsDNA substrates, only four nucleotides from a single strand fit into the enzyme active site. Within the TREX1 active site, four highly conserved acidic residues coordinate two divalent cations that are important for DNA binding and catalysis. A conserved hisitidine residue likely has a role in the generation of the nucleophile and may also be involved in product release (91, 94, 102, 103, 106). The C-terminal 70 amino acids of TREX1 contain a putative transmembrane helix important for the localization of TREX1 at the endoplasmic reticulum (62). Other structural elements within the TREX1 active site vicinity may contribute to TREX1 function. A mobile loop adjacent to the active site probably plays a role in substrate binding, but how this loop interacts with DNA is unknown. Additionally, arginine 128 has a putative role in melting a dsDNA substrate for catalysis (94).

Several cellular roles have been proposed for TREX1, each of which suggest different DNA structures as the primary substrate for the enzyme. TREX1 is a member of the ER-associated SET complex (110). This complex is composed of other proteins including the nucleosome assembly protein, SET; the endonucleases APE1 and NM23H1; and the DNA bending protein, HMG-2 (110, 111, 115). During GzmA cell death, the SET complex translocates to the nucleus and NM23H1 generates single-strand nicks in the DNA that TREX1 binds and degrades (110–112). Roles for TREX1 have additionally been proposed in the degradation of DNA derived from endogenous retroelements in the cytosol and in the disposal of aberrant replication intermediates to promote cell cycle progression. How each of these functions for TREX1 contribute to TREX1-mediated autoimmune disease remains unclear.
Human autoimmune disease mutations have been identified throughout the TREX1 sequence including the active site, dimer interface, and C-terminal tail. Multiple AGS and FCL mutations are positioned within the TREX1 active site vicinity. The *de novo* heterozygous mutation of D200, a metal coordinating residue in the TREX1 active site, to either a histidine or an asparagine causes dominant AGS. The mutation of V201 to aspartate, adjacent to the metal-coordinating D200, causes AGS as a recessive mutation. The heterozygous mutation of the metal-coordinating D18 residue to asparagine causes FCL.

The mechanism of recessive AGS remains unclear and may follow different pathways depending upon the mutation. The recessive mutation of the amino acid R114 to histidine at the TREX1 dimer interface causes AGS. This mutant has a 34-fold reduction in activity on a 30mer ssDNA and a 300-fold decrease in activity on a nicked double-stranded plasmid substrate (123). The recessive V201D TREX1 mutant, which also causes AGS, exhibits a consistent, approximately 4-fold reduction in activity on all substrates tested. However, an assay for 3′ exonuclease activity in fibroblast cell extracts derived from an AGS patient with the recessive V201D mutation showed no detectable catalytic activity (56). Thus, it remains unclear whether the 75% reduction in catalytic activity observed for the V201D mutant is sufficient to cause disease or if there are other means of shutting down its cellular function. Altered protein-protein interactions or changes in the pattern of post-translational modification of the TREX1 V201D mutant could play a role in the loss of TREX1 function *in vivo*. 
In FCL and dominant AGS, allele expression probably determines the cellular form of the TREX1 dimer. Wt/wt or mutant/mutant homodimer, as well as wt/mutant heterodimer TREX1 might all be expressed in these patients’ cells. The mechanism of dominant TREX1-mediated disease remains a significant question. Specifically, the effect of heterodimers on TREX1 activity is unknown. There are several types of potential TREX1 substrates in the cell including 60-65 nucleotide ssDNA, 100-10,000 nucleotide ssDNA, nicked dsDNA, and blunt-ended dsDNA substrates. In order to understand the effect of heterodimers on activity and gain a better picture of how they might function on potential substrates we evaluated the activities of the dominant TREX1 mutants on substrates that model proposed TREX1 substrates. The heterodimer mutants inhibit wt TREX1 activity on dsDNA regardless of the context of the 3’ terminus (nick or blunt end). Furthermore, these mutants do not inhibit wt activity on ssDNA substrates. An electrophoretic mobility shift assay supports the model that TREX1 inhibits wt exonuclease activity on double-stranded substrates by binding the DNA and protecting the 3’ termini.

3.2 Experimental Procedures

Materials

The synthetic 30mer oligonucleotide 5’-ATACGACGGTGTTGTCAGACAGGT-3’ was 5’-labeled with 6-carboxyfluorescein (6-FAM) (Operon). Plasmid 1 (9.4 kb) is a derivative of the pMYB5 plasmid (New England Biolabs) and contains a single Nt.BbvCI restriction enzyme site. The plasmid was purified from bacterial cultures and digested with the Nt.BbvCI restriction enzyme (New England Biolabs). Plasmid 2 (6.4 kb) is
a derivative of the pET19b plasmid (Novagen) and contains a single Bam HI restriction enzyme site. Plasmid 2 was purified from bacterial cultures and digested with the Bam HI restriction enzyme (Promega). The 3 kb pBluescript II KS+ phagemid DNA (Stratagene) was prepared as described (152). Linear single-stranded KS+ was prepared by hybridizing a 15mer oligonucleotide (5′-GCCCGGGGGATCCAC-3′) to generate a double-stranded region, followed by digestion with the Bam HI restriction enzyme. Primer 1 5′-AGACGCCAGCGGCGCCATCT-3′, 5′-labeled with 6-FAM and Primer 2 5′-ATTGTGCGTTGCGTGCCGGCG-3′ (Integrated DNA Technologies) were used to amplify a 209-bp DNA for the electrophoretic mobility shift assays.

3.2.1 Enzyme Preparation

The human and mouse TREX1 3′ exonuclease gene fragments encoding the N-terminal 242 amino acid residues were expressed in bacteria as a fusion with MBP in a modified pMAL-C2 vector that we termed the pLM303 plasmid. This modified vector encodes a 6x polyhistidin sequence on the N-terminus of MBP and the Rhinovirus 3C recognition sequence is positioned in between the MBP and the TREX1 gene sequences. The pLM303 constructs containing the wt, D18N, D200N, D200H, or V201D TREX1 gene sequences were transformed into Escherichia coli BL21*(DE3) Rosetta 2 cells (Novagen) for overexpression. Cells were grown to an A_{600}=0.5 and cooled quickly on ice to 17°C. Cells were induced with 1 mM isopropyl-β-D-thiogalactopyranoside and allowed to grow for 15 hours at 17°C. The MBP-TREX1 fusion protein was bound to amylose resin (New England Biolabs), washed, and incubated overnight with PreScission Protease (GE Biosciences) to separate MBP from TREX1. The TREX1 protein
was collected from the amylose column flow-through, dialyzed into low salt buffer, and purified to homogeneity using phosphocellulose chromatography as previously described (94, 123).

Heterodimer human TREX1 protein was obtained by cloning the TREX1 mutant gene sequence into the pLM303x plasmid, in which the N-terminal 6x polyhistidine sequence has been removed, and the TREX1 wt gene sequence into the pCDFDuet-1 plasmid (Novagen). Expression of the pLM303x construct produces MBP-TREX1 and expression of the pCDFDuet-1 construct results in HisNusAHis-TREX1 protein; both constructs encode a Rhinovirus 3C recognition sequence in between the affinity tag and the TREX1 protein. The wt pCDFDuet-1 construct and the mutant pLM303x plasmid were both transformed into *Escherichia coli* Rosetta 2 cells and overexpressed as described for the homodimer TREX1 protein. Overexpression of the pLM303x and the pCDFDuet-1 constructs results in the formation of three forms of the TREX1 dimer: HisNusAHis-TREX1wt/HisNusAHis-TREX1wt homodimer, MBP-TREX1mutant/MBP-TREX1mutant homodimer, and HisNusAHis-TREX1wt/MBP-TREX1mutant heterodimer. The heterodimer is purified from the homodimer TREX1 via column chromatography where the protein was first bound to a nitriloacetic acid resin, which would bind any TREX1 containing the polyhistidine sequence, followed by affinity purification to an amylose resin which would only bind heterodimer protein containing both the polyhistidine sequence and MBP. The amylose resin was washed following the initial binding and then exposed to PreScission Protease for 12 hours at 4°C. The TREX1 wt/mutant heterodimer was then collected
in the column flow-through and purified to homogeneity using phosphocellulose chromatography.

### 3.2.2 Exonuclease Assays

The exonuclease reactions contained 20 mM Tris-HCl (pH 7.5), 5 mM MgCl$_2$, 2 mM dithiothreitol, 100 µg/ml bovine serum albumin, 50 nM fluorescein-labeled 30mer oligonucleotide (ssDNA oligonucleotide assays) or 10 µg/ml plasmid (either plasmid 1 or 2) DNA (dsDNA assays), or 20 µg/ml phagemid (ssDNA phagemid assays) and human TREX1 protein as indicated. TREX1 enzyme and variant mixtures were prepared on ice at 10 times the final concentrations. Incubations were for the indicated times at 25°C. Following incubation, samples were quenched by the addition of three volumes of 100% ethanol and dried in vacuo. For ssDNA oligonucleotide assays, the fluorescently labeled bands were visualized using a Typhoon Trio Phosphorimager (GE Healthcare) and quantified as previously described (153). For dsDNA assays and ssDNA phagemid assays, reaction products were resuspended in 20 µl of TAE-agarose gel running solution and electrophoresed on 0.8% agarose gels (dsDNA) or 0.5% agarose gels (ssDNA phagemid) containing ethidium bromide. DNA was visualized using a BIO-RAD ChemiDoc XRS. Enzyme activity on both dsDNA substrates and the single-stranded phagemid DNA was determined by calculating the time for the wt TREX1 (7.6 nM) to completely degrade the nicked plasmid DNA to dNMP (deoxynucleoside monophosphate) as indicated by the loss of ethidium bromide staining material. For plasmid 1, enzyme activity was further determined by the accumulation of the un-nicked ssDNA visible in the agarose gels.
3.2.3 Electrophoretic Mobility Shift Assays (EMSAs)

A 209-bp DNA fragment was amplified using Taq polymerase (Promega) and Primers 1 and 2 in order to obtain a 5′-fluorescein-labeled product. One and a half µM 209-bp DNA was incubated in TREX1 reaction buffer (20 mM Tris-HCl [pH7.5], 5 mM MgCl₂, 2 mM dithiothrietol [4 mM for the D200N and D18N mutant protein reactions], 100 µg/ml bovine serum albumin) and mouse TREX1 as indicated for 20 minutes at 25°C. The mouse TREX1 protein, which has 77% sequence identity with the human protein, has a pI more favorable for native gel electrophoresis than human TREX1. Five mM MgCl₂ was replaced with either 5 mM CaCl₂ or 5 mM ethylenediaminetetraacetic acid (EDTA) in the control experiments with wt and V201D protein. TREX1 enzyme and variant mixtures were prepared on ice at 5 times the final concentrations. Glycerol was added to each sample to a final concentration of 9% immediately prior to gel loading. Reaction mixtures were electrophoresed on 8% polyacrylamide gels in Tris-borate-EDTA (pH 8.0) buffer at 4°C. The gels were analyzed with a Typhoon Phosphorimager (GE Healthcare). DNA binding was determined by a shift in migration distance of the fluorescently labeled DNA in the presence of TREX1 protein.

3.3 Results

3.3.1 Exonuclease activities of TREX1 AGS and FCL mutants using a 30mer oligonucleotide ssDNA

Amino acids D18 and D200 are both active site residues involved in the coordination of two magnesium ions, and are important for DNA binding and catalysis within
the TREX1 active site (Figure 3.1). The dominant mutation of D18 to asparagine (D18N) has been identified in patients with FCL. The heterozygous mutation of D200 to either asparagine (D200N) or histidine (D200H) accounts for the two known cases of dominant AGS. For patients heterozygous for disease mutations, the cellular form of TREX1 is dependent on allele expression and patients could express either wt/wt homodimer or mutant/mutant homodimer protein. We considered the theoretical possibility that based on allele expression heterozygous patients might also have TREX1 heterodimer (wt/mutant) proteins formed \textit{in vivo}. The interplay between these three forms of the TREX1 dimer in the cell may contribute to the disease phenotype in these patients and there is already some experimental evidence that this may be the case (123).

As a model for the putative short ssDNA nucleotide substrates derived from aberrant replication intermediates, we tested the activity of the D200H and V201D mu-
tants on a 30mer ssDNA. The exonucleolytic activity of the D200H AGS mutant homodimer is reduced by more than 10,000-fold compared with the wt protein on a 30mer ssDNA (Figure 3.2), based on a standard exonuclease assay (123). The wt/D200H heterodimer exhibits a 1.7-fold decrease in activity, suggesting an inactive mutant monomer and a fully active wt monomer within the TREX1 dimer. The activities of the D18N and D200N mutants on the 30mer ssDNA were published previously (62, 94, 123) and show that these homodimer mutants have a greater than $10^4$-fold decrease in exonuclease activity compared with the wt TREX1 protein (Table 3.1) (61, 94). The wt/D18N and wt/D200N TREX1 heterodimer proteins have a 2.6-fold and 1.5-fold reduction in activity, respectively (61, 123).

The V201D homodimer activity on the 30mer oligonucleotide retains 25% of wt TREX1 exonuclease activity and still efficiently degrades the ssDNA substrate (Table 3.1) (94). The wt/V201D heterodimer protein, which may be present in the parents of AGS patients, exhibits a 1.4-fold reduction in activity (data not shown). To date, there is no evidence that persons heterozygous for the V201 mutation exhibit any symptoms.

3.3.2 The TREX1 V201D homodimer mutant exhibits an approximately 4-fold reduction in activity on models of all putative biological substrates

V201 is in the vicinity of the active site and is a residue directly adjacent to the important metal-coordinating residue, D200. The activities of this AGS mutant on a nicked double-stranded plasmid DNA, a blunt-ended dsDNA, and a linear single-stranded phagemid DNA were determined in order to evaluate the activity of the
Figure 3.2: Activities of TREX1 wt and D200H variants on a 30mer ssDNA. Reaction products from standard exonuclease reactions (20 µl) were electrophoresed on 23% urea-polyacrylamide gels (A) and quantified (B) as described under Experimental Procedures. Relative activities of the D200H/D200H homodimer and wt/D200H heterodimer TREX1 are compared with wt TREX1 dimers. The relative activity was calculated as relative activity = 100 x [(fmol of dNMP released/s/fmol of mutant enzyme)/(fmol of dNMP released/s/fmol wt enzyme)].
Table 3.1: Relative Activities of TREX1 Wt and Disease Mutants on a 30mer ssDNA

<table>
<thead>
<tr>
<th>TREX1</th>
<th>Relative Activity</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>1</td>
<td>This study and Refs (93) and (123)</td>
</tr>
<tr>
<td>D18N/D18N</td>
<td>1/160,000</td>
<td>Ref (61)</td>
</tr>
<tr>
<td>wt/D18N</td>
<td>1/2.6</td>
<td>Ref (61)</td>
</tr>
<tr>
<td>D200H/D200H</td>
<td>1/13,000</td>
<td>This study</td>
</tr>
<tr>
<td>wt/D200H</td>
<td>1/1.7</td>
<td>This study</td>
</tr>
<tr>
<td>D200N/D200N</td>
<td>1/20,000</td>
<td>Ref (123)</td>
</tr>
<tr>
<td>wt/D200N</td>
<td>1/1.5</td>
<td>Ref (123)</td>
</tr>
<tr>
<td>V201D/V201D</td>
<td>1/4.1</td>
<td>Ref (94)</td>
</tr>
<tr>
<td>wt/V201D</td>
<td>1/1.4</td>
<td>This study</td>
</tr>
</tbody>
</table>

V201D TREX1 mutant on the putative biological substrates: nicked genomic DNA during GzmA cell death, dsDNA and ssDNA derived from endogenous retroviral elements, respectively. At a concentration of 31.2 nM V201D TREX1, approximately 4-fold higher than the concentration used in the wt TREX1 reactions (7.6 nM), the V201D mutant degrades the nicked double-stranded plasmid DNA, the blunt-ended dsDNA, and the 3 kb ssDNA at a rate similar to that observed for the wt TREX1 protein (Figure 3.3). This suggests that the V201D homodimer TREX1 mutant retains about 25% of wt TREX1 activity on all three substrates. The activity of the heterodimeric wt/V201D TREX1 mutant on these substrates was essentially indistinguishable from that of the wt TREX1 homodimer in these assays (data not shown).

3.3.3 D18N, D200H, and D200N active site mutant homodimers and heterodimers are inactive on nicked or blunt-ended dsDNA substrates and inhibit wt TREX1 activity on either substrate

The context of the DNA 3’ terminus may be an important factor in determining the ability of a TREX1 mutant to degrade a dsDNA substrate. Therefore, the activities of
Figure 3.3: Activities of the wt and V201D mutant homodimers on dsDNA and ssDNA substrates. The TREX1 wt concentration was 7.6 nM and the TREX1 V201D concentration was 31.2 nM for all assays. The times sampled during each time course reaction are indicated above the lanes. (A) Activity on a nicked, 9.4 kb dsDNA plasmid substrate, similar to the putative cellular substrate – nicked genomic DNA. The migration position of the supercoiled dsDNA, nicked dsDNA, and circular ssDNA are indicated. (b) Activity on a linear, blunt-ended dsDNA substrate (6.4 kb) as a model of a dsDNA-derived from an endogenous retroelement. The migration of the supercoiled and linear dsDNAs are indicated. (c) The activity of wt and V201D TREX1 on a 3 kb linear, ssDNA that models a ssDNA retroelement. The migration positions of the linear and circular ssDNAs are indicated.
the TREX1 active site mutants were tested using a dsDNA exonuclease assay (123) on two dsDNA substrates: a circular plasmid DNA with a single nick on one strand and a linearized plasmid with blunt ends. These substrates mimic either nicked genomic DNA from the GzmA cell death pathway (nicked plasmid) or a dsDNA intermediate of the endogenous retroelement replication cycle (blunt-ended DNA).

Nicked plasmid dsDNA

For the single-nicked plasmid assay, TREX1 activity was evaluated based on the disappearance of the nicked dsDNA band and the appearance of a ssDNA band (Figure 3.3a, compare lanes 6 and 7). Over the course of 60 minutes, 7.6 nM TREX1 wt degrades the majority of the nicked strand of the plasmid (Figure 3.3a lane 7). Neither the homodimer D200H/D200H nor the wt/D200H heterodimer TREX1 are active on the nicked plasmid DNA, even at 76 nM protein (Figure 3.4a and data not shown). Both the TREX1 D200H/D200H and the wt/D200H mutants are able to inhibit the activity of wt TREX1 in a competition assay (Figure 3.4b). The mutant TREX1 homodimer and heterodimer protein exhibit inhibition when there is half as much mutant protein present as wt TREX1 protein (Figure 3.4b compare lane 3 with lanes 6 and 11). Furthermore, our data on the D18N and D200N homodimer and heterodimer mutants confirm the defective dsDNA degradation activity previously observed for these mutants on the single-nicked plasmid (data not shown) (123).
Figure 3.4: The TREX1 D200H/D200H homodimer and wt/D200H heterodimer inhibit wt TREX1 activity on a nicked dsDNA plasmid substrate. Exonuclease reactions contain the nicked dsDNA plasmid substrate at 10 µg/ml (1.6 nM nicks) (A) Exonuclease reactions containing nicked dsDNA plasmid and no enzyme (lane 2) or increasing concentrations of the TREX1 wt/D200H heterodimer are indicated. (B) Exonuclease reactions were prepared containing nicked dsDNA plasmid and no enzyme (lane 2), the indicated concentration of wt TREX1 only (lane 3) or a mixture of TREX1 wt with the indicated increased concentrations of TREX1 D200H/D200H (lanes 4-8) or wt/D200H (lanes 9-13).
Blunt-ended dsDNA

The activity of TREX1 on the blunt-ended dsDNA was determined by the disappearance of the dsDNA band (Figure 3.3b). Wt TREX1 (7.6 nM) degrades 10 µg/ml of a blunt-ended dsDNA substrate within 30 minutes (Figure 3.4b compare lanes 2 and 6). The dominant TREX1 homodimers and heterodimers are unable to degrade the blunt-ended dsDNA substrate (Figure 3.5, data not shown). Even at concentrations as high as 230 nM, 30 times more protein than required for the wt TREX1 to degrade the substrate (Figure 3.3b), the TREX1 wt/D18N, wt/D200H, and wt/D200N heterodimers are unable to degrade the blunt-ended dsDNA (Figure 3.5). In competition assays with the wt protein, the dominant mutants are able to inhibit the wt activity on the blunt-ended dsDNA substrate (Figure 3.6). The mutant TREX1 homodimers achieved inhibition of wt protein at approximately 57 nM protein, about three-fourths of the concentration of wt protein present (Figure 3.6).

Figure 3.5: The wt/D18N, wt/D200H, and wt/D200N TREX1 heterodimers are inactive on a linear, blunt-ended dsDNA substrate. The exonuclease reactions were prepared containing linear, blunt-ended dsDNA (10 µg/ml = 4.8 nM free 3' ends) and no enzyme (lanes 2, 9, and 16) or the indicated increased concentrations (in nM) of TREX1 wt/D18N (lanes 3-7), wt/D200H (lanes 10-14), or wt/D200N (lanes 17-21). The positions of migration of the supercoiled dsDNA and linear dsDNA are indicated.
A, B, and C, compare lanes 2 and 7). The wt/D18N heterodimer is close to fully inhibiting the wt protein at a 1:1 ratio (Figure 3.6A compare lanes 2 and 13). The TREX1 wt/D200H mutant fully inhibits wt protein by 57 nM (Figure 3.6B compare lanes 2 and 12) and the wt/D200N mutant fully inhibits wt TREX1 when they are present at equal concentrations (Figure 3.6C, compare lanes 2 and 13).

3.3.4 Dominant TREX1 D18N, D200H, and D200N homo- and heterodimer activities on single-stranded phagemid DNA

A linearized 3 kb single-stranded phagemid DNA was utilized to test the activity of the dominant mutants on a substrate similar to an endogenous ssDNA retroelement. TREX1 activity is indicated by the disappearance of the linear ssDNA band. Wt TREX1 at 7.6 nM fully degrades the ssDNA substrate within 30 minutes (Figure 3.3C). In a control experiment, the wt TREX1 protein is unable to degrade the circular single-stranded phagemid, which lacks an available 3' terminus (data not shown). The D18N, D200H, and D200N homodimers are inactive on the 3 kb linearized, single-stranded phagemid DNA substrate (Figure 3.7). Thirty-eight nM of the wt/D18N, wt/D200H, or wt/D200N heterodimer completely degrades this large single-stranded DNA substrate within 30 minutes (Figure 3.7, compare lanes 2 and 5, 9 and 12, 16 and 19). These results are consistent with the observed trends in activities of these homodimer mutants on the 30mer ssDNA.
Figure 3.6: The D18N, D200H, and D200N TREX1 variants inhibit wt TREX1 activity on a linear, blunt-ended dsDNA. The exonuclease reactions were prepared containing linear, blunt-ended dsDNA (10 µg/ml = 4.8 nM free 3’ ends) and no enzyme (A, B and C, lane 2) the indicated concentration of wt TREX1 only (A, B and C, lane 3) or a mixture of TREX1 wt with the indicated increased concentrations of TREX1 D18N/D18N (A, lanes 4-8), wt/D18N (A, lanes 9-13), D200H/D200H (B, lanes 4-8), wt/D200H (B, lanes 9-13), D200N/D200N (C, lanes 4-8), or wt/D200N (C, lanes 9-13). The positions of migration of the supercoiled dsDNA and linear dsDNA are indicated.

<table>
<thead>
<tr>
<th>TREX1 wt [nM]</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
</tr>
</thead>
<tbody>
<tr>
<td>TREX1 D18N/D18N [nM]</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
<td>38</td>
<td>57</td>
<td>76</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
<td>38</td>
</tr>
<tr>
<td>TREX1 wt/D18N</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
<td>38</td>
<td>57</td>
<td>76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TREX1 wt [nM]</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
</tr>
</thead>
<tbody>
<tr>
<td>TREX1 D200H/D200H [nM]</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
<td>38</td>
<td>57</td>
<td>76</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
<td>38</td>
</tr>
<tr>
<td>TREX1 wt/D200H</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
<td>38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TREX1 wt [nM]</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
<th>76</th>
</tr>
</thead>
<tbody>
<tr>
<td>TREX1 D200N/D200N [nM]</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
<td>38</td>
<td>57</td>
<td>76</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
</tr>
<tr>
<td>TREX1 wt/D200N</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.6</td>
<td>19</td>
</tr>
</tbody>
</table>
3.3.5 Dominant TREX1 active site disease mutants do not inhibit wt TREX1 activity on single-stranded phagemid DNA

Wt TREX1 at 76 nM was incubated in increasing concentrations with homodimer or heterodimer TREX1 active site mutants in the presence of the single-stranded phagemid DNA. The heterodimer, which has activity on this substrate, does not
inhibit the wt activity on the ssDNA substrate, which is fully degraded during the 30 minute course of the reaction (Figure 3.8, lanes 9-13). The homodimer mutants that do not have intrinsic activity on this ssDNA substrate also are unable to inhibit the wt activity on the single-stranded phagemid DNA (Figure 3.8, lanes 3-8). This is in stark contrast to the dsDNA substrates, on which both homo- and heterodimeric active site mutants are able to inhibit wt TREX1 activity.

3.3.6 The dominant TREX1 autoimmune disease mutants inhibit wt TREX1 activity on dsDNA by protecting the DNA 3′ termini

An electrophoretic mobility shift assay (EMSA) demonstrates that, in the presence of magnesium, the dominant TREX1 active site mutants are not only inactive on dsDNA, but they bind it in a stable complex (Figure 3.9). This is indicated by the shift of the free DNA band up to a fully bound state (7 µM D18N, 4 µM D200H, and 3 µM D200N) and after boiling the TREX1-bound DNA to denature the protein, the DNA is still present with the same migration distance as the free DNA. In contrast, 76 nM wt TREX1 degrades the dsDNA substrate as indicated by the disappearance of the DNA band from the wt TREX1 lanes (Figure 3.9). Once the dsDNA substrate is fully bound by the mutant TREX1 protein, it is bound in a stable complex that does not enable wt TREX1 activity. In a competition experiment where wt and mutant TREX1 were added simultaneously to the reaction mix, the homodimer TREX1 mutants form a stable complex with the dsDNA and inhibit the activity of wt TREX1 (Figure 3.9).

In the presence of the divalent metal calcium, which enables TREX1 DNA binding, but does not support catalysis, the wt protein stably binds DNA (Figure 3.10).
Figure 3.8: The D18N, D200H, and D200N TREX1 variants do not inhibit wt TREX1 activity on a linear ssDNA. The exonuclease reaction mixtures were prepared containing linear, ssDNA (20 µg/ml = 20.5 nM free 3' ends) and no enzyme (A, B, and C, lane 2) the indicated concentration of wt TREX1 only (A, B, and C, lane 3) or a mixture of TREX1 wt with the indicated increased concentrations of TREX1 D18N/D18N (A, lanes 4-8), wt/D18N (A, lanes 9-13), D200H/D200H (B, lanes 4-8), wt/D200H (B, lanes 9-13), D200N/D200N (C, lanes 4-8), or wt/D200N (C, lanes 9-13). The positions of migration of the circular ssDNA and linear ssDNA are indicated.
Furthermore, in the presence of 5 mM EDTA, wt TREX1 is still capable of binding dsDNA (Figure 3.10). This suggests that metal may not be required for DNA binding by TREX1 as previously thought, but that the metal is required only for catalysis. The recessive TREX1 V201D mutant shows a similar trend in activity and DNA binding in the presence of magnesium, calcium, and EDTA. TREX1 V201D is active with magnesium and forms a stable complex with DNA in the presence of calcium or when no metal is present (5 mM EDTA) (Figure 3.10).

3.4 Discussion

3.4.1 Dominant TREX1 autoimmune disease is likely triggered by the cells inability to dispose of dsDNA

Four distinct substrates have been proposed for TREX1. This study tested the possibility that TREX1 mutants, specifically the heterodimer form, might have a dominant effect on these substrates. The activities of the dominant TREX1 active site mutants on dsDNA and ssDNA suggest a primary role for TREX1 in the degradation of dsDNA in the cell. The expression of the mutant allele in patients with FCL or dominant AGS would result in the competitive inhibition of the dsDNA degradation activity of wt TREX1, regardless of whether the mutant is expressed as a homodimer or a heterodimer. Importantly, neither the heterodimer nor the homodimer is capable of inhibiting the wt TREX1 activity on a ssDNA substrate. This signifies a departure from the understanding that TREX1 exhibits similar behavior on both dsDNA and ssDNA.

These results implicate TREX1 in a dsDNA degradation pathway. Three cellular
Figure 3.9: The D18N, D200H, and D200N dominant TREX1 variants inhibit TREX1 wt activity by binding and protecting the 3’ termini of a dsDNA. The migration of the free 209-bp dsDNA is indicated in lane 1 of each polyacrylamide gel. DNA binding reactions were prepared under conditions where the wt TREX1 protein is active (refer to Experimental Procedures for more detail). Each reaction contained 1.5 µM 209-bp dsDNA and no enzyme (free 209mer), wt TREX1 at the indicated concentration, TREX1 D18N/D18N, D200H/D200H, or D200N/D200N at the indicated increased concentrations (D18N, D200H, and D200N lanes 3-5). Boiling the fully bound 209mer regenerates the original substrate. The addition of the indicated concentrations of wt TREX1 at a concentration of D18N, D200H, or D200N sufficient to fully bind the substrate, protects the DNA from wt exonuclease activity. Enzyme concentrations are given as the concentration of TREX1 monomer.

roles have been proposed for TREX1 including the degradation of nicked, dsDNA during GzmA cell death, degradation of ssDNA or dsDNA endogenous retroviral elements, and disposal of single-stranded DNA replication intermediates. The context of the free 3’ end, whether it is at a nick site or at a blunt ended DNA terminus, does not influence the inability of the dominant TREX1 mutants to degrade dsDNA. Therefore, it would be possible for accumulation of dsDNA to occur from loss of TREX1 activity during GzmA cell death, from dsDNA derived from endogenous retroelements, or possibly through another unidentified pathway. The TREX1 wt/mutant
heterodimers retain activity on ssDNA and the dominant homodimer mutants, which are inactive on ssDNA, do not inhibit wt activity on ssDNA. Thus, the likelihood that the primary cellular role of TREX1 can be attributed to a ssDNA degradation activity is much less probable.

3.4.2 EMSA assays reveal a mechanism where dominant TREX1 mutants inhibit wt TREX1 activity by protecting the DNA 3′ termini of dsDNA substrates

The inhibition of wt TREX1 exonuclease activity by the TREX1 mutants could be achieved by multiple mechanisms. A novel protein-protein interaction between the wt and mutant TREX1 that sequesters the wt protein could result in the observed loss in wt activity. Another possibility is that the mutant homodimers and wt/mutant heterodimers bind the DNA at a 3′ terminus and trap it in an unproductive complex, inhibiting wt TREX1 exonucleolytic activity. The EMSA assays support the second model by showing wt inhibition when the TREX1 mutants are bound to the dsDNA.

The ability of dominant TREX1 autoimmune disease mutants to bind the 3′ termini of a dsDNA and inhibit wt TREX1 activity, but not on ssDNA suggests a unique binding interaction between TREX1 and dsDNA. There are possible structural explanations for the stabilization of dsDNA in the TREX1 active site. The flexible DNA binding loop, adjacent to the TREX1 active site, may interact exclusively with the complementary strand of a dsDNA substrate during TREX1 degradation. This
Figure 3.10: The wt TREX1 and recessive V201D TREX1 mutant protein bind dsDNA in the absence of metal. Both wt and V201D TREX1 are active on the dsDNA substrate in the presence of magnesium. In the presence of calcium or without any metal present (5 mM EDTA) both wt TREX1 and the V201D TREX1 mutant bind dsDNA. Each reaction contains 1.5 µM 209-bp dsDNA, no enzyme (free 209mer), wt or V201D TREX1 at the indicated concentrations.

loop is disordered in the X-ray crystal structure of TREX1 in complex with a four-nucleotide DNA substrate indicating flexibility (94), but may adopt a more stable conformation upon binding dsDNA. Another possibility is that the amino acid R128, which has a putative role in melting dsDNA for TREX1 degradation (94), stabilizes the interaction between TREX1 and dsDNA. It is also possible that the DNA binding loop and R128 together contribute to a stabilized interaction between TREX1 and dsDNA. More work is required to better understand the difference between how TREX1 interacts with double-stranded versus single-strand DNA substrates.
3.4.3 The recessive TREX1 V201D mutant may follow a distinct pathway to autoimmunity

The V201D homodimer has reduced activity relative to wt TREX1, but still efficiently degrades both ssDNA and dsDNA substrates. Wt TREX1 protein (7.6 nM) nearly fully degrades a single strand of a nicked 9.4 kb plasmid in 60 minutes (Figure 3.3A). The V201D protein is able to degrade the same substrate in 60 minutes with only a 4-fold increase in the protein concentration. This TREX1 mutant exhibits similar activities on the blunt-ended dsDNA and the linear ssDNA substrates. Even with a 75% reduction in catalytic activity relative to the wt protein, the V201D protein retains a robust exonuclease activity. However, no 3’ exonuclease activity was detected in fibroblast cell extracts derived from a patient with this recessive mutation (56). While there is a possibility that even this modest reduction in TREX1 activity is sufficient to result in the AGS phenotype, it is more likely that the V201D mutation may result in the abrogation of TREX1 cellular function via a different mechanism.

So far neither exonuclease activity nor protein structure studies clearly explain the induction of the severe AGS phenotype by the recessive V201D mutation in the TREX1 dimer. Therefore, there is a distinct possibility that recessive AGS mutants follow a different pathway to disease than the dominant active site mutants. Furthermore, all recessive TREX1 mutations that cause AGS may not result in the AGS phenotype via a single mechanism. The recessive R114H mutation at the TREX1 dimer interface causes AGS. This mutant exhibits a 34-fold decrease in catalytic activity on a single-stranded 30mer DNA and an approximately 300-fold decrease in
activity on a nicked plasmid DNA substrate (123). The AGS phenotype in patients with the R114H mutation might be explained by the defective catalytic activity alone. The V201D mutant, however, retains a robust catalytic activity on all ssDNA and dsDNA substrates tested. The observed lack of exonuclease activity in cell extracts derived from a patient with the V201D mutation may be explained by a modification in protein-protein interactions or post-translational modification that inhibits TREX1 cellular function.

Post-translational modification may be important for TREX1 activity (62). Altered patterns of TREX1 modification at putative phosphorylation or ubiquitination sites due to disease mutations could result in TREX1 activity being turned off or the lack of a modification could be a missing signal for TREX1 translocation to the nucleus. Post-translational modifications could result in the targeting of TREX1 to the proteasome for degradation. The effective loss of TREX1 cellular function due to changes in post-translational modification would result in the accumulation of self DNA. Nucleic acid receptors in the cell do not have a sensitive mechanism for distinguishing self DNA from the DNA of an invading pathogen. Therefore, the accumulation of a patients’ DNA could stimulate an aberrant immune response.

TREX1 homodimer and heterodimer mutant inhibition of wt TREX1 activity may be important for disease progression in patients with dominant, active site mutations. When TREX1 activity is required, these mutants will bind the DNA without degrading it, and prevent any wt protein present from degrading the DNA as well.
Over time, this would result in an aberrant accumulation of the patient’s own genetic material that may initiate an autoimmune response and could be the major trigger for AGS and FCL. Furthermore, the presence of TREX1-dsDNA complexes in the cell may also result in the production of autoantibodies specific to these complexes. More work is required to determine whether autoantibodies to a TREX1-dsDNA complex are relevant in patients.

Acknowledgements

This work was supported in whole or in part by an American Heart Association predoctoral fellowship to SLB, American Cancer Society Grant RSG-04-187-01-GMC (TH) and National Institutes of Health Grant R01-GM069962 (FWP).
Chapter 4: Conclusions
Fifteen - twenty-five million Americans suffer from some type of autoimmune disease and its incidence in the United States is on the rise. Unfortunately there are very limited resources for the treatment of autoimmune disorders and there is no cure. In order to improve the ability of clinicians to treat autoimmune disorders it is important to understand the pathophysiology of autoimmunity. Antinuclear antibodies including those to DNA are commonly detected in patients with autoimmune diseases including systemic lupus erythematosus, Aicardi-Goutières syndrome, and familial chilblain lupus. This suggests a role for the accumulation of host nucleic acids in the aberrant stimulation of an immune response by self DNA or RNA. Furthermore, the deficiency in nuclease activity has been implicated in autoimmune disease in both mice and humans. These results suggest that nucleases may play a role in degrading host DNA and RNA within cellular compartments and the extracellular environment to avoid autoimmune activation.

There is little structural basis for the differentiation between self and nonself nucleic acids, providing a unique role for nucleases in the prevention of autoimmunity. These nucleases may be critical for degrading self DNA in the cytosol (TREX1), lysosomes (DNase II), and extracellular space (DNase I) to avoid host DNA recognition by cellular sensors. TREX1 null mice die prematurely from progressive cardiomyopathy and exhibit an autoimmune phenotype (55, 107). Mutations within the trex1 gene have been identified in SLE, AGS, FCL, and RVCL autoimmune disease patients, implicating TREX1 with a function critical for avoiding self DNA accumulation, recognition, and stimulation of autoimmunity.
4.1 New developments in the understanding of the role of TREX1 in autoimmune disease

TREX1 is a homodimeric 3′ exonuclease that is localized to the cytosol until a stimulus (such as oxidative stress) triggers its translocation to the nucleus. TREX1 has a putative interaction with the SET protein within the SET complex (110); other candidate binding partners for TREX1 in the cell are unknown. The question of whether or not TREX1 activity is modulated by post-translational modification remains unanswered. Wt TREX1 is fully active on ssDNAs and double-stranded DNAs with a nick, blunt end, or 3′ mispairs and has putative roles in degrading nicked genomic DNA during granzyme A cell death (110), eliminating small, 60-65 nucleotide ssDNA aberrant replication intermediates (117), and disposing of either ssDNA or dsDNA derived from endogenous retroelements (107).

Studies to evaluate homodimeric and heterodimeric TREX1 mutant protein were first undertaken to evaluate TREX1 exonuclease activity on a single-stranded 30mer DNA substrate (61, 94). These studies provided only limited correlation between the observed diminished TREX1 catalytic activity and the autoimmune phenotype. Further studies have examined the possible role of a nicked double-stranded plasmid DNA substrate in disease (123) (Chapter 3). The study by Lehtinen et al. demonstrated an ability of the dominant D18N and D200N homodimer and heterodimer TREX1 mutants to inhibit the activity of wt TREX1 on a nicked plasmid substrate. The ability of TREX1 to degrade larger ssDNA substrates and linear, blunt-ended dsDNA substrates as models for endogenous retroelement DNA was determined here.
The recessive AGS V201D mutation

The recessive V201D mutant, which causes AGS, retains 25% of wt activity on a ssDNA 30mer (94). This mutant exhibits a similar activity on a large ssDNA, and both nicked and blunt-ended dsDNA substrates (Chapter 3). The wt/V201D heterodimer exhibits an activity that is essentially indistinguishable from wt exonuclease activity on all ssDNA and dsDNA substrates tested. The crystal structure revealed that the V201D homodimer mutant does not exhibit any global structural changes within the protein, at least in its truncated form, that might play a role in autoimmune disease (Chapter 2). Morphological changes within the homodimeric V201D active site likely explain the observed reduction in catalytic activity. In the structure of wt TREX1 in complex with single-stranded DNA, there are two active site conformations within the TREX1 dimer. In one protomer, H195 is flipped out of the active site and the metal ions are separated by 3.1 Å, this is termed the resting state of the enzyme. In what is called the active state, H195 is in the inward conformation and the active site metals are separated by 3.6 Å. The increase in inter-metal distance from the resting to the active state results from a motion in metal A away from metal B and towards the substrate, a motion that may contribute to directing nucleophilic attack during catalysis. In the V201D homodimer structure, the catalytic histidine (H195) is in the inward conformation in both monomers of the dimer. Furthermore, the inter-metal distance in the V201D active site is increased to between 3.7 and 3.8 Å in both protomers. Both of these structural features suggest a loss of coordinated motion between the two active sites and reduced catalytic activity by the V201D
Neither biochemical nor structural data on the V201D homodimer provide a definitive explanation for the severity of the observed AGS phenotype in patients recessive for this mutation. There is no 3′ exonuclease activity detected in cell extracts derived from a patient homozygous for the V201D mutation (56), which suggests a different biological explanation for how TREX1 cellular function is lost due to this mutation. TREX1 may be targeted to the proteasome for degradation or its activity may be turned off via altered post-translational modification or protein-protein interactions in patients with the V201D mutant. Sequence analysis suggests an ubiquitination site in the flexible DNA binding loop adjacent to the TREX1 active site. The V201D mutation is within the active site vicinity and may alter the pattern of TREX1 post-translational modification within the DNA binding loop, such that cellular TREX1 exonucleotic activity is lost. Two additional putative ubiquitination sites and possible sites of phosphorylation (62) are located between the TREX1 catalytic domain and the putative transmembrane helix within the TREX1 C-terminal region. Another possible explanation for the loss of TREX1 cellular activity is that the V201D mutation may alter the contacts TREX1 mediates within the SET complex which results in aberrant trafficking of the enzyme.

**Dominant active site TREX1 mutations**

Dominant mutations that have been identified in patients with FCL and *de novo* AGS may result in the presence of three different cellular forms of TREX1 depending
upon allele expression. There may be wt/wt homodimer, mutant/mutant homodimer, and wt/mutant heterodimer TREX1 present in the patients’ cells. For this reason, it is important to evaluate the possible roles of each of these three forms of the TREX1 protein in the autoimmune disease phenotype. The exonuclease activities of TREX1 autoimmune disease mutants positioned within the active site vicinity were evaluated on different DNA substrates as models for putative in vivo substrates. A 3 kb ssDNA, blunt-ended dsDNA, nicked dsDNA, and a single-stranded 30mer DNA were tested as models for a ssDNA retroelement, dsDNA retroelement, nicked genomic DNA during cell death, and ssDNA aberrant replication intermediates, respectively.

The dominant homodimer and heterodimer TREX1 mutants are unable to degrade dsDNA, regardless of the context of the 3′ terminus. In contrast, the dominant heterodimer mutants retain approximately 50% activity on both a small 30mer ssDNA and a 3 kb ssDNA substrate. A major discovery, however, was that both the dominant homodimer and heterodimer TREX1 mutants were only able to inhibit wt TREX1 activity on the dsDNA substrates. Even the homodimer mutants which are inactive on both double- and single-stranded DNA were unable to inhibit wt activity on the ssDNA substrate. This result suggests that double-stranded DNA may be the primary TREX1 substrate in vivo.

The mechanism by which dominant TREX1 mutants inhibit wt activity could be explained by at least two possibilities. First, it is possible that the TREX1 mutants mediate a novel protein-protein interaction with the wt TREX1 protein that inhibits TREX1 activity. The fact that the TREX1 mutants are only able to in-
hibit wt TREX1 activity on dsDNA and not ssDNA makes this possibility less likely, although it is possible that a change in conformation upon binding dsDNA would mediate the interaction. A second possibility is that the dominant mutants have a stabilized interaction with dsDNA substrates and that they bind and protect the DNA 3′ termini of dsDNA, preventing wt TREX1 from degrading the substrate. An EMSA demonstrated that the dominant TREX1 mutants inhibit wt TREX1 activity by binding and not releasing the substrate (Chapter 3). This suggests that the dominant TREX1 mutant inhibits wt activity by protecting the DNA 3′ termini.

The structures of the dominant active site mutants (as described in chapter 2) provide additional insight into the activities of the dominant FCL and AGS TREX1 mutants. As for the V201D homodimer mutant, there are no global structural changes in the dominant mutants. The structure of the D18N mutation, which causes familial chilblain lupus, revealed that the catalytic histidine was in the inward conformation in both monomers. The motion of this histidine residue may be important for TREX1 catalysis and the consistent position of this residue in both monomers implies a possible loss of coordination between the active sites. The inter-metal distance in the D18N active sites was increased to greater than 4 Å in both protomers. This increase of between 0.5 and 1.0 Å relative to the distances observed in the wt substrate structure may no longer enable positioning of the nucleophile for attack on the scissile phosphate. The increased metal ion spacing may stabilize the inward conformation of the catalytic histidine, explaining the single conformation observed in the structure. Together, the increased metal spacing and single H195 conformation explain the loss
of exonuclease activity observed for the D18N mutant.

The structures of the dominant AGS D200H and D200N mutants reveal that these TREX1 proteins have similar active site morphologies that likely explain their observed loss in catalytic function. There is no density for metal ion A in the active site of either D200 mutant. Metal ion A is believed to be important for lowering the pK$_a$ of a water molecule and helping direct nucleophilic attack on the scissile phosphate. Loss of this metal from the active site prevents the coordination of a water molecule in position for nucleophilic attack. This alone could explain the observed lack of activity by the homodimer mutant enzymes on all substrates tested. The loss of metal A from the active site may additionally explain the position of H195, which is flipped out of the active site in both protomers of the D200H and D200N mutant structures. If the positive charge in the active site helps stabilize the inward conformation of the histidine, this metal may be required for that motion. Furthermore, D200 forms a hydrogen bond with the backbone amide of H195, stabilizing the inward conformation observed in the wt structure as well as the D18N and V201D structures. The mutation of aspartate 200 to either asparagine or histidine would preclude the formation of this hydrogen bond and may hinder the normal function of the catalytic histidine residue.

The dominant TREX1 D18N, D200H, and D200N homodimer and heterodimer TREX1 mutants are all able to inhibit wt TREX1 activity on dsDNA substrates, regardless of the context of the DNA 3’ terminus. These mutants do not have the same inhibitory affect with a ssDNA substrate. Furthermore, the R114H mutant, which causes SLE as a heterozygous mutation and AGS as a recessive mutation, is
unable to inhibit wt TREX1 activity on dsDNA. This indicates a unique interaction between dsDNA and the dominant active site mutants that is not present in other TREX1 mutants that are capable of causing dominant TREX1-mediated autoimmune disease.

Within the TREX1 active site, there may be a delicate balance between the coordination of the two metals and charge repulsion between the ions. During normal TREX1 catalysis, metal A may get bumped out of the active site, which could in turn knock the substrate out of the active site. This could explain the observed non-processive nature of TREX1. The increase in the inter-metal distance to over 4 Å in the D18N structure and the absence of metal A from the D200 mutant structures may therefore explain the ability of these mutants to protect the 3′ termini from wt TREX1 activity. The R114H mutation, which is located at the TREX1 dimer interface, is unlikely to result in a change in metal coordination within the enzyme active site, although this is not known for certain since there is no reported structure for this TREX1 variant. The increased metal distance observed in the D18N structure may stabilize the position of metal A within the active site, resulting in a more stable interaction with the DNA and enabling this mutant to exert dominance over the wt protein. However, the fact that D18N retains metal A within the active site may also provide an explanation for the less severe phenotype associated with FCL. The D200 mutants, which lack metal A in the active site, may virtually never release dsDNA once they bind it and this may explain their dominance and the severe AGS phenotype.
The structure of the TREX1 apoprotein in combination with the structures of TREX1 wt and mutants in complex with a ssDNA substrate reveal a mechanism for TREX1 that may also be relevant for other DEDDh nucleases. In the absence of metal and DNA, H195 is disordered, but this residue adopts one of two conformations upon metal and DNA binding. There is evidence for cooperation between the TREX1 dimers based on the observation of two active site conformations within the wt substrate complex, where there is a putative active protomer and a resting protomer. The metals are coordinated at 3.1 Å when H195 is in the flipped out (resting) conformation and 3.6 Å when H195 is in the inward (active) conformation in the structure of wt TREX1 in complex with a four-nucleotide substrate. The TREX1 monomers may alternate between the active and resting states throughout catalysis, resulting in only one protomer being active at a time. Both metal A and the substrate may be lost from the active site between each nucleotide removed, resulting in the non-processive nature of TREX1.

The work in this dissertation demonstrates:

- Dominant mutations within the TREX1 active site further stabilize the contacts between TREX1 and a dsDNA substrate enabling these mutants to bind and protect the 3’ termini of dsDNA substrates and inhibit wt TREX1 activity. The increased inter-metal distance observed in the D18N mutant structure and the loss of metal A from the active site in both D200 mutant structures suggests that metal A may play a role in the release of substrate from the TREX1 active site.
• TREX1 may have more extensive interactions or a more stable interaction with a dsDNA substrate than with ssDNA. This is based on results demonstrating that the dominant TREX1 active site mutants that cause FCL and de novo AGS are able to inhibit wt TREX1 activity on a dsDNA substrate, but not on a ssDNA substrate.

• AGS caused by the recessive V201D mutation may not be explained by the modest loss in catalytic activity or the truncated TREX1 protein does not mimic full length TREX1 in this case. There may be changes in protein-protein interactions or post-translational modification to the TREX1 protein due to this mutation that result in the functional loss of TREX1 3′ exonuclease activity.

**Future Directions**

Understanding the underlying etiology of a disease enables the development of therapeutics targeted more specifically to patients with specific disease-linked mutations. This study provides a more thorough understanding of autoimmune disease caused by dominant active site mutations in the TREX1 3′ exonuclease. Furthermore, these results suggest an undiscovered mechanism for disease caused by the recessive V201D mutation. There are further experiments that would contribute to our understanding of the involvement of TREX1 in these complex autoimmune diseases.

In AGS, some cases of SLE, and RVCL, the mutation of TREX1 causes severe neurological damage in patients, specifically in the white matter of the brain. This suggests a unique role for TREX1 function in the brain, in particular in cell types such
as oligodendrocytes that form the myelin sheath around neurons and are referred to as the “white matter.” There is a possibility of differential regulation of TREX1 activity in cells such as oligodendrocytes that distinguish them from other cell types and explains the common manifestation of TREX1 disease with white matter abnormalities. The study of TREX1 transcription and protein expression, cellular localization, and post-translational modification in these cells relative to other cell types may provide continued insight into the underlying mechanism of TREX1 mutation in autoimmune disease.

Post-translational control of TREX1 function has never been explored and may play a critical role in how some TREX1 variants lead to autoimmune disease. The recessive V201D TREX1 mutation that causes the severe autoimmune disease, AGS, retains approximately 25% of wt TREX1 exonuclease activity on all ssDNA and dsDNA substrates tested. Exploring the possibility of differences in post-translational modification of the V201D mutant enzyme and the amount or location of the protein in patient cells may provide a clearer picture of autoimmune disease in patients with this mutation. V201D, which is directly adjacent to one of the metal-coordinating residues within the TREX1 active site, is also in close proximity to the flexible DNA binding loop near the TREX1 active site. This loop contains a putative ubiquitination site. If ubiquitination within this loop is important for TREX1 function or becomes a novel ubiquitination site in the V201D variant, then the cellular function of the V201D TREX1 protein could be inhibited.

The differences between how TREX1 interacts with single- and double-stranded
DNA substrates remain unclear. Determining the X-ray crystal structure of TREX1 in complex with dsDNA as well as with a larger ssDNA substrate would contribute to a better understanding of these differences. Electron microscopy of TREX1 using either a ssDNA or a dsDNA may reveal important differences in morphology between TREX1-ssDNA and TREX1-dsDNA complexes. DNA footprinting studies with both ssDNA and dsDNA substrates may enable the determination of differences in the quantity or pattern of DNA bases protected from cleavage due to the interaction with TREX1.

The evidence for a stable interaction between dominant TREX1 active site mutants and dsDNA suggests that these complexes may exist in the cells of patients with these mutations. There is evidence for antibodies to protein-DNA complexes, including nucleosomes (9, 47, 121, 154, 155), in patients with autoimmune disease. Patients with dominant active site TREX1 mutations should be tested for the presence of autoantibodies to a TREX1-dsDNA complex. Structural data show an overall conservation of the TREX1 RNase H-like fold in each of these TREX1 mutants, therefore, antibodies to the mutant may additionally result in a lack of tolerance for the wt TREX1 protein in these patients that would further amplify the autoimmune response.

Diminished TREX1 3’ exonuclease activity is a common theme among TREX1 mutants identified in patients with autoimmune disease. Gene therapy or enzyme therapy, where TREX1 activity may be restored, is a possible therapeutic treatment
for these patients in the future. Before this is possible, it would be critical to define specific mutations where there is a lack of normal TREX1 activity, without intolerance for the TREX1 enzyme. This requires further research to increase our understanding of the underlying etiology of autoimmune disease caused by TREX1 mutations in order to develop safe and effective treatments.
Bibliography


basis for proofreading during replication of the Escherichia coli chromosome, ”


[124] C. M. Hedrich, B. Fiebig, F. H. Hauck, S. Sallmann, G. Hahn, C. Pfeiffer,


Scholastic Vitae

Suzanna L. Bailey

Birthplace and Date:  Greensboro, North Carolina
                   September 12, 1982

Undergraduate Study:  Furman University
                      Greenville, South Carolina
                      B.S., Chemistry, 2005

Graduate Study:  Wake Forest University School of Medicine
                 Winston-Salem, North Carolina
                 Ph.D., Biochemistry, 2010

Research Experience

   Doctoral Research, Wake Forest University School of Medicine;
   Winston-Salem, NC. 08/2005–05/2010. Dr. Thomas Hollis advisor

   Undergraduate Research Assistant, Furman University; Greenville, SC.
   05/2003–06/2005. Dr. Moses Lee advisor

Grants

   2008  American Heart Association Predoctoral Fellowship

Academic Awards
2009  Herbert C. Cheung Award for Outstanding Graduate Student in Biochemistry

2008  Virginia Tech Structural Biology Symposium Poster Prize
2008  Artom Fellowship for Outstanding Graduate Student in Biochemistry

**Relevant Courses**

2009  Attended NSLS RapiData course in macromolecular structure solution

**Teaching Experience**

2009  Teaching Assistant for the Department of Biochemistry
       Drug Discovery Laboratory Course

**Service/Leadership**

2007-2008  Graduate Student Association Departmental Representative

**Publications**

*Bailey, SL.*, Harvey, S., Perrino, FW, Hollis, T. Dominant TREX1 autoimmune disease mutants inhibit wt TREX1 exonuclease activity on double-stranded DNA by competitive binding and protecting DNA 3 termini. In preparation.

*Bailey, SL.*, Lee-Kirsch, MA, Harvey, S., Perrino, FW, Hollis, T.

Insight into TREX1 catalysis from the crystal structures of autoimmune disease mutants. In preparation.


Abstracts/Poster Presentations


Buchmueller, KL., **Bailey, SL.**, Matthews, DA., Taherbhai, ZT., Register, JK., Davis, ZS., Bruce, CD., O’Hare, CO., Hartley, JA., Lee, M. (2004) DNA Base Pair Recognition by Formamido/Pyrrole and Formamide/Imidazole Pairings in Stacked Polyamides. *Southeastern Regional Meeting of the American Chemical Society, Raleigh, NC.*