7-31-2023

UvrD1 Helicase of Mycobacterium tuberculosis that can be Activated by Multiple Unique Mechanisms

Ankita Chadda
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Developmental, Regenerative and Stem Cell Biology

Dissertation Examination Committee:
Eric Galburt, Chair
Michael Caparon
Douglas Chalker
Guy Genin
Michael Greenberg

UvrD1 Helicase of *Mycobacterium tuberculosis* that can be Activated by Multiple Unique Mechanisms

by
Ankita Chadda

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2023
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I thank my family, and friends for being supportive during my PhD. I thank my mentor Eric Galburt for his patience, optimism, and advise on how to prepare presentations, making hypotheses and testing them. I thank members of the Galburt Lab for the supportive lab environment and getting me started on protein biochemistry. I thank my thesis committee members for their time and support.

I thank the lab of Tim Lohman for providing lab space and reagents for my experiments. I would also like to thank members of Lohman lab for their help in experiments and in scientific discussions. I would specially like to thank Tim Lohman for his constant guidance and interest in the project throughout my thesis.

Ankita Chadda

Washington University in St. Louis

August 2023

ABSTRACT OF THE DISSERTATION

UvrD1 Helicase of *Mycobacterium tuberculosis* that can be Activated

by Multiple Unique Mechanisms

by

Ankita Chadda

Doctor of Philosophy in Biology and Biomedical Sciences

Developmental, Regenerative and Stem Cell Biology

Washington University in St. Louis, 2023

Associate Professor Eric Galburt, Chair

Mycobacterium tuberculosis (*Mtb*) causes tuberculosis disease in humans and is one of the leading causes of death worldwide due to infectious agents. During infection, *Mtb* is exposed to reactive oxygen species and reactive nitrogen intermediates from the host immune response that causes DNA damage. UvrD-like helicases are involved in DNA repair and use energy from ATP hydrolysis to translocate on single stranded DNA (ssDNA) or unwind double-stranded DNA (dsDNA) to remove the damaged DNA strand. Previous studies on UvrD-like helicases have shown that they exist in a monomer-dimer equilibrium and unwind only as dimers in absence of accessory factors. However, *Mtb* UvrD1 had previously been suggested to perform unwinding as a monomer. Using quantitative biophysical assays including analytical ultracentrifugation and stopped-flow kinetics, I demonstrated that *Mtb* UvrD1 exists in monomer, dimer, and higher-order oligomeric forms, where the dimer is redox-dependent. I identified a cysteine in the 2B domain

that is required for oxidative dimerization, thus demonstrating that the 2B domain is directly involved in dimerization. Furthermore, although the monomer of UvrD1 can bind and translocate on ssDNA, only the dimer unwinds dsDNA supporting the notion that helicase activity requires more than just translocation on ssDNA. My results suggest a model where UvrD1 senses the oxidative conditions within human macrophages during infection through dimerization, resulting in the activation of its DNA-unwinding activity needed for DNA repair and other DNA metabolic pathways. In addition, it also highlights the dimerization interface of the UvrD-family helicases that was previously unknown.

In related work using cryo-electron microscopy, I have visualized the structure of the *Mtb* UvrD1 dimer where the disulfide bond is clearly visible. Data analysis resulted in 4-5 Å resolution structures of UvrD1 in two different conformations (open and closed) that might represent two different conformations of a dimer.

Independent studies have shown that the helicase activity of *Mtb* UvrD1 is activated by *Mtb* Ku which is a homodimer homologous to the eukaryotic Ku70/Ku80 heterodimer and which plays a role in double-stranded break repair via non-homologous end-joining. I followed up on this observation in the context of my results of the redox-sensitive dimeric structure of UvrD1 and showed that Ku specifically activates the unwinding activity of a monomer of UvrD1. Importantly, this activity is weak compared to that of the dimer and is only observed under multi-round conditions and therefore depends on multiple interactions between UvrD1 monomer, Ku, and DNA. Dissecting the interaction further, I revealed that the Ku-UvrD1 interaction occurs via the C-terminal Tudor domain of UvrD1. I also found that 3 Ku dimers are bound to the DNA substrate and that Ku appears to be able to load on dsDNA even in the presence of ssDNA overhangs. Since

the maximal Ku-dependent unwinding occurs under saturating conditions of Ku, my results suggest a model where maximum number of Ku dimers bound to DNA results in more unwinding.

In summary, my characterization of *Mtb* UvrD1 highlights its potential role in multiple DNA repair pathways through different mechanisms of activation and contributes to our understanding of the UvrD-family of helicases which are found across biology.

Chapter 1: Introduction

The maintenance of genome and its replication are central features in all kingdoms of life. Nucleic acid metabolism during the process of replication, recombination, and repair requires the separation of double stranded DNA (dsDNA) to single stranded DNA (ssDNA)[1]. For instance, during DNA replication the dsDNA separates into single strands so the DNA polymerase can bind and synthesize the new polynucleotide strand[2]. Likewise, during DNA repair the DNA damage is being sensed by recognition enzymes and endonucleases are recruited that excise the lesion and the ssDNA with the lesion is separated and removed to synthesize a new polynucleotide strand without damage[3–5]. Because of the stability of the DNA duplex the rate of spontaneous single strand separation is very slow. Nature has evolved specialized enzymes called helicases that perform this function by coupling chemical potential energy of the nucleoside triphosphates (NTP) with strand separation[6]. Thus, helicases use free energy of binding or/and hydrolysis and product release of NTP to perform work required to separate dsDNA to ssDNA[6,7]. For performing unwinding of dsDNA substrates, helicases have the properties to bind nucleic acids. The binding process may either be completed when helicase must bind and dissociate from DNA multiple times to unwind DNA, or it may bind and unwind DNA in a single round condition. In addition, these specialized enzymes can also translocate on the single stranded DNA[8,9].

Mycobacterium tuberculosis (*Mtb*) causes tuberculosis (TB) disease in humans[10]. The lack of an effective vaccine, the long and expensive drug regimens and few diagnostic tools has led TB to be one of the deadliest diseases around the globe. *Mtb* infection occurs when tubercle bacilli dispersed in the air from a patient with active pulmonary disease reaches the alveoli of the host. Here, they are quickly phagocytized by alveolar macrophages that can kill the entering

bacteria by innate immune response. If the bacilli survive this first line of defense, they start actively replicating in macrophages, diffuse to nearby cells including epithelial and endothelial cells, reaching in few weeks of exponential growth and a high bacterial burden[11]. During these early steps of infection, *Mtb* can diffuse to other organs through the lymphatics and by hematogenous dissemination where it can infect other cells and cause active disease. Thereafter, migration to the site of primary infection of neutrophils, lymphocytes and other immune cells form a cellular infiltrate that later assume a granuloma where it resides in a dormant, non-metabolically active state, for years, decades, or even for lifetime[11,12].

Pathogenic bacteria like *Mtb* are constantly exposed to a multitude of hostile conditions, with host defense systems and antibiotic treatments and with a continuous change in their environments. The potentially DNA-damaging assaults *in vivo*, are primarily from host-generated antimicrobial reactive oxygen intermediates (ROS) and reactive nitrogen intermediates (RNI)[13–15]. It is therefore important for bacteria to have an effective DNA damage repair and reversal mechanisms that can efficiently counteract the detrimental effects of these challenges. However, due to the technical difficulties in working with a slow-growing pathogen such as *Mtb*, the study of DNA repair systems in this organism has advanced more slowly than for other bacteria. *Mtb* has genes encoding proteins for various DNA repair pathways such as Nucleotide excision repair (NER), Base excision repair (BER) and double strand break repair (DSB). However, it's unknown how the proteins coordinate with each other and are activated to perform DNA repair.

One of the pathways of DNA repair is NER pathway which is responsible for removing a wide range of DNA lesions occurring in the single nucleotide strand, such as cyclobutane–pyrimidine dimers and 6–4 photoproducts induced by UV radiation[16,17]. The NER pathway in bacteria involves UvrA, UvrB, UvrC, and UvrD proteins. NER is initiated when a UvrA dimer

binds to a damaged DNA site (Figure 1.1). The lesion recognition is verified by binding of a UvrB dimer which triggers the release of UvrA from the DNA. UvrB remains tightly bound to the lesion and activates downstream repair. This includes recruitment of a UvrC endonuclease which nicks the ssDNA strand containing the lesion on both the 3' (~4-5 nucleotides away) and 5' (~8 nucleotides away) side. Following the incision, UvrD is recruited to the nick and thought to displace the ~12 nucleotides ssDNA containing the lesion. Missing bases are re-synthesized by DNA polymerase I, utilizing the undamaged complementary strand as a template. The DNA repair is completed by sealing two nicks in the phosphodiester backbone of DNA with the help of DNA ligase.

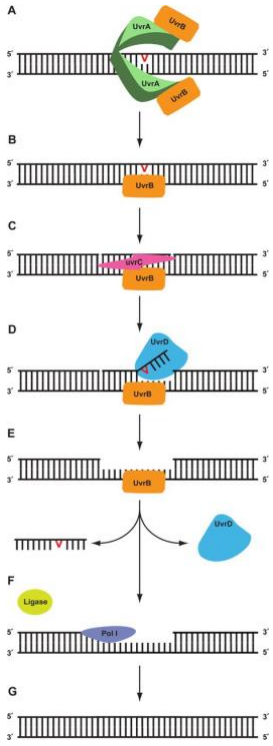


Figure 1.1: Schematic of nucleotide excision repair pathway where DNA damage in one of the nucleotide strands is recognized by UvrA dimer as part of UvrA2B2 tetramer complex. UvrB is then recruited by UvrA to scan the DNA damage. Once an altered base is found UvrB remains bound to the undamaged strand forming the UvrB-DNA pre-incision complex while UvrA2 is released. UvrB then recruits UvrC endonuclease that excise the lesion and the damaged strand is unwound and removed by helicase UvrD. The resulting gap is filled by DNA polymerase and joined by ligase[18,19].

DNA damage can also affect both strands and cause double strand breaks that are induced by endogenous sources such as reactive by-products of cellular metabolism, replication across nicks, and exogenous agents like ionizing radiation and genotoxic chemicals. One pathway of double strand break repair present in *Mtb* is non-homologous end-joining pathway (NHEJ) that uses two proteins Ku that binds to DNA at the sites of damage and recruits DNA ligase D (LigD) to seal linear plasmids and chromosomal DSB's (Figure 1.2)[20,21].

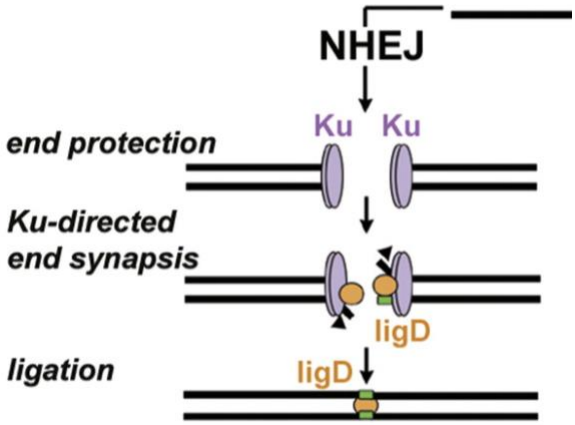


Figure 1.2: Schematic of bacterial non-homologous end joining pathway (NHEJ) where double strand breaks lead to Ku binding which recruits Ligase D that seals the breaks and repairs double strand breaks.

Although biochemical studies have shown the proteins involved in various repair pathways, we still don't understand how these proteins are activated both in the absence and presence of accessory factors to perform their function. The goal of my dissertation is to use *in-vitro* biophysical and biochemical approaches to characterize a helicase involved in DNA repair from *Mtb* called UvrD1. Helicases have been classified in 5 superfamilies (SF1-SF5). All proteins bind ATP and, consequently, all of them carry conserved residues that interact with DNA, the Walker A (phosphate-binding loop or P-loop) and Walker B (Mg^{2+}binding aspartic acid) residues. UvrD1 belongs to the family of UvrD like helicases belonging to the SF1 family of helicases that are identified by presence of seven conserved helicase sequence motifs and variable N and C-

terminal domains important for protein-protein interactions [22]. They have been shown to translocate on ssDNA as they appear to bind to the DNA substrate via hydrophobic interactions with the bases. They can also unwind dsDNA to ssDNA strands and play important roles in nucleic acid metabolism. UvrD like helicases have been well characterized from model organisms, including Escherichia coli (*E. coli*) and Bacillus subtilis (*B. subtilis*). Previous studies have shown that while monomers of UvrD family members (UvrD, Rep, and PcrA) are ssDNA translocases, dimeric forms of these enzymes are required to unwind duplex DNA in the absence of accessory factors or force[23–26]. In the following paragraphs, I describe biochemical and structural studies that describe the activation mechanism of UvrD like helicases.

1.1 Activation of UvrD-like helicases

UvrD like helicases belong to SF1 family of helicases[22]. Extensive biochemical and kinetic studies have tried to determine the active species of helicases. For the class of helicases belonging to hexameric ring structures family SF3 to SF6 it has been proposed that they encircle the DNA and thus link the helicase to the DNA substrate topologically[27]. However, among the non-hexameric DNA helicases, primarily in the SF1 and SF2 superfamilies, models comprising both functional monomeric and oligomeric structures have been proposed[28]. Monomeric helicase models have been proposed for the SF1 *Bacillus stearothermophilus* PcrA, and *E. coli* UvrD as well as the SF2 hepatitis C viral NS3 RNA helicase[29–31]. In contrast, models requiring dimerization have been proposed for the *E. coli* Rep, and UvrD helicases[32,33]. The monomers of the helicases have been shown to translocate in a particular direction either 3’-5’or 5’-3’ manner on ssDNA however oligomeric species mostly dimers are involved in unwinding. In the case of the Rep helicase, which shares around 40% sequence similarity with UvrD and PcrA, pre-steady-

state single turnover kinetic experiments, as well as single-molecule fluorescence studies indicate that Rep monomers are unable to initiate DNA unwinding, and that Rep oligomerization is required for DNA unwinding *in vitro*[34,35]. In the case of UvrD, pre-steady-state single turnover DNA-unwinding experiments performed at high concentrations of DNA substrate (100 nM) revealed a sigmoidal dependence of the extent of DNA unwinding on the concentration of UvrD, indicating that multiple UvrD monomers are required to bind the DNA substrate for optimal unwinding activity. These observations suggested that UvrD oligomerization is required for its helicase activity *in-vitro*[32]. In contrast with *E. coli* Rep, which remains monomeric in the absence of DNA, UvrD can self-assemble in the absence of DNA[36] The self-association properties of UvrD have been characterized using a range of solution conditions like salt, glycerol and temperature and have shown that UvrD can form dimers and tetramers over a range of solution conditions, although it changes to monomers upon increasing the concentration of NaCl and/or glycerol[36]. The stoichiometry of UvrD binding to a DNA substrate containing a 3' dT20 tail attached to an 18 bp duplex DNA was determined and in conditions when the total DNA concentration is at least twofold molar excess over the total UvrD concentration, the UvrD–DNA complex corresponds to a monomer of UvrD bound to the DNA. Therefore, if a UvrD monomer can function as a helicase *in-vitro*, one would expect to observe significant DNA unwinding in a single turnover helicase experiment performed under conditions of DNA substrate in excess over UvrD[37]. However, that was not the case. In addition, no DNA unwinding was observed for a tail length up to 10 nucleotides when only a monomeric species can bind to DNA[37]. Thus, above single turnover kinetic/single molecule studies established that both *E. coli* Rep and UvrD oligomers are the active unwinding species. For another helicase belonging to same family called PcrA it has been shown to translocate in 3' to 5' direction, rapidly and processively, but display no

detectable helicase activity under the same solution conditions *in vitro*[24]. Other studies have reported processive unwinding by PcrA by increasing DNA tail length or addition of accessory proteins suggesting the requirement of an oligomeric species for unwinding[38,39]. Studies on Hepatitis NS3 helicase which is SF2 have been shown that it is active as both monomer and oligomers, in conditions where substrate is saturating it functions as a monomer and is a unprocessive helicase while limiting substrate allows the formation of oligomers that is a processive helicase[40]. Similarly for bacteriophage T4 DdA helicase which is a SF1 helicase with 5'-3' directionality of translocation the processivity of unwinding increases with tail length suggesting oligomeric species to be responsible for unwinding[9,41].

Mtb has two UvrD like helicases UvrD1 and UvrD2, with UvrD1 being the SF1 family helicase with 39% sequence identity with *E. coli* UvrD, and 42% with Rep and conserved helicase residues called motifs and a C-terminal Tudor domain[42]. UvrD2 however has a different structure resembling RecQ an SF2 helicase with a C-terminal HRDC helicase and RNAseD-like C-terminal domain and an intervening tetra cysteine motif[43,44]. In case of *M. smegmatis* UvrD2 the HRDC domain is not required for helicase activity but uvrD2 and not uvrD1 has been shown to be essential in *M. smeg*[45]. One of the earlier studies that characterized *Mtb* UvrD1 showed that it is a SF1 family helicase with 3'-5' polarity and a DNA dependent ATPase. The size exclusion chromatography and sedimentation equilibrium performed at three different speeds in presence of reducing agent showed *Mtb* UvrD1 to be a monomer. Unwinding studies were performed using nicked, tailed, and forked DNA substrate in the absence of reducing agent show that it unwinds as a monomer[42]. In addition, its ability to unwind nicked and forked DNA substrates suggested its role in NER and clearance of stalled replication forks[42]. Another study published in the same year showed that *Mtb* UvrD1 is a monomeric ATP dependent translocase

but a weak DNA helicase and cannot unwind DNA. Ku an NHEJ protein stimulates UvrD1 to catalyze ATP-dependent unwinding of 3'-tailed DNAs[46]. In addition, UvrD1, Ku, and DNA form a stable ternary complex in the absence of ATP[46]. Ku was also shown to increase the helicase activity of UvrD2 deleted of tetracysteine module and the C-terminal domain[45]. Both *Mtb* UvrD1 and UvrD2 were also shown to be efficient in unwinding G4 structures derived from the *Mtb* genome and may suggest its role in the maintenance of genome integrity via G4 DNA resolution[47]. Since the solution conditions were different for purification and unwinding it was unclear what is the active species of *Mtb* UvrD1 and how does UvrD1 unwinds DNA.

1.2 Dimerization domain of UvrD-like helicases

UvrD like helicases are composed of two domains 1 and 2 which are further subdivided into subdomains 1A, 2A called the RecA like domains which bear the ATP binding site and conserved helicase motifs along with inserted domains 1B, 2B[7]. The structures of the Rep protein provided the first detailed view of a helicase bound to dT-15 ssDNA. These structures demonstrated that the 2B domain of Rep monomer can undergo large movements relative to the other three domains. This resulted in two different conformations closed and open in which the relative orientation of 2B domain appears different as it undergoes a 130° rotation about the hinge region[48]. This movement results in the closing of the large cleft formed in the open configuration of the Rep monomer by the four subdomains of the protein (Figure 3). In addition to closing off the cleft, this rotation partially buries the ssDNA binding groove between domains 1A and 1B and consequently buries the 3′ end of the ssDNA[48]. The presence of two structurally distinct conformations of Rep suggests that the functional asymmetries in the DNA binding and ATPase properties of the subunits of the Rep dimer may be reflected in between the open and closed forms

of the monomers observed in the crystal[48]. Furthermore, deleting or crosslinking the 2B domain made Rep an active helicase suggesting that it is not required for DNA unwinding or is auto-inhibitory *in-vitro*[34,49]. Single molecule studies on *E. coli* UvrD has showed that the 2B domain can exists in four distinct conformations[50]. Presence of high salt favors an open conformation and low salt favors a closed conformation. In addition, binding of the second UvrD monomer on DNA after the first or DNA induced dimerization favors a closed conformation[50]. This suggests a hypothesis that dimerization of UvrD may relieve the autoinhibition by 2B domain by changing its conformation to closed and may increase its helicase activity. However, it still proposes a hypothesis of presence of an asymmetric dimer as the first monomer that is bound to DNA is in open conformation and the second is closed. UvrD monomers can be activated for unwinding by accessory factor MutL and the mechanism is by rotation of 2B domain to a closed conformation[51]. MutL exists as a dimer in solution and one MutL dimer binds to *E. coli* UvrD monomer[4,51]. Since single-molecule experiment with FRET-labeled UvrD showed that the 2B subdomain of a UvrD monomer bound to a 3′-ssDNA duplex displays dynamic transitions between an open and closed state. Binding of a MutL dimer, the 2B subdomain assumes an intermediate state between open and closed conformations, and this state is on the pathway to forming the active helicase[51]. These results suggest that dimerization of UvrD or binding of an accessory factor to monomeric subunit both can lead to change in the conformation of UvrD like helicases with 2B domain into a more closed state. One of the models proposed for dimerization of UvrD like helicases through crystal studies from Rep was via 2B-2B domain as a dimer formed by same domain may limit the formation of higher order assembly states and the change in conformation of 2B domain relative to hinge region may create large movements that might be relevant for protein translocation. For instance, if Rep translocate on DNA by subunit switching mechanism a

model proposed for DNA unwinding, involving transient binding of both subunits, then one way to accommodate this is to couple DNA and/or ATP binding or hydrolysis to large relative movements of the two DNA binding sites within the dimer. Asymmetry observed in the two Rep monomeric units in the crystal structure is also explained by equilibrium and kinetic studies where negative cooperativity between two subunits shows evidence for a weak and a high affinity binding site for Rep is found between two subunits of Rep. Thus, subunit switching mechanism could explain the interconversion of a weak affinity to a higher affinity binding site and a conformational transition on DNA would then convert weak affinity site of Rep to a higher affinity site and vice versa.

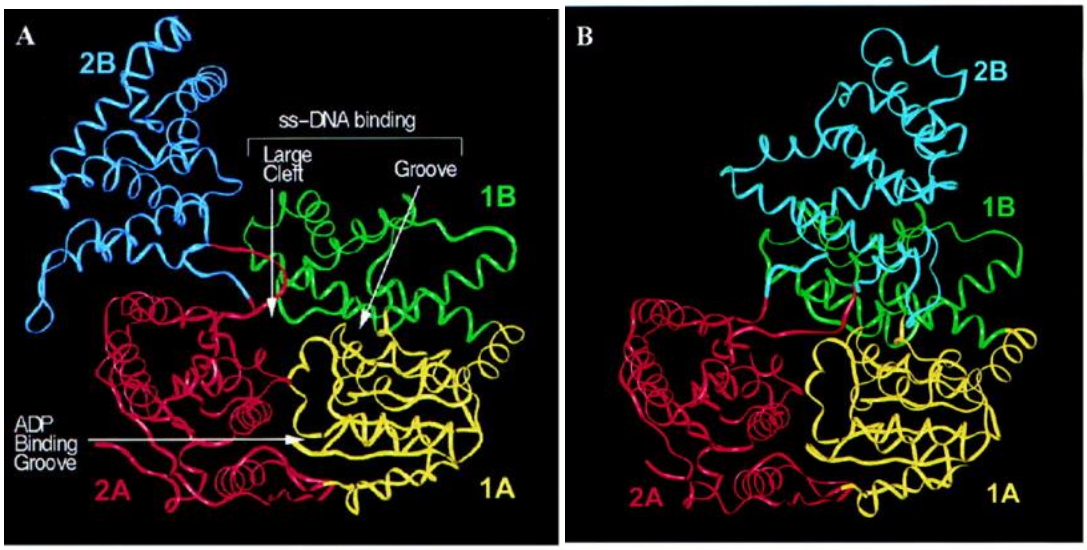


Figure 1.3: Open (left) and closed (right) conformations of *E. coli* Rep protein (PDB 1UAA).

1.3 Role of accessory factors in unwinding of UvrD-like helicases

UvrD-like helicase monomers and dimers can be activated for unwinding by means of accessory factors. *In-vitro* studies have shown that helicases can be activated by interacting with

accessory proteins in presence of DNA and are involved in important cellular processes like replication and repair. This activation has been shown for SF1 helicases like *E. coli* UvrD, PcrA, Rep and *Mtb* UvrD1. For instance, MutL is an accessory factor that increases the helicase activity of both a monomer and dimer of *E. coli* UvrD[52]. This stimulation of helicase activity requires MutL to interact with ssDNA. One MutL dimer was sufficient to increase the helicase activity of a monomer of UvrD and removing the C-terminal Tudor domain of UvrD did not affect the helicase activation of a monomer by MutL suggesting their interaction does not depend on Tudor domain[52]. Furthermore, single molecule FRET studies and kinetic analysis have shown that UvrD monomer bound to DNA has the 2B domain in open conformation but binding of MutL dimer to a UvrD monomer bound to DNA favors a closed conformation state of 2B domain[51]. This activation of UvrD monomer via rotation of 2B domain is specific for UvrD because replacing the 2B domain of UvrD with 2B domain of Rep did not stimulate its helicase activity[51]. This interaction of UvrD with MutL suggests its role in mismatch repair. Analogous studies on PcrA have shown that RepD a replication initiation protein stimulates the helicase activity and processivity of PcrA a helicase belonging to the family of UvrD family[53]. RepD was shown to load at the plasmid with sequence containing origin site where it covalently links with the phosphate group at the nick formed at the site of origin and then helps load the helicase PcrA. PcrA alone was shown to translocate processively but unwind poorly and only in presence of RepD dimer it is stimulated for directional unwinding[54]. The bulk studies were also confirmed using single molecule approaches where either PcrA or plasmid with origin replication sequences was bound to surface. In both conditions monomeric PcrA was unable to unwind DNA on its own but presence of RepD increased the processivity of unwinding[39]. This suggests the role of PcrA with RepD to be involved in resolution of stalled replication forks. Another protein called Rep

belonging to same family of helicases and to be important for replication restart at stalled replication forks along with other proteins such as PriA, B and C[55]. *In-vivo* studies showed that rep and priC are involved in same pathway of clearing stalled replication forks[56]. These results were later confirmed by *in-vitro* studies that show PriC-can stimulate the helicase activity of Rep monomer and dimer and increases its processivity on a forked DNA substrate [57]. One of the earlier studies published on *Mtb* UvrD1 showed that it is a strong ATP dependent translocase but a poor helicase[46]. Ku an accessory factor involved in double strand break repair stimulates the helicase activity of UvrD1[46]. This suggests the role of *Mtb* UvrD1 along with Ku in double strand break repair. Thus, all the studies reported above suggest that although the oligomeric form of helicases can function independently as a helicase the helicase activity of monomers/dimers can be further stimulated by means of accessory proteins.

1.4 Structural studies on UvrD like helicases

One of the first structures published to show a SF1 helicase bound to DNA was the crystal structure of Rep protein bound to dT-15 ssDNA[48]. Each monomeric subunit as for a conventional SF1 helicase is composed of 4 domains 1A,1B,2A and 2B. One of the crystallized structures show two asymmetric monomeric subunits bound to dT-16 ssDNA in presence of ADP. The two conformations of Rep were called as the “open” and “closed” conformations, since a cleft within which ssDNA binds is open in the open form but closed in the other[48]. The relative orientation of the domain 2B is different in both conformations. In closed conformation the motion in 2B domain results in the closing of the large cleft formed in the open configuration of the Rep monomer by the four subdomains of the protein. This structure suggests that the asymmetry seen in two rep monomers bound to DNA might be functionally relevant and the two monomers with

different conformation might have different affinity for the DNA substrate. This is also reflected in data showing negative cooperativity in binding of DNA to the second subunit of the dimer[58]. Since binding of ADP favors both subunits to bind to ssDNA and ATP instead favors binding to the dsDNA an active rolling subunit switching model was proposed for DNA unwinding. In this model one subunit of the dimer binds to ssDNA and presence of ATP increases the affinity for the second unligated subunit for binding to the dsDNA. The unwinding because of translocation and ATP hydrolysis then gives rise to an intermediate with ADP where both subunits are bound to ssDNA and the release of ADP and phosphate that leads to dissociation of initially bound monomeric subunit from ssDNA. Thus, both subunits alternate between binding to ss and dsDNA[48]. The crystal structure of another helicase PcrA belonging to the same family of helicases was determined bound to 10 base-pair DNA with 7 nucleotide tail. The structure obtained bound to AMPPNP (non-hydrolysable ATP analog) was called the substrate and sulphate ion was called the product complex[31]. Since the sulphate ion occupies the same place as the phosphate this is rereferred as the product complex. The structures show that ssDNA binds across the surface of the two RecA-like domains (1A and 2A). The domains 1A and 2A rotate toward each other when bound to ATP and open after ATP hydrolysis, and with this movement of domains PcrA translocate along the ssDNA as a monomer like an inchworm with alternating tight and loose interactions[31]. The crystal structures of *E. coli* UvrD binary complexes of UvrD-DNA and ternary with UvrD-DNA-AMPPNP and UvrD-DNA-ADP.MgF3 with 18bp-dT7 tail were determined that show a monomer of UvrD bound to DNA[59]. The ATP analog, AMPPNP, binds at the cleft separating the 1A and 2A sub-domains and interacts with all the seven conserved helicase. The duplex and single-stranded regions of a 3'-ssDNA/duplex show a ~90° bend in a UvrD monomer-DNA complex[59]. The 3'-ssDNA tail binds across the 1A and 2A sub-domains

with a 3’ to 5’ orientation along the 1A and 2A sub-domains, respectively. The ssDNA interacts with the conserved residues located in between the RecA domains. The 2B sub-domain in a closed conformation contacts the duplex region of DNA through a Glycine-isoleucine-glycine (GIG) motif which was proposed to facilitate DNA unwinding (Figure 1.4) [59].

The crystal structures of both PcrA and UvrD show monomeric subunits bound to DNA. However, whether this UvrD-DNA structure reflects a functional complex has been questioned. First a UvrD/PcrA monomer alone cannot unwind a 18bp duplex DNA with dT7 tail on which it was crystallized[36,37]. Second, deletion of the 2B sub-domain of the *E. coli* Rep monomer activates its latent helicase activity indicating that the 2B sub-domain rather plays a regulatory role[7,34].

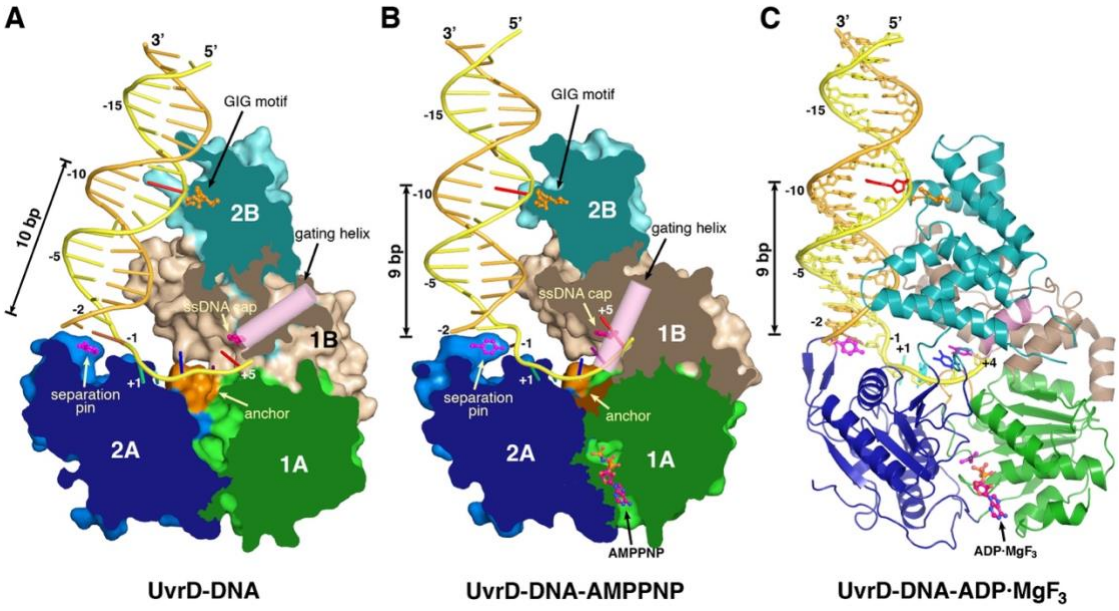


Figure 1.4: Binary and ternary complexes of *E. coli* UvrD bound to DNA that show the interaction of GIG motif in 2B domain with dsDNA (PDB 2IS1).

Structural studies on *E. coli* UvrD have also reported a crystal structure of the apo form of UvrD that has its 2B sub-domain in an open conformation indicating that the 2B sub-domain of UvrD, just as for Rep and PcrA, can undergo a similar large rotational conformational change[60].

Furthermore, the rotational conformational state of the 2B sub-domain of UvrD can be influenced by binding of ligands (e.g., nucleotides, DNA) and changes in solution conditions such that low salt concentration favors the closed conformation and high salt concentration favoring the open conformation in the absence of DNA. The fact that rotational motion of the 2B sub-domain is coupled to the binding of nucleotides and DNA suggests that it is likely to be functionally important for some activities of this enzyme and their regulation (Figure1.5).

In the following chapters, I will characterize the helicase from *Mtb* called UvrD1 that is 40% identical to *E. coli* UvrD and Rep proteins. I will also show how it exists as a monomer and dimer populations as a function of redox. The dimeric form of the protein is an active helicase but the monomer of UvrD1 can be activated by an accessory protein called Ku. This suggests the role of UvrD1 in multiple DNA repair pathways by different activation mechanisms.

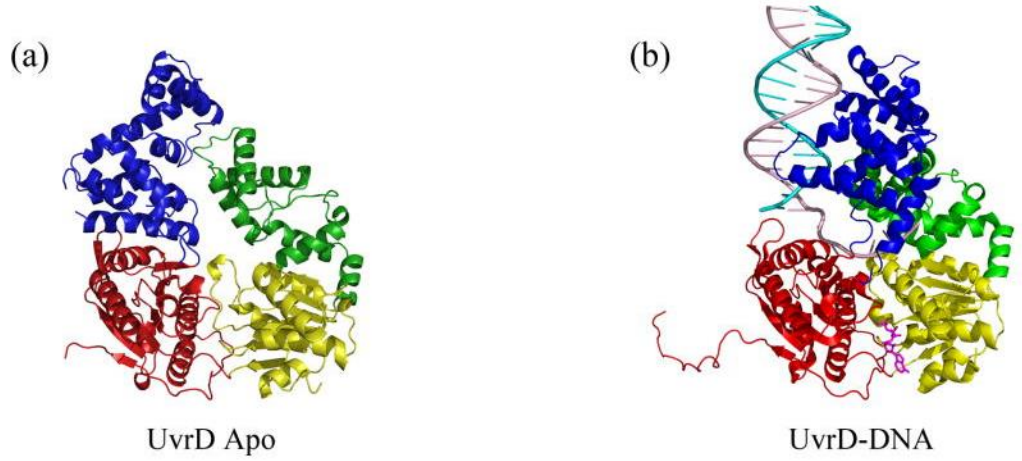


Figure 1.5: UvrD in apo form with domains 1A, 1B, 2A and 2B with yellow, green, red and blue colors. The 2B domain is in open conformation in 5a and closed where the 2B domain undergoes a rotation by 160° (PDB 3LFU).

1.5 References

1. S.W. Matson, D.W. Bean, J.W. George, DNA helicases: Enzymes with essential roles in all aspects of DNA metabolism, Bioessays. 16 (1994) 13–22. https://doi.org/10.1002/bies.950160103.

2. M. O'Donnell, L. Langston, B. Stillman, Principles and Concepts of DNA Replication in Bacteria, Archaea, and Eukarya, Csh Perspect Biol. 5 (2013) a010108. https://doi.org/10.1101/cshperspect.a010108.

3. I.O. Petruseva, A.N. Evdokimov, O.I. Lavrik, Molecular mechanism of global genome nucleotide excision repair., Acta Naturae. 6 (2014) 23–34.

4. L. Worth, S. Clark, M. Radman, P. Modrich, Mismatch repair proteins MutS and MutL inhibit RecA-catalyzed strand transfer between diverged DNAs., Proc National Acad Sci. 91 (1994) 3238–3241. https://doi.org/10.1073/pnas.91.8.3238.

5. A. Sancar, Structure and function of DNA photolyase, Biochemistry-Us. 33 (1994) 2–9. https://doi.org/10.1021/bi00167a001.

6. T.M. Lohman, K.P. Bjornson, Mechanisms of Helicase-Catalyzed DNA Unwinding, Annu Rev Biochem. 65 (1996) 169–214. https://doi.org/10.1146/annurev.bi.65.070196.001125.

7. T.M. Lohman, E.J. Tomko, C.G. Wu, Non-hexameric DNA helicases and translocases: mechanisms and regulation, Nat Rev Mol Cell Bio. 9 (2008) 391–401. https://doi.org/10.1038/nrm2394.

8. E.J. Enemark, L. Joshua-Tor, Mechanism of DNA translocation in a replicative hexameric helicase, Nature. 442 (2006) 270–275. https://doi.org/10.1038/nature04943.

9. A.K. Byrd, D.L. Matlock, D. Bagchi, S. Aarattuthodiyil, D. Harrison, V. Croquette, K.D. Raney, Dda Helicase Tightly Couples Translocation on Single-Stranded DNA to Unwinding of Duplex DNA: Dda Is an Optimally Active Helicase, J Mol Biol. 420 (2012) 141–154. https://doi.org/10.1016/j.jmb.2012.04.007.

10. J. Chakaya, M. Khan, F. Ntoumi, E. Aklillu, R. Fatima, P. Mwaba, N. Kapata, S. Mfinanga, S.E. Hasnain, P.D.M.C. Katoto, A.N.H. Bulabula, N.A. Sam-Agudu, J.B. Nachega, S. Tiberi, T.D. McHugh, I. Abubakar, A. Zumla, Global Tuberculosis Report 2020 – Reflections on the Global TB burden, treatment and prevention efforts, Int J Infect Dis. 113 (2021) S7–S12. https://doi.org/10.1016/j.ijid.2021.02.107.

11. A.J. Wolf, L. Desvignes, B. Linas, N. Banaiee, T. Tamura, K. Takatsu, J.D. Ernst, Initiation of the adaptive immune response to Mycobacterium tuberculosis depends on antigen production in the local lymph node, not the lungs, J Exp Medicine. 205 (2008) 105–115. https://doi.org/10.1084/jem.20071367.

12. A.J. Lenaerts, D. Hoff, S. Aly, S. Ehlers, K. Andries, L. Cantarero, I.M. Orme, R.J. Basaraba, Location of Persisting Mycobacteria in a Guinea Pig Model of Tuberculosis Revealed by R207910, Antimicrob Agents Ch. 51 (2007) 3338–3345. https://doi.org/10.1128/aac.00276-07.

13. J.D. MacMicking, R.J. North, R. LaCourse, J.S. Mudgett, S.K. Shah, C.F. Nathan, Identification of nitric oxide synthase as a protective locus against tuberculosis, Proc National Acad Sci. 94 (1997) 5243–5248. https://doi.org/10.1073/pnas.94.10.5243.

14. E.A. Rich, M. Torres, E. Sada, C.K. Finegan, B.D. Hamilton, Z. Toossi, Mycobacterium tuberculosis (MTB)- stimulated production of nitric oxide by human alveolar macrophages and relationship of nitric oxide production to growth inhibition of MTB, Tubercle Lung Dis. 78 (1997) 247–255. https://doi.org/10.1016/s0962-8479(97)90005-8.

15. T. Akaki, H. Tomioka, T. Shimizu, S. Dekio, K. Sato, Comparative roles of free fatty acids with reactive nitrogen intermediates and reactive oxygen intermediates in expression of the anti-microbial activity of macrophages against Mycobacterium tuberculosis, Clin Exp Immunol. 121 (2000) 302–310. https://doi.org/10.1046/j.1365-2249.2000.01298.x.

16. L. Grossman, S. Thiagalingam, Nucleotide excision repair, a tracking mechanism in search of damage, J Biol Chem. 268 (1993) 16871–16874. https://doi.org/10.1016/s0021-9258(19)85273-0.

17. J.T. Reardon, A. Sancar, Nucleotide Excision Repair, Prog Nucleic Acid Re. 79 (2005) 183–235. https://doi.org/10.1016/s0079-6603(04)79004-2.

18. C. Kisker, J. Kuper, B.V. Houten, Prokaryotic Nucleotide Excision Repair, Csh Perspect Biol. 5 (2013) a012591. https://doi.org/10.1101/cshperspect.a012591.

19. B.V. Houten, N. Kad, Investigation of bacterial nucleotide excision repair using single-molecule techniques, Dna Repair. 20 (2014) 41–48. https://doi.org/10.1016/j.dnarep.2013.10.012.

20. C. Gong, P. Bongiorno, A. Martins, N.C. Stephanou, H. Zhu, S. Shuman, M.S. Glickman, Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C, Nat Struct Mol Biol. 12 (2005) 304–312. https://doi.org/10.1038/nsmb915.

21. R. Gupta, M. Ryzhikov, O. Koroleva, M. Unciuleac, S. Shuman, S. Korolev, M.S. Glickman, A dual role for mycobacterial RecO in RecA-dependent homologous recombination and RecA-independent single-strand annealing, Nucleic Acids Res. 41 (2013) 2284–2295. https://doi.org/10.1093/nar/gks1298.

22. M.E. Fairman-Williams, U.-P. Guenther, E. Jankowsky, SF1 and SF2 helicases: family matters, Curr Opin Struc Biol. 20 (2010) 313–324. https://doi.org/10.1016/j.sbi.2010.03.011.

23. M.S. Dillingham, D.B. Wigley, M.R. Webb, Direct Measurement of Single-Stranded DNA Translocation by PcrA Helicase Using the Fluorescent Base Analogue 2-Aminopurine †, Biochemistry-Us. 41 (2002) 643–651. https://doi.org/10.1021/bi011137k.

24. A. Niedziela-Majka, M.A. Chesnik, E.J. Tomko, T.M. Lohman, Bacillus stearothermophilus PcrA Monomer Is a Single-stranded DNA Translocase but Not a Processive Helicase in Vitro *, J Biol Chem. 282 (2007) 27076–27085. https://doi.org/10.1074/jbc.m704399200.

25. E.J. Tomko, C.J. Fischer, A. Niedziela-Majka, T.M. Lohman, A Nonuniform Stepping Mechanism for E. coli UvrD Monomer Translocation along Single-Stranded DNA, Mol Cell. 26 (2007) 335–347. https://doi.org/10.1016/j.molcel.2007.03.024.

26. K.S. Lee, H. Balci, H. Jia, T.M. Lohman, T. Ha, Direct imaging of single UvrD helicase dynamics on long single-stranded DNA, Nat Commun. 4 (2013) 1878. https://doi.org/10.1038/ncomms2882.

27. S.S. Patel, K.M. Picha, STRUCTURE AND FUNCTION OF HEXAMERIC HELICASES1, Annu Rev Biochem. 69 (2000) 651–697. https://doi.org/10.1146/annurev.biochem.69.1.651.

28. A.E. Gorbalenya, E.V. Koonin, Helicases: amino acid sequence comparisons and structure-function relationships, Curr Opin Struc Biol. 3 (1993) 419–429. https://doi.org/10.1016/s0959-440x(05)80116-2.

29. J.L. Kim, K.A. Morgenstern, J.P. Griffith, M.D. Dwyer, J.A. Thomson, M.A. Murcko, C. Lin, P.R. Caron, Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding, Structure. 6 (1998) 89–100. https://doi.org/10.1016/s0969-2126(98)00010-0.

30. L.E. Mechanic, M.C. Hall, S.W. Matson, Escherichia coli DNA Helicase II Is Active as a Monomer*, J Biol Chem. 274 (1999) 12488–12498. https://doi.org/10.1074/jbc.274.18.12488.

31. S.S. Velankar, P. Soultanas, M.S. Dillingham, H.S. Subramanya, D.B. Wigley, Crystal Structures of Complexes of PcrA DNA Helicase with a DNA Substrate Indicate an Inchworm Mechanism, Cell. 97 (1999) 75–84. https://doi.org/10.1016/s0092-8674(00)80716-3.

32. J.A. Ali, T.M. Lohman, Kinetic Measurement of the Step Size of DNA Unwinding by Escherichia coli UvrD Helicase, Science. 275 (1997) 377–380. https://doi.org/10.1126/science.275.5298.377.

33. I. Wong, T.M. Lohman, Allosteric Effects of Nucleotide Cofactors on Escherichia coli Rep Helicase&DNA Binding, Science. 256 (1992) 350–355. https://doi.org/10.1126/science.256.5055.350.

34. W. Cheng, K.M. Brendza, G.H. Gauss, S. Korolev, G. Waksman, T.M. Lohman, The 2B domain of the Escherichia coli Rep protein is not required for DNA helicase activity, Proc National Acad Sci. 99 (2002) 16006–16011. https://doi.org/10.1073/pnas.242479399.

35. T. Ha, I. Rasnik, W. Cheng, H.P. Babcock, G.H. Gauss, T.M. Lohman, S. Chu, Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase, Nature. 419 (2002) 638–641. https://doi.org/10.1038/nature01083.

36. N.K. Maluf, T.M. Lohman, Self-association Equilibria of Escherichia coli UvrD Helicase Studied by Analytical Ultracentrifugation, J Mol Biol. 325 (2003) 889–912. https://doi.org/10.1016/s0022-2836(02)01276-7.

37. N.K. Maluf, C.J. Fischer, T.M. Lohman, A Dimer of Escherichia coli UvrD is the Active Form of the Helicase In Vitro, J Mol Biol. 325 (2003) 913–935. https://doi.org/10.1016/s0022-2836(02)01277-9.

38. Y. Yang, S.-X. Dou, H. Ren, P.-Y. Wang, X.-D. Zhang, M. Qian, B.-Y. Pan, X.G. Xi, Evidence for a functional dimeric form of the PcrA helicase in DNA unwinding, Nucleic Acids Res. 36 (2008) 1976–1989. https://doi.org/10.1093/nar/gkm1174.

39. L.T. Chisty, C.P. Toseland, N. Fili, G.I. Mashanov, M.S. Dillingham, J.E. Molloy, M.R. Webb, Monomeric PcrA helicase processively unwinds plasmid lengths of DNA in the presence of the initiator protein RepD, Nucleic Acids Res. 41 (2013) 5010–5023. https://doi.org/10.1093/nar/gkt194.

40. T.A. Jennings, S.G. Mackintosh, M.K. Harrison, D. Sikora, B. Sikora, B. Dave, A.J. Tackett, C.E. Cameron, K.D. Raney, NS3 Helicase from the Hepatitis C Virus Can Function as a Monomer or Oligomer Depending on Enzyme and Substrate Concentrations*, J Biol Chem. 284 (2009) 4806–4814. https://doi.org/10.1074/jbc.m805540200.

41. A.K. Byrd, K.D. Raney, Increasing the Length of the Single-Stranded Overhang Enhances Unwinding of Duplex DNA by Bacteriophage T4 Dda Helicase †, Biochemistry-Us. 44 (2005) 12990–12997. https://doi.org/10.1021/bi050703z.

42. E. Curti, S.J. Smerdon, E.O. Davis, Characterization of the Helicase Activity and Substrate Specificity of Mycobacterium tuberculosis UvrD, J Bacteriol. 189 (2007) 1542–1555. https://doi.org/10.1128/jb.01421-06.

43. A. Williams, C. Güthlein, N. Beresford, E.C. Böttger, B. Springer, E.O. Davis, UvrD2 Is Essential in Mycobacterium tuberculosis, but Its Helicase Activity Is Not Required, J Bacteriol. 193 (2011) 4487–4494. https://doi.org/10.1128/jb.00302-11.

44. M.P. Killoran, J.L. Keck, Structure and function of the regulatory C-terminal HRDC domain from Deinococcus radiodurans RecQ, Nucleic Acids Res. 36 (2008) 3139–3149. https://doi.org/10.1093/nar/gkn143.

45. K.M. Sinha, N.C. Stephanou, M.-C. Unciuleac, M.S. Glickman, S. Shuman, Domain Requirements for DNA Unwinding by Mycobacterial UvrD2, an Essential DNA Helicase, Biochemistry-Us. 47 (2008) 9355–9364. https://doi.org/10.1021/bi800725q.

46. K.M. Sinha, N.C. Stephanou, F. Gao, M.S. Glickman, S. Shuman, Mycobacterial UvrD1 Is a Ku-dependent DNA Helicase That Plays a Role in Multiple DNA Repair Events, Including Double-strand Break Repair*, J Biol Chem. 282 (2007) 15114–15125. https://doi.org/10.1074/jbc.m701167200.

47. T. Saha, K. Shukla, R.S. Thakur, A. Desingu, G. Nagaraju, Mycobacterium tuberculosis UvrD1 and UvrD2 helicases unwind G-quadruplex DNA, Febs J. 286 (2019) 2062–2086. https://doi.org/10.1111/febs.14798.

48. S. Korolev, J. Hsieh, G.H. Gauss, T.M. Lohman, G. Waksman, Major Domain Swiveling Revealed by the Crystal Structures of Complexes of E. coli Rep Helicase Bound to Single-Stranded DNA and ADP, Cell. 90 (1997) 635–647. https://doi.org/10.1016/s0092-8674(00)80525-5.

49. S. Arslan, R. Khafizov, C.D. Thomas, Y.R. Chemla, T. Ha, Engineering of a superhelicase through conformational control, Science. 348 (2015) 344–347. https://doi.org/10.1126/science.aaa0445.

50. B. Nguyen, Y. Ordabayev, J.E. Sokoloski, E. Weiland, T.M. Lohman, Large domain movements upon UvrD dimerization and helicase activation, Proc National Acad Sci. 114 (2017) 12178–12183. https://doi.org/10.1073/pnas.1712882114.

51. Y.A. Ordabayev, B. Nguyen, A.G. Kozlov, H. Jia, T.M. Lohman, UvrD helicase activation by MutL involves rotation of its 2B subdomain, Proc National Acad Sci. 116 (2019) 16320–16325. https://doi.org/10.1073/pnas.1905513116.

52. Y.A. Ordabayev, B. Nguyen, A. Niedziela-Majka, T.M. Lohman, Regulation of UvrD Helicase Activity by MutL, J Mol Biol. 430 (2018) 4260–4274. https://doi.org/10.1016/j.jmb.2018.08.022.

53. P. Soultanas, M.S. Dillingham, D.B. Wigley, F. Papadopoulos, S.E.V. Phillips, C.D. Thomas, Plasmid replication initiator protein RepD increases the processivity of PcrA DNA helicase, Nucleic Acids Res. 27 (1999) 1421–1428. https://doi.org/10.1093/nar/27.6.1421.

54. W. Zhang, M.S. Dillingham, C.D. Thomas, S. Allen, C.J. Roberts, P. Soultanas, Directional Loading and Stimulation of PcrA Helicase by the Replication Initiator Protein RepD, J Mol Biol. 371 (2007) 336–348. https://doi.org/10.1016/j.jmb.2007.05.050.

55. S.J. Sandler, J.D. McCool, T.T. Do, R.U. Johansen, PriA mutations that affect PriA–PriC function during replication restart, Mol Microbiol. 41 (2001) 697–704. https://doi.org/10.1046/j.1365-2958.2001.02547.x.

56. S.J. Sandler, Multiple Genetic Pathways for Restarting DNA Replication Forks in Escherichia coli K-12, Genetics. 155 (2000) 487–497. https://doi.org/10.1093/genetics/155.2.487.

57. B. Nguyen, M.K. Shinn, E. Weiland, T.M. Lohman, Regulation of E. coli Rep helicase activity by PriC, J Mol Biol. 433 (2021) 167072. https://doi.org/10.1016/j.jmb.2021.167072.

58. K.P. Bjornson, I. Wong, T.M. Lohman, ATP Hydrolysis Stimulates Binding and Release of Single Stranded DNA from Alternating Subunits of the DimericE. coliRep Helicase: Implications for ATP-driven Helicase Translocation, J Mol Biol. 263 (1996) 411–422. https://doi.org/10.1006/jmbi.1996.0585.

59. J.Y. Lee, W. Yang, UvrD Helicase Unwinds DNA One Base Pair at a Time by a Two-Part Power Stroke, Cell. 127 (2006) 1349–1360. https://doi.org/10.1016/j.cell.2006.10.049.

60. H. Jia, S. Korolev, A. Niedziela-Majka, N.K. Maluf, G.H. Gauss, S. Myong, T. Ha, G. Waksman, T.M. Lohman, Rotations of the 2B Sub-domain of E. coli UvrD Helicase/Translocase Coupled to Nucleotide and DNA Binding, J Mol Biol. 411 (2011) 633–648. https://doi.org/10.1016/j.jmb.2011.06.019.

Chapter 2: DNA Damage Detected Through Direct, Indirect, and Inference Mechanisms

2.1 Abstract

Although DNA damage underlies the generation of mutations that serve to generate diversity required for evolution, it is also a constant threat to genomic stability. Many different pathways, often conserved throughout biology, have evolved to identify lesions in the DNA and recruit the appropriate proteins needed for repair. Here we address overarching strategies that have evolved from a bird's eye view and provide examples of each. Specifically, we divide these strategies into three mechanisms we call direct, indirect, and inferred. Direct mechanisms depend on the physical binding of a factor to the lesion. For example, UvrA in bacterial NER. Indirect mechanisms depend instead on the interruption of a normal process and include transcription coupled and replication coupled repair. Finally, we describe a mechanism where no DNA lesion is detected at all. Instead, DNA damage is inferred via the sensing of conditions in the cell permissive for DNA damage such as the presence of reactive oxygen species.

2.2 Introduction to DNA repair

Life depends on the chemical pliability of biological molecules and changes in molecular structures allow for information to be integrated and transmitted via biological signaling pathways. However, this also means that these molecular components are susceptible to reactions that interfere with their optimal function as dictated by their evolved roles. Selective pressure to utilize this malleability while limiting the deleterious effects of undesirable chemical reactions has led to the existence of a suite of control processes constantly surveying, repairing, or recycling molecules

to maintain this delicate balance. These quality control mechanisms exist for nearly all biological molecules. Here, we focus on those responsible for protecting the genetic information known broadly as DNA repair mechanisms[1,2].

DNA serves as the physical storehouse of genetic information in the form of a linear sequence of four nucleobases. Its two-stranded double-helical structure provides the ability to divide the strands and make two identical copies, providing the mechanism for replication and the passage of this information during development and reproduction. In addition, the double helicity represents a back-up of genetic information in case the information in one strand is compromised. In fact, this information is constantly being compromised and a single human cell must deal with modified or damaged information tens of thousands of times a day[3].

Certain modifications to DNA structure have been co-opted by evolution and are enacted enzymatically to mark the genome both globally and locally for the purposes of self-recognition or gene regulation. For example, DNA methylation allows for the identification of the parental strand in replication and of viral DNA in immune responses[4,5]. Local modifications can also interface with epigenetic mechanisms to determine gene expression levels (i.e., 5'-methylcytosine and 5'-hydroxymethylcystosine)[6,7]. However, when modifications are randomly caused throughout the genome by unregulated chemical reactions, they are referred to as DNA damage.

DNA damage can be caused by a variety of processes and take many forms. The simplest example may be the hydrolysis of the glycosidic bond that links a nucleobase to the sugar ring resulting in an empty position within the linear sequence or an abasic site[8,9]. Furthermore, oxidation reactions can result in over 20 kinds of chemically modified bases[10,11]. In addition, UV light can indirectly and directly cause damage through the generation of reactive chemical species and the stimulation of base-base crosslinks[11,12].

The consequences of damage can be broken down into two categories. The first is mutation. A damaged site may lead to misinterpretation of, or a permanent change in, the genetic information. Very few of these changes will be advantageous to the organism, some will have no immediate consequence, and others will be disadvantageous and may lead to disease states. The second category involves the effects of DNA lesions which impede normal nucleic acid metabolism. The processes of replication and transcription require enzyme motors to translocate processively across large distances. Certain lesions result in a significant change in the conformation of the DNA such that these motors cannot pass through and result in the stalling of these central processes. Due to the essentiality of replication and transcription, cells must either pause and repair the damage or die.

The plethora of DNA damage types have precipitated an equal variety in DNA repair mechanisms. While there is some crossover, these mechanisms are classically grouped according to the kinds of damage that they sense and repair. Base excision repair (BER) attends to small modifications to nucleobases that can be remedied simply by replacing the base itself[8,9]. Mismatch repair (MMR) recognizes non-complementary base pairs within duplex DNA and repairs the newer strand to preserve the genetic code[13]. Nucleotide excision repair (NER) is sensitive to larger lesions where the shape of the double-helix itself is compromised to the point that replication and transcription are affected[14,15]. Double-strand break repair (DSBR) responds to situations where both strands have been cut requiring either repair by homologous recombination or via non-homologous end-joining[16,17].

Here, instead of dividing repair mechanisms by the types of lesions they repair, we categorize the sensing mechanism itself and define three strategies: direct, indirect, and inferred. Direct mechanisms are those where a dedicated repair factor makes molecular contact with a

lesion. Indirect mechanisms are those where lesions are detected by virtue of their interruption of normal cellular processes. Lastly, inferred mechanisms are not dependent on the existence of a specific lesion at all. Instead, general cellular conditions under which DNA lesions are likely to be generated, such as the presence of reactive oxygen species, are detected and enzymatic activities are regulated so as to best support repair.

2.3 Direct Mechanisms

DNA damage fundamentally results in the modification of the structure of the DNA. This can be as subtle as the introduction of a single atom or as dramatic as a double-stranded break. Direct mechanisms describe those pathways whereby the new structural feature is recognized by the binding of a factor which then recruits repair machinery (Fig. 1). We illustrate this with examples taken from double-strand break repair (DSBR), base-excision repair (BER), and mismatch repair (MMR).

Double-strand break repair

DNA damage due to exogenous agents such as ultraviolet light, mutagens, and free radicals, or endogenous processes such as replication and transcription can cause double-strand breaks (DSB)[18]. If left unrepaired, DSBs result in loss of genetic information, chromosomal aberrations, or cell death[16]. There are two main mechanisms for repairing DSB, homologous recombination (HR) or nonhomologous end-joining (NHEJ)[17,19]. In many organisms, DSB repair is initiated via binding of Ku proteins that recognize the break directly. This sensor is evolutionarily conserved from bacteria to eukaryotes although it is not found in all bacteria[20]. Ku exists as a homodimer in prokaryotes and a heterodimer in eukaryotes formed of two subunits: Ku70 and Ku80[21]. Specifically, the Ku dimer forms a ring structure that must be threaded onto a

free DNA end[22], allowing it to distinguish broken from unbroken DNA and sense this type of damage.

Base excision repair

Damage to individual bases is the most common type of DNA damage in humans[8]. BER responds to lesions that alter the DNA base but do not significantly distort the DNA helix. These lesions can arise from endogenous processes such as deamination, oxidation, or methylation[8]. In the BER pathway, DNA glycosylases are responsible for the direct detection of DNA damage. There are many flavors of these enzymes with varying specificity to a certain type of modification or damage. The glycosylase activity cleaves the glycolytic bond between the sugar and the base and results in the removal of the damaged base and the generation of an abasic site[8]. Examples include the eukaryotic hOGG1 and prokaryotic MutM, which are responsible for removing 8-oxo-guanine or AlkD[9,23]. While each glycosylase possesses subtle variations, they share basic characteristics that allow them to directly identify their respective targets. First, they efficiently scan the genome via 1D diffusion[24]. Second, most importantly perhaps is that target recognition and catalysis typically require the DNA substrate to adopt a distorted structure. Since oxidized and alkylated bases do not fit into the canonical DNA helix and thus exhibit lower base-pair stability, they can be recognized via their propensity to adopt such distorted structures. In these structures, the helical axis is kinked, and the damaged base is flipped out of the helix[25]. Subsequently, extrahelical bases can be further identified through structural compatibility with the enzyme active site and neighboring binding pockets[24].

Mismatch repair

The mismatch repair pathway is conserved from bacteria to humans, corrects base substitution and insertion-deletion mismatches (IDLs) generated during replication[13,26], and

depend on lesion sensing by a MutS dimer[27]. After binding DNA, the MutS dimer forms a theta-shaped structure with upper and lower channels where the DNA residues are scanned for mismatch in the lower channel[28]. The MutS dimer binds to DNA and bends it using energy from nonspecific protein-DNA interactions[28,29]. This facilitates DNA bending to search for mismatch sites in DNA. The presence of a mismatch is detected by a conserved glutamate residue[29] and leads to the formation of a kinked DNA recognition complex. Subsequent nucleotide exchange results in MutS sliding and the recruitment of downstream proteins for repair[28,30].

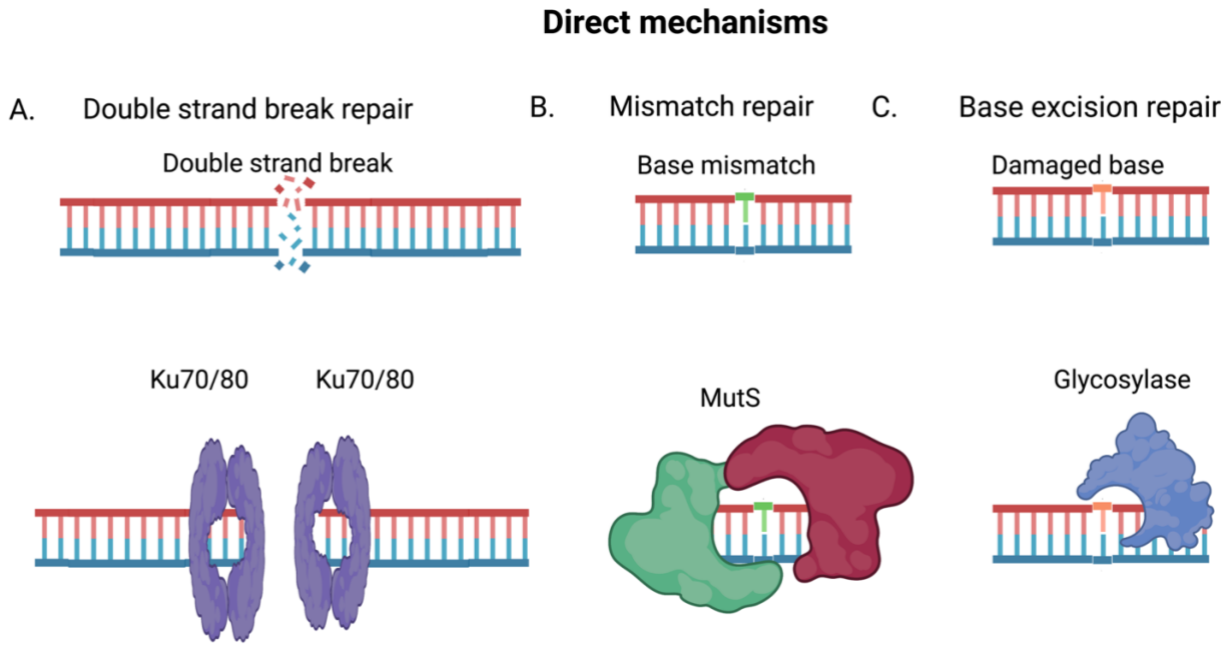


Figure 2.1: Examples of direct mechanisms where factors bind specifically to DNA lesions. (A) Ku70/Ku80 heterodimer recognizes double-stranded DNA ends. (B) MutS dimer identifies base mismatches. (C) DNA glycosylases recognize a damaged base.

2.4 Indirect Mechanisms

The processes of genome replication and expression utilize highly precise and processive molecular motors that translocate along the DNA. Pauses in the continuous motion of these motors occur stochastically, are biased by sequence[31–33], and are stimulated by template-bound barriers[34] and the incorporation of incorrect nucleotides[35,36]. The later fact underlies proofreading mechanisms that increase fidelity by orders of magnitude. However, even in the absence of proof-reading mechanisms, these enzymes possess error rates of 10^{-3}-10^{-6} primarily due to the exquisite structural fit of correct base-pairs in the active sites of these enzymes[36,37]. As DNA damage can introduce even larger structural distortions in the template, they often prevent the advancement of these motors leading to a molecular signal of their presence.

Transcription-coupled DNA repair

Transcription coupled repair (TCR) is one example of an indirect mechanism. Here, lesions are sensed by the inability of RNA polymerase (RNAP) to transcribe through them. Originally, this mechanism was revealed by the observation that certain types of lesions in transcribed strands (i.e., the strand being read by RNAP) were repaired more rapidly than those in other regions of the genome or even than those found on the complementary non-template strands[38,39]. As RNAP can bypass lesions small enough to pass through the active site such as abasic sites or mismatches, BER and MMR are not stimulated through this mechanism. In contrast, bulky lesions repaired by nucleotide excision repair (NER) do result in RNAP stalling. As such, TCR can be thought an NER-specific pathway. Interestingly, our understanding of the proportion of NER-repaired lesions detected indirectly in this manner as opposed to that detected directly may be changing[40,41].

In indirect mechanisms, the factor “recognizing” the lesion is not a part of a DNA repair pathway *per se*. In the case of TCR, the stalled RNAP serves as a signal of damage, but also

obscures the lesion within its active site and prevents future rounds of transcription through the site, blocking gene expression. This complex must be remodeled to both repair the lesion and allow for transcription to proceed. Two molecular motors have been shown to be capable of performing this function in Bacteria: Mfd/TCRF and UvrD (Fig. 2A). These enzymes interact with stalled RNAPs and act to translocate the transcriptional complex away from the lesion. The Mfd motor pushes RNAP forward[42,43] while UvrD pulls it backward[44]. At the same time, these motors serve as links between the polymerase and NER factors. Mfd binds to both the polymerase and the NER factor responsible for the direct recognition of damage, UvrA[43,45,46]. In contrast, UvrD causes backtracking and conformational changes in a lesion-stalled polymerase that stimulate the recruitment of a UvrA dimer and UvrB[40]. While both processes appear to contribute to TCR in different conditions, recent work suggests that the majority of TCR and, due to pervasive transcription perhaps the majority of NER, occurs via the UvrD-dependent mechanism[40,41].

Replication-coupled DNA repair

There are a multitude of mechanisms for maintaining genome integrity during replication and different types of lesions are repaired by distinct pathways. In particular, the replicative helicase may proceed through single-stranded forms of damage that the replicative polymerase cannot. This leads to a de-coupling of DNA unwinding and DNA synthesis and the potential generation of large regions of single-stranded DNA. In these cases, replication restart can be stimulated by translesion synthesis, template-switching, or repriming[47,48]. However, other mechanisms may serve to couple replication to repair instead of skipping over lesions to be dealt with later. In bacteria, the SOS response activates the expression of DNA repair factors upon the detection of ssDNA formed by stalling of the polymerase via the combined activities of RecA and

LexA[49]. These pathways can be considered as replication-coupled repair (RCR) as the damaged DNA is identified via consequences of its interruption of the replication process[48].

In another example of RCR, lesions lead to the formation of broken replication forks and double-stranded breaks. In the context of our discussion here, the conversion of a lesion into a more toxic intermediate represents an indirect mechanism of detection. The resulting double-stranded break may be repaired via either homologous recombination, break-induced replication, or non-homologous end-joining mechanisms[50]. Although logically consistent with the idea of an indirect detection mechanism, this pathway has significant drawbacks as double-stranded breaks represent a substantial problem for the cell and often lead to the chromosomal rearrangements or the outright loss of genetic information.

Instead of creating more damage, lesion-induced stalling of the eukaryotic replisome can instead lead to an intermediate known as a reversed fork[51]. In humans, fork reversal can be catalyzed by ATP-dependent DNA translocases (i.e., ZRANB3 and SMARCAL1)[52] and results in the pairing of the recently replicated DNA strands, the backwards movement of the replisome, and the reannealing of the parental strands downstream of the replisome[53]. This allows for the replisome to reverse direction and places a lesion which may have been recognized in a single-strand form into a double-stranded context where it may be addressed by canonical excision repair pathways. Interestingly, the fork-reversal translocase ZRANB3 possesses endonuclease activity that appears dependent on the activity of the ATPase motor domain[54]. This leads to a hypothesis that only when ZRANB3 is actively remodeling stalled replication forks, does it cleave neighboring DNA and stimulate the recruitment of downstream repair factors (Fig. 2B)[54,55]. This mechanism is perhaps most analogous to transcription-coupled repair as it involves a backtracked intermediate that can be re-started once repair has taken place.

Interstrand crosslinks (ICLs) come in many varieties and are particularly toxic to cells as they completely prevent the separation of the two DNA strands and block the progress of the replicative helicase (i.e., CMG in Eukaryotes). Although mechanisms may be present to allow for ICL bypass without repair[56], they more frequently result in replisome stalling. In this case, the convergence of two replisomes headed in opposite directions signals for downstream repair[57]. A current model suggests a tiered approach that depends on the chemical nature of the ICL and the levels of TRAIP-catalyzed CMG ubiquitination that build up over time[58]. For example, an acetaldehyde crosslink is repaired rapidly by a yet unidentified factor prior to ubiquitylation[59], short ubiquitin chains recruit the NEIL3 glycosylase for repair of psoralen and abasic ICLs, and longer ubiquitin chains result in p97-dependent CMG helicase segregation and recruitment of the Fanconia Anemia pathway endonucleases for repair of cisplatin ICLs[60].

In summary, the stalling of RNAP, DNAP, and replicative helicases are examples of indirect mechanisms for sensing DNA damage. The critical feature being the interruption of a central process unrelated to DNA repair that has been efficiently coupled to the recruitment of repair proteins. One can also imagine the pressure behind the evolution of these indirect mechanisms as lesions that are disrupting essential cellular processes must have a greater priority for repair than others and need to be addressed as quickly as possible.

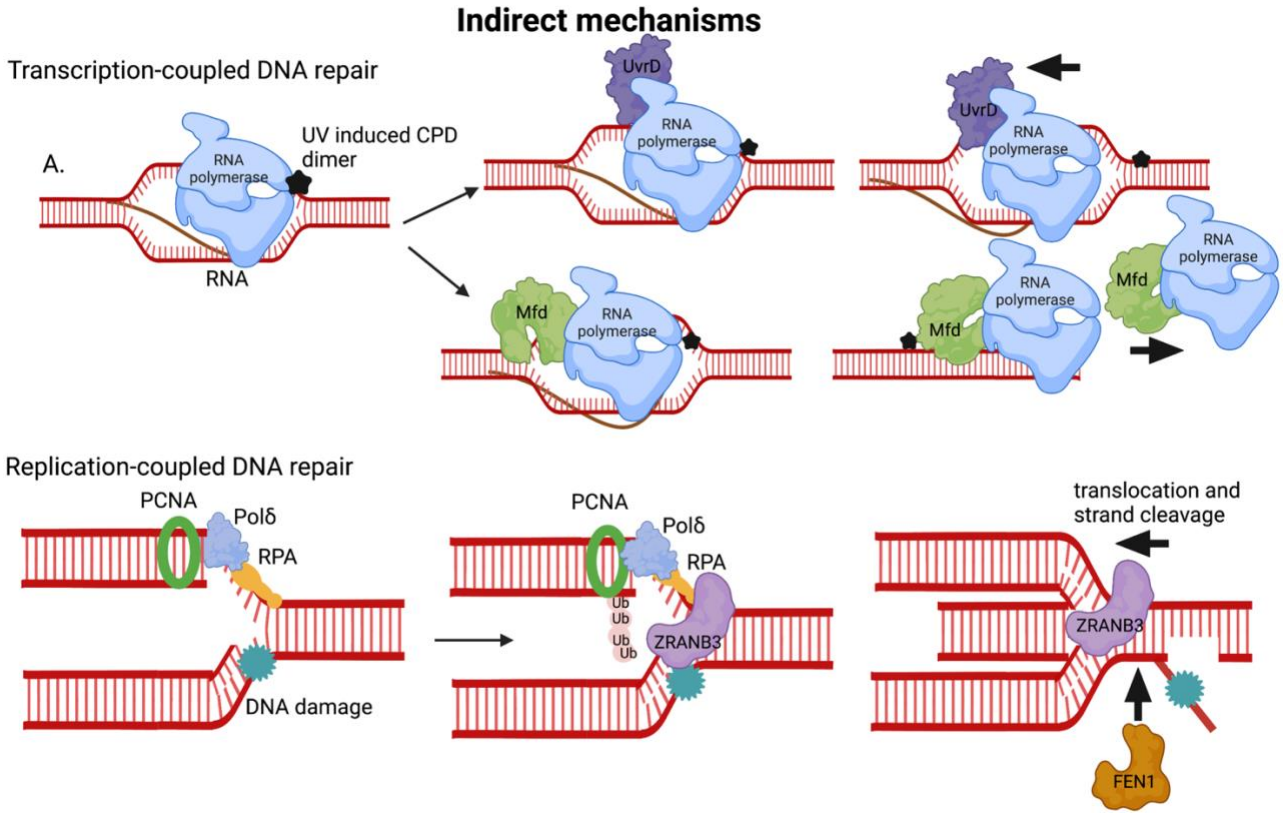


Figure 2.2: Examples of indirect mechanisms where downstream consequences of lesions are recognized. **(A)** RNAP stalled at the site of lesion can serve as a signal of damage and obscure the lesion within its active site. The UvrD helicase can bind RNAP and pull it backward so the lesion can be accessed by the repair enzymes or the Mfd translocase can bind and push RNAP forward leading to RNAP dissociation and termination of transcription. **(B)** During replication, the separation of DNA strands exposes a DNA lesion that blocks the polymerase. The subsequent polyubiquitinylation of PCNA leads to the recruitment of the ZRANB3 translocase. ZRANB3 stimulates fork reversal, the movement of the lesion to a double-stranded region, strand cleavage, and recruitment of downstream repair factors such as FEN1.

2.5 Inference Mechanisms

Here we describe a third potential sensing mechanism where neither the lesion itself nor its consequences are detected. Instead, conditions that will likely lead to damage are detected and acted on. Oxygen radicals are generated as a byproduct of cellular respiration or in response to ultraviolet or ionizing radiation and represent a major source of DNA damage. Monovalent reduction of oxygen produces the superoxide ion ($O_2^{\bullet}$) and can be readily converted to hydrogen peroxide which can then produce additional free radicals including $^{\bullet}OH$, $^{\bullet}NO$, and $^{\bullet}NO_2$

collectively referred to as reactive oxygen and nitrogen species (RONS)[61]. RONS can be reduced by small molecules such as glutathione and enzymes such as superoxide dismutase[62]. However, oxidative stress occurs when endogenously or exogenously produced RONS exceed cellular antioxidant defenses[63,64]. The resulting DNA lesions are varied and include base modifications, single-strand, or double-strand breaks[65].

However, RONS can also chemically modify proteins by oxidation of solvent accessible sulfhydryl side chains of cysteine residues[66–68]. Specific examples of the resulting post-translational modifications include conversion of cysteine thiols to disulfide or further oxidation to sulfinic, sulfenic, and sulfonic acid derivatives[69]. Disulfide bonds can stabilize proteins structurally and can also lead to changes in protein activity. The latter case can be thought of as a redox switch which elicits functional effects in response to changes in oxidative state[70]. In addition to the modification of cysteines, redox potential can modify protein structure and function through the electrochemistry of iron sulfur clusters[71]. Here, we discuss examples where enzymes may infer the current or future presence of DNA damage by detecting the presence of oxidative conditions even though actual damage has not been directly or indirectly sensed.

We want to distinguish this inference mechanism from that of more general stress responses where entire genomic programs are enacted in the face of environmental stress. For example, in Eukaryotes, ATM has been reported to form redox-dependent disulfide-bonded dimers that are activated to phosphorylate downstream regulatory targets including p53[72]. Although the activation of ATM in this case is caused by an environment likely to cause DNA damage, the redox dependent dimer of ATM has not been shown to be involved directly in the machinery of repair.

The redox-dependent activation of the DNA repair helicase UvrD1 from Mtb

DNA helicases are enzymes that couple ATP binding and hydrolysis to translocate on single stranded ssDNA and unwind dsDNA and are involved in DNA replication, recombination, and repair[73]. The SF1 superfamily of helicases is the largest group of known helicases and includes UvrD, Rep, and PcrA[74]. They consist of two RecA-like domains (1A and 2A) and two accessory subdomains (1B and 2B)[74]. SF1 helicase monomers have been shown to possess ssDNA translocase activity but no helicase activity in the absence of accessory factors or force exerted on the DNA. In other words, helicase activity requires a dimer as shown in the case of *E. coli* UvrD[75–77].

Recently, we described a redox-sensing dimerization and activation mechanism for UvrD1 from *Mycobacterium tuberculosis* (Mtb) that is not possessed by its well-studied homolog from *E. coli*[78]. Mtb UvrD1 is a 771 amino acid protein, and the predicted structure shows two core helicase domains 1A and 1B coupled with auxiliary 2A and 2B subdomains[74]. It has 3 cysteines, two that are partially buried and located in the 1A and 1B domains and one solvent exposed located in the regulatory 2B domain. Mutation of the 2B cysteine prevented both the dimerization and helicase activity in the presence of oxidative conditions. Interestingly, the 2B domain cysteine in UvrD1 is conserved in helicases among the order Corynebacteriales of class Actinobacteria that includes several pathogenic bacteria suggesting a common redox-dependent activation mechanism amongst these organisms[78].

Mtb is an intracellular pathogen that causes tuberculosis disease in humans[79]. During infection, it faces a variety of host-mediated stress responses including RONS produced by macrophages[80,81]. This RONS leads to the formation of DNA damage[82] and UvrD1 has been shown to be important in a subset of DNA repair mechanisms[83]. Since helicase activity requires dimerization via disulfide bond formation, we proposed that oxidative stress is being detected by

the enzyme and being used to stimulate all the repair pathways in which its unwinding activity is required. In this way, UvrD1 would be inferring the current or future presence of DNA damage without any direct or indirect sensing of actual lesions (Fig. 3).

Potential inference mechanisms by iron-sulfur cluster proteins

Iron-sulfur clusters are structural elements in a large array of protein regulatory factors[71]. In a significant subset of cases, the proteins are involved in nucleic acid metabolism and the oxidation state of 4Fe4S clusters has been found to affect protein activities. Some studies point to the potential role of the clusters in identifying DNA damage through DNA-dependent oxidation[84]. Interestingly, and more germane to our discussion here, other studies link the oxidation state of the cluster directly to protein activity. Examples include the yeast DNA polymerase delta (Pol δ) and the *E. coli* helicase DinG.

The redox state of the iron-sulfur cluster of Polδ has been shown to directly affect its enzymatic activity[85]. More specifically, the oxidation of the cluster from +2 to +3 leads to 6X slower polymerization rates. This effect may be explained via a tighter association with DNA in the oxidized form leading to more stalling and lower processivity. This idea is consistent with increased DNA binding observed with oxidized clusters in EndoIII and DNA primase[86–88] . In terms of a DNA repair inference mechanism, one could imagine that slower polymerization inhibits replication under oxidative conditions where DNA damage is likely. This could provide more time for repair pathways to act and avoid problematic conflicts between the replisome and DNA lesions.

The *E. coli* DNA damage-inducible helicase, DinG, contains a 4Fe4S cluster that affects its DNA unwinding activity[89]. In contrast to Pol δ, and as in the case of UvrD1, reduction of the cluster in DinG inhibits unwinding while oxidation activates its helicase activity and primes it to

participate in DNA repair pathways. These observations taken together suggest a model wherein redox-sensitive replication enzymes are inhibited, and redox-sensitive DNA repair enzymes are activated via oxidative conditions. While more specific mechanisms have been suggested for the role of iron sulfur clusters in detecting DNA damage via electrochemical coupling[90], this general mechanism could represent an overall inference of DNA damage and a shift in the activities of these enzymes in preparation for repair.

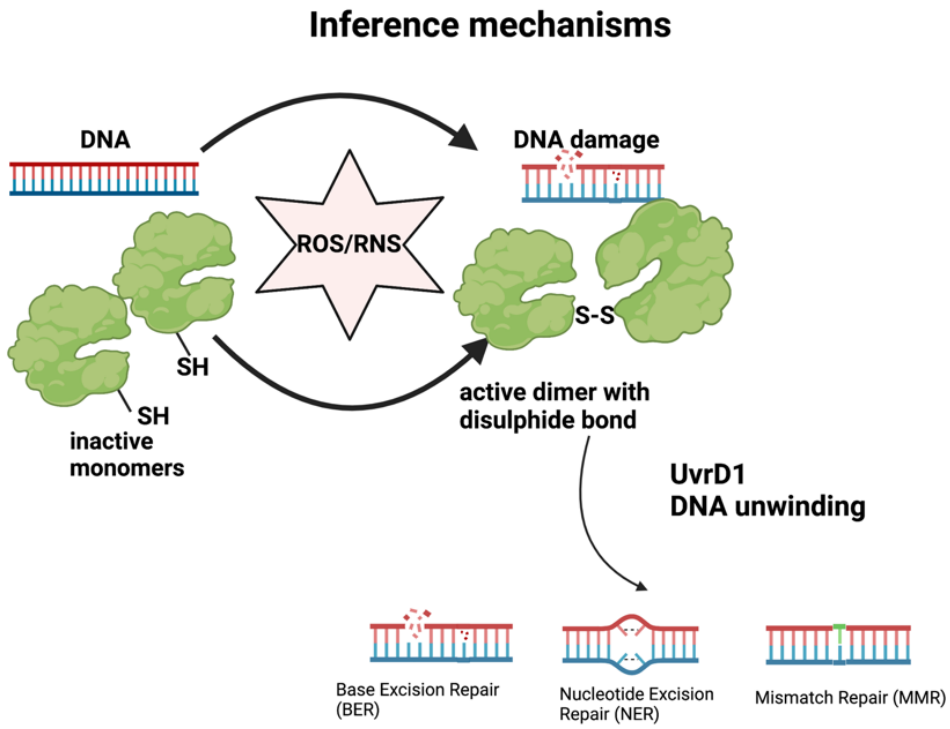


Figure 2.3: An example of a potential inference mechanism. The presence of reactive oxygen and nitrogen species that may lead to DNA damage stimulates the cysteine-dependent dimerization of UvrD1. The active dimer's is now primed to contribute to multiple pathways of DNA repair.

2.6 Conclusion/Summary

We have discussed three broad strategies to maintain the integrity of the genome in the face of DNA damage. Direct mechanisms rely on the principles of molecular recognition of a DNA lesion. Here, members of dedicated repair pathways specifically bind to the lesion itself thus signaling its presence. This approach allows for constant surveillance of the current state of the genome regardless of the conditions. Indirect mechanisms rely on the activity or consequence of a lesion instead of its presence. Specifically, the ability of lesions to interrupt cellular processes such as transcription and replication lead to stalled molecular motors that must be attended to. In this way, indirect mechanisms sense lesions that took place in the past and were not caught by direct mechanisms. These pathways provide an ability for the cell to address lesions that are currently inhibiting its growth and survival. Inference mechanisms may provide another level of protection wherein cellular conditions that may eventually lead to the creation of lesions are detected via coupled changes in protein chemistry. These mechanisms may allow the cell to prepare for the eventual presence of damage in the future.

The logic of the evolution of a set of mechanisms dependent on different, but related, signals to detect and address genome insults is clear. By not depending on any single strategy for recognizing and repairing damage, the cell is made more robust in the face of the variety of conditions encountered and the constant battle to maintain genome integrity.

Acknowledgement: Many thanks to Peter Burgers, Tim Lohman, Nima Mossamaparast, Alessandro Vindigini, Eric Tomko, and Ana Ruiz Manzano for thoughtful conversations and suggestions during the writing of this manuscript.

Figures created with BioRender.com

2.7 References

1. Chatterjee, N., & Walker, G. C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. *Environmental and Molecular Mutagenesis*, *58*(5), 235–263. https://doi.org/10.1002/em.22087

2. Friedberg, E. C. (2003). DNA damage and repair. *Nature*, *421*(6921), 436–440. https://doi.org/10.1038/nature01408

3. Jackson, S. P., & Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature*, *461*(7267), 1071–1078. https://doi.org/10.1038/nature08467

4. Marinus, M. G., & Løbner-Olesen, A. (2014). DNA Methylation. *EcoSal Plus*, *6*(1). https://doi.org/10.1128/ecosalplus.esp-0003-2013

5. Mattei, A. L., Bailly, N., & Meissner, A. (2022). DNA methylation: a historical perspective. *Trends in Genetics*, *38*(7), 676–707. https://doi.org/10.1016/j.tig.2022.03.010

6. Heras, S., Smits, K., Schauwer, C. D., & Soom, A. V. (2017). Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during pronuclear development in equine zygotes produced by ICSI. *Epigenetics & Chromatin*, *10*(1), 13. https://doi.org/10.1186/s13072-017-0120-x

7. Salvaing, J., Aguirre-Lavin, T., Boulesteix, C., Lehmann, G., Debey, P., & Beaujean, N. (2012). 5-Methylcytosine and 5-Hydroxymethylcytosine Spatiotemporal Profiles in the Mouse Zygote. *PLoS ONE*, *7*(5), e38156. https://doi.org/10.1371/journal.pone.0038156

8. Bauer, N. C., Corbett, A. H., & Doetsch, P. W. (2015). The current state of eukaryotic DNA base damage and repair. *Nucleic Acids Research*, *43*(21), 10083–10101. https://doi.org/10.1093/nar/gkv1136

9. Krokan, H. E., & Bjørås, M. (2013). Base Excision Repair. *Cold Spring Harbor Perspectives in Biology*, *5*(4), a012583. https://doi.org/10.1101/cshperspect.a012583

10. Hofer, A., Liu, Z. J., & Balasubramanian, S. (2019). Detection, Structure and Function of Modified DNA Bases. *Journal of the American Chemical Society*, *141*(16), 6420–6429. https://doi.org/10.1021/jacs.9b01915

11. Cadet, J., & Wagner, J. R. (2013). DNA Base Damage by Reactive Oxygen Species, Oxidizing Agents, and UV Radiation. *Cold Spring Harbor Perspectives in Biology*, *5*(2), a012559. https://doi.org/10.1101/cshperspect.a012559

12. Guven, M., Barnouin, K., Snijders, A. P., & Karran, P. (2016). Photosensitized UVA-Induced Cross-Linking between Human DNA Repair and Replication Proteins and DNA Revealed by Proteomic Analysis. *Journal of Proteome Research*, *15*(12), 4612–4623. https://doi.org/10.1021/acs.jproteome.6b00717

13. Kunkel, T. A., & Erie, D. A. (2005). DNA MISMATCH REPAIR*. *Annual Review of Biochemistry*, *74*(1), 681–710. https://doi.org/10.1146/annurev.biochem.74.082803.133243

14. Schärer, O. D. (2013). Nucleotide Excision Repair in Eukaryotes. *Cold Spring Harbor Perspectives in Biology*, *5*(10), a012609. https://doi.org/10.1101/cshperspect.a012609

15. Kusakabe, M., Onishi, Y., Tada, H., Kurihara, F., Kusao, K., Furukawa, M., Iwai, S., Yokoi, M., Sakai, W., & Sugasawa, K. (2019). Mechanism and regulation of DNA damage recognition in nucleotide excision repair. *Genes and Environment*, *41*(1), 2. https://doi.org/10.1186/s41021-019-0119-6

16. Varga, T., & Aplan, P. D. (2005). Chromosomal aberrations induced by double strand DNA breaks. *DNA Repair*, *4*(9), 1038–1046. https://doi.org/10.1016/j.dnarep.2005.05.004

17. Mao, Z., Bozzella, M., Seluanov, A., & Gorbunova, V. (2008). Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair*, *7*(10), 1765–1771. https://doi.org/10.1016/j.dnarep.2008.06.018

18. Rastogi, R. P., Richa, Kumar, A., Tyagi, M. B., & Sinha, R. P. (2010). Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair. *Journal of Nucleic Acids*, *2010*, 592980. https://doi.org/10.4061/2010/592980

19. Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y., & Takeda, S. (2006). Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair*, *5*(9–10), 1021–1029. https://doi.org/10.1016/j.dnarep.2006.05.022

20. Aravind, L., & Koonin, E. V. (2001). Prokaryotic Homologs of the Eukaryotic DNA-End-Binding Protein Ku, Novel Domains in the Ku Protein and Prediction of a Prokaryotic Double-Strand Break Repair System. *Genome Research*, *11*(8), 1365–1374. https://doi.org/10.1101/gr.181001

21. Lieber, M. R. (2008). The Mechanism of Human Nonhomologous DNA End Joining*. *Journal of Biological Chemistry*, *283*(1), 1–5. https://doi.org/10.1074/jbc.r700039200

22. Walker, J. R., Corpina, R. A., & Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, *412*(6847), 607–614. https://doi.org/10.1038/35088000

23. Banda, D. M., Nuñez, N. N., Burnside, M. A., Bradshaw, K. M., & David, S. S. (2017). Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: Mechanism, metals and medicine. *Free Radical Biology and Medicine*, *107*, 202–215. https://doi.org/10.1016/j.freeradbiomed.2017.01.008

24. Lee, A. J., Warshaw, D. M., & Wallace, S. S. (2014). Insights into the glycosylase search for damage from single-molecule fluorescence microscopy. *DNA Repair*, *20*, 23–31. https://doi.org/10.1016/j.dnarep.2014.01.007

25. Hollis, T., Ichikawa, Y., & Ellenberger, T. (2000). DNA bending and a flip-out mechanism for base excision by the helix–hairpin–helix DNA glycosylase, Escherichia coli AlkA. *The EMBO Journal*, *19*(4), 758–766. https://doi.org/10.1093/emboj/19.4.758

26. Hsieh, P., & Yamane, K. (2008). DNA mismatch repair: Molecular mechanism, cancer, and ageing. *Mechanisms of Ageing and Development*, *129*(7–8), 391–407. https://doi.org/10.1016/j.mad.2008.02.012

27. Wang, H., Yang, Y., Schofield, M. J., Du, C., Fridman, Y., Lee, S. D., Larson, E. D., Drummond, J. T., Alani, E., Hsieh, P., & Erie, D. A. (2003). DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proceedings of the National Academy of Sciences*, *100*(25), 14822–14827. https://doi.org/10.1073/pnas.2433654100

28. Tessmer, I., Yang, Y., Zhai, J., Du, C., Hsieh, P., Hingorani, M. M., & Erie, D. A. (2008). Mechanism of MutS Searching for DNA Mismatches and Signaling Repair*. *Journal of Biological Chemistry*, *283*(52), 36646–36654. https://doi.org/10.1074/jbc.m805712200

29. Salsbury, F. R., Clodfelter, J. E., Gentry, M. B., Hollis, T., & Scarpinato, K. D. (2006). The molecular mechanism of DNA damage recognition by MutS homologs and its consequences for cell death response. *Nucleic Acids Research*, *34*(8), 2173–2185. https://doi.org/10.1093/nar/gkl238

30. Hingorani, M. M. (2016). Mismatch binding, ADP–ATP exchange and intramolecular signaling during mismatch repair. *DNA Repair*, *38*, 24–31. https://doi.org/10.1016/j.dnarep.2015.11.017

31. Vvedenskaya, I. O., Vahedian-Movahed, H., Bird, J. G., Knoblauch, J. G., Goldman, S. R., Zhang, Y., Ebright, R. H., & Nickels, B. E. (2014). Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science (New York, NY)*, *344*(6189), 1285–1289. https://doi.org/10.1126/science.1253458

32. Saba, J., Chua, X. Y., Mishanina, T. V., Nayak, D., Windgassen, T. A., Mooney, R. A., & Landick, R. (2019). The elemental mechanism of transcriptional pausing. *ELife*, *8*, e40981. https://doi.org/10.7554/elife.40981

33. Vvedenskaya, I. O., Vahedian-Movahed, H., Bird, J. G., Knoblauch, J. G., Goldman, S. R., Zhang, Y., Ebright, R. H., & Nickels, B. E. (2014). Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science*, *344*(6189), 1285–1289. https://doi.org/10.1126/science.1253458

34. Chen, Z., Gabizon, R., Brown, A. I., Lee, A., Song, A., Díaz-Celis, C., Kaplan, C. D., Koslover, E. F., Yao, T., & Bustamante, C. (2019). High-resolution and high-accuracy topographic and transcriptional maps of the nucleosome barrier. *ELife*, *8*, e48281. https://doi.org/10.7554/elife.48281

35. Sydow, J. F., & Cramer, P. (2009). RNA polymerase fidelity and transcriptional proofreading. *Current Opinion in Structural Biology*, *19*(6), 732–739. https://doi.org/10.1016/j.sbi.2009.10.009

36. Sydow, J. F., Brueckner, F., Cheung, A. C. M., Damsma, G. E., Dengl, S., Lehmann, E., Vassylyev, D., & Cramer, P. (2009). Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Molecular Cell*, *34*(6), 710–721. https://doi.org/10.1016/j.molcel.2009.06.002

37. Kunkel, T. A., & Bebenek, K. (2000). DNA REPLICATION FIDELITY1. *Annual Review of Biochemistry*, *69*(1), 497–529. https://doi.org/10.1146/annurev.biochem.69.1.497

38. Bohr, V. A., Smith, C. A., Okumoto, D. S., & Hanawalt, P. C. (1985). DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell*, *40*(2), 359–369. https://doi.org/10.1016/0092-8674(85)90150-3

39. Mellon, I., Spivak, G., & Hanawalt, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell*, *51*(2), 241–249. https://doi.org/10.1016/0092-8674(87)90151-6

40. Bharati, B. K., Gowder, M., Zheng, F., Alzoubi, K., Svetlov, V., Kamarthapu, V., Weaver, J. W., Epshtein, V., Vasilyev, N., Shen, L., Zhang, Y., & Nudler, E. (2022). Crucial role and mechanism of transcription-coupled DNA repair in bacteria. *Nature*, 1–8. https://doi.org/10.1038/s41586-022-04530-6

41. Martinez, B., Bharati, B. K., Epshtein, V., & Nudler, E. (2022). Pervasive Transcription-coupled DNA repair in E. coli. *Nature Communications*, *13*(1), 1702. https://doi.org/10.1038/s41467-022-28871-y

42. Graves, E. T., Duboc, C., Fan, J., Stransky, F. ccedil ois, Leroux-Coyau, M., & Strick, T. R. (2015). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. *Nature Structural & Molecular Biology*, *22*(6), 1–9. https://doi.org/10.1038/nsmb.3019

43. Fan, J., Leroux-Coyau, M., Savery, N. J., & Strick, T. R. (2016). Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. *Nature*, *536*(7615), 234–237. https://doi.org/10.1038/nature19080

44. Epshtein, V., Kamarthapu, V., McGary, K., Svetlov, V., Ueberheide, B., Proshkin, S., Mironov, A., & Nudler, E. (2014). UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature*, *505*(7483), 372. https://doi.org/10.1038/nature12928

45. Selby, C. P., & Sancar, A. (1995). Structure and Function of Transcription-Repair Coupling Factor. *Journal of Biological Chemistry*, *270*(9), 4882–4889. https://doi.org/10.1074/jbc.270.9.4882

46. Selby, C. P. (2017). Mfd Protein and Transcription-Repair Coupling in Escherichia coli. *Photochemistry and Photobiology*, *93*(1), 280–295. https://doi.org/10.1111/php.12675

47. Quinet, A., Tirman, S., Cybulla, E., Meroni, A., & Vindigni, A. (2021). To skip or not to skip: choosing repriming to tolerate DNA damage. *Molecular Cell*, *81*(4), 649–658. https://doi.org/10.1016/j.molcel.2021.01.012

48. Cortez, D. (2019). Replication-Coupled DNA Repair. *Molecular Cell*, *74*(5), 866–876. https://doi.org/10.1016/j.molcel.2019.04.027

49. Maslowska, K. H., Makiela-Dzbenska, K., & Fijalkowska, I. J. (2019). The SOS system: A complex and tightly regulated response to DNA damage. *Environmental and Molecular Mutagenesis*, *60*(4), 368–384. https://doi.org/10.1002/em.22267

50. Chang, H. H. Y., Pannunzio, N. R., Adachi, N., & Lieber, M. R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology*, *18*(8), 495–506. https://doi.org/10.1038/nrm.2017.48

51. Neelsen, K. J., & Lopes, M. (2015). Replication fork reversal in eukaryotes: from dead end to dynamic response. *Nature Reviews Molecular Cell Biology*, *16*(4), 207–220. https://doi.org/10.1038/nrm3935

52. Taglialatela, A., Alvarez, S., Leuzzi, G., Sannino, V., Ranjha, L., Huang, J.-W., Madubata, C., Anand, R., Levy, B., Rabadan, R., Cejka, P., Costanzo, V., & Ciccia, A. (2017). Restoration of Replication Fork Stability in BRCA1- and BRCA2-Deficient Cells by Inactivation of SNF2-Family Fork Remodelers. *Molecular Cell*, *68*(2), 414-430.e8. https://doi.org/10.1016/j.molcel.2017.09.036

53. Quinet, A., Lemaçon, D., & Vindigni, A. (2017). Replication Fork Reversal: Players and Guardians. *Molecular Cell*, *68*(5), 830–833. https://doi.org/10.1016/j.molcel.2017.11.022

54. Weston, R., Peeters, H., & Ahel, D. (2012). ZRANB3 is a structure-specific ATP-dependent endonuclease involved in replication stress response. *Genes & Development*, *26*(14), 1558–1572. https://doi.org/10.1101/gad.193516.112

55. Badu-Nkansah, A., Mason, A. C., Eichman, B. F., & Cortez, D. (2016). Identification of a Substrate Recognition Domain in the Replication Stress Response Protein Zinc Finger Ran-binding Domain-containing Protein 3 (ZRANB3)*. *Journal of Biological Chemistry*, *291*(15), 8251–8257. https://doi.org/10.1074/jbc.m115.709733

56. Huang, J., Liu, S., Bellani, M. A., Thazhathveetil, A. K., Ling, C., de Winter, J. P., Wang, Y., Wang, W., & Seidman, M. M. (2013). The DNA Translocase FANCM/MHF Promotes Replication Traverse of DNA Interstrand Crosslinks. *Molecular Cell*, *52*(3), 434–446. https://doi.org/10.1016/j.molcel.2013.09.021

57. Zhang, J., Dewar, J. M., Budzowska, M., Motnenko, A., Cohn, M. A., & Walter, J. C. (2015). DNA interstrand cross-link repair requires replication-fork convergence. *Nature Structural & Molecular Biology*, *22*(3), 242–247. https://doi.org/10.1038/nsmb.2956

58. Wu, R. A., Pellman, D. S., & Walter, J. C. (2020). The Ubiquitin Ligase TRAIP: Double-Edged Sword at the Replisome. *Trends in Cell Biology*, *31*(2), 75–85. https://doi.org/10.1016/j.tcb.2020.11.007

59. Hodskinson, M. R., Bolner, A., Sato, K., Kamimae-Lanning, A. N., Rooijers, K., Witte, M., Mahesh, M., Silhan, J., Petek, M., Williams, D. M., Kind, J., Chin, J. W., Patel, K. J., & Knipscheer, P. (2020). Alcohol-derived DNA crosslinks are repaired by two distinct mechanisms. *Nature*, *579*(7800), 603–608. https://doi.org/10.1038/s41586-020-2059-5

60. Wu, R. A., Semlow, D. R., Kamimae-Lanning, A. N., Kochenova, O. V., Chistol, G., Hodskinson, M. R., Amunugama, R., Sparks, J. L., Wang, M., Deng, L., Mimoso, C. A., Low, E., Patel, K. J., & Walter, J. C. (2019). TRAIP is a master regulator of DNA interstrand cross-link repair. *Nature*, *567*(7747), 267–272. https://doi.org/10.1038/s41586-019-1002-0

61. Winterbourn, C. C. (2008). Reconciling the chemistry and biology of reactive oxygen species. *Nature Chemical Biology*, *4*(5), 278–286. https://doi.org/10.1038/nchembio.85

62. Aoyama, K., & Nakaki, T. (2015). Glutathione in Cellular Redox Homeostasis: Association with the Excitatory Amino Acid Carrier 1 (EAAC1). *Molecules*, *20*(5), 8742–8758. https://doi.org/10.3390/molecules20058742

63. Finkel, T. (2011). Signal transduction by reactive oxygen species. *The Journal of Cell Biology*, *194*(1), 7–15. https://doi.org/10.1083/jcb.201102095

64. Finkel, T., & Holbrook, N. J. (2000). Oxidants, oxidative stress and the biology of ageing. *Nature*, *408*(6809), 239–247. https://doi.org/10.1038/35041687

65. Fleming, A. M., & Burrows, C. J. (2020). Interplay of Guanine Oxidation and G-Quadruplex Folding in Gene Promoters. *Journal of the American Chemical Society*, *142*(3), 1115–1136. https://doi.org/10.1021/jacs.9b11050

66. Bonetto, V., & Ghezzi, P. (2006). *Redox Proteomics: From Protein Modifications to Cellular Dysfunction and Diseases*. 101–122. https://doi.org/10.1002/0471973122.ch4

67. Bak, D. W., & Weerapana, E. (2014). Cysteine-mediated redox signalling in the mitochondria. *Molecular BioSystems*, *11*(3), 678–697. https://doi.org/10.1039/c4mb00571f

68. Sun, M., Zhang, Q., Wang, Y., Ge, W., & Guo, D. (2016). Prediction of redox-sensitive cysteines using sequential distance and other sequence-based features. *BMC Bioinformatics*, *17*(1), 316. https://doi.org/10.1186/s12859-016-1185-4

69. Meng, Y., & Li, L. (2021). Cysteine post-translational modifications: ten years from chemical proteomics to bioinformatics. *ArXiv*.

70. Fra, A., Yoboue, E. D., & Sitia, R. (2017). Cysteines as Redox Molecular Switches and Targets of Disease. *Frontiers in Molecular Neuroscience*, *10*, 167. https://doi.org/10.3389/fnmol.2017.00167

71. Crack, J. C., & Brun, N. E. L. (2018). Redox-Sensing Iron–Sulfur Cluster Regulators. *Antioxidants & Redox Signaling*, *29*(18), 1809–1829. https://doi.org/10.1089/ars.2017.7361

72. Guo, Z., Kozlov, S., Lavin, M. F., Person, M. D., & Paull, T. T. (2010). ATM activation by oxidative stress. *Science (New York, N.Y.)*, *330*(6003), 517–521. https://doi.org/10.1126/science.1192912

73. Raney, K. D., Byrd, A. K., & Aarattuthodiyil, S. (2012). DNA Helicases and DNA Motor Proteins. *Advances in Experimental Medicine and Biology*, *767*, 17–46. https://doi.org/10.1007/978-1-4614-5037-5_2

74. Lohman, T. M., Tomko, E. J., & Wu, C. G. (2008). Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nature Reviews Molecular Cell Biology*, *9*(5), 391–401. https://doi.org/10.1038/nrm2394

75. Niedziela-Majka, A., Chesnik, M. A., Tomko, E. J., & Lohman, T. M. (2007). Bacillus stearothermophilus PcrA Monomer Is a Single-stranded DNA Translocase but Not a Processive Helicase in Vitro *. *Journal of Biological Chemistry*, *282*(37), 27076–27085. https://doi.org/10.1074/jbc.m704399200

76. Maluf, N. K., Fischer, C. J., & Lohman, T. M. (2003). A Dimer of Escherichia coli UvrD is the Active Form of the Helicase In Vitro. *Journal of Molecular Biology*, *325*(5), 913–935. https://doi.org/10.1016/s0022-2836(02)01277-9

77. Nguyen, B., Ordabayev, Y., Sokoloski, J. E., Weiland, E., & Lohman, T. M. (2017). Large domain movements upon UvrD dimerization and helicase activation. *Proceedings of the National Academy of Sciences*, *114*(46), 12178–12183. https://doi.org/10.1073/pnas.1712882114

78. Chadda, A., Jensen, D., Tomko, E. J., Manzano, A. R., Nguyen, B., Lohman, T. M., & Galburt, E. A. (2022). Mycobacterium tuberculosis DNA repair helicase UvrD1 is activated by redox-dependent dimerization via a 2B domain cysteine. *Proceedings of the National Academy of Sciences*, *119*(8), e2114501119. https://doi.org/10.1073/pnas.2114501119

79. Delogu, G., Sali, M., & Fadda, G. (2013). THE BIOLOGY OF MYCOBACTERIUM TUBERCULOSIS INFECTION. *Mediterranean Journal of Hematology and Infectious Diseases*, *5*(1), 2013070. https://doi.org/10.4084/mjhid.2013.070

80. Shastri, M. D., Shukla, S. D., Chong, W. C., Dua, K., Peterson, G. M., Patel, R. P., Hansbro, P. M., Eri, R., & O'Toole, R. F. (2018). Role of Oxidative Stress in the Pathology and Management of Human Tuberculosis. *Oxidative Medicine and Cellular Longevity, 2018*, 7695364. https://doi.org/10.1155/2018/7695364

81. Voskuil, M. I., Bartek, I. L., Visconti, K., & Schoolnik, G. K. (2011). The Response of Mycobacterium Tuberculosis to Reactive Oxygen and Nitrogen Species. *Frontiers in Microbiology*, *2*, 105. https://doi.org/10.3389/fmicb.2011.00105

82. Vultos, T. D., Mestre, O., Tonjum, T., & Gicquel, B. (2009). DNA repair in Mycobacterium tuberculosis revisited. *FEMS Microbiology Reviews*, *33*(3), 471–487. https://doi.org/10.1111/j.1574-6976.2009.00170.x

83. Houghton, J., Townsend, C., Williams, A. R., Rodgers, A., Rand, L., Walker, K. B., Böttger, E. C., Springer, B., & Davis, E. O. (2012). Important Role for Mycobacterium tuberculosis UvrD1 in Pathogenesis and Persistence apart from Its Function in Nucleotide Excision Repair. *Journal of Bacteriology*, *194*(11), 2916–2923. https://doi.org/10.1128/jb.06654-11

84. Lukianova, O. A., & David, S. S. (2005). A role for iron–sulfur clusters in DNA repair. *Current Opinion in Chemical Biology*, *9*(2), 145–151. https://doi.org/10.1016/j.cbpa.2005.02.006

85. Bartels, P. L., Stodola, J. L., Burgers, P. M. J., & Barton, J. K. (2017). A Redox Role for the [4Fe4S] Cluster of Yeast DNA Polymerase δ. *Journal of the American Chemical Society*, *139*(50), 18339–18348. https://doi.org/10.1021/jacs.7b10284

86. Gorodetsky, A. A., Boal, A. K., & Barton, J. K. (2006). Direct Electrochemistry of Endonuclease III in the Presence and Absence of DNA. *Journal of the American Chemical Society*, *128*(37), 12082–12083. https://doi.org/10.1021/ja064784d

87. Tse, E. C. M., Zwang, T. J., & Barton, J. K. (2017). The Oxidation State of [4Fe4S] Clusters Modulates the DNA-Binding Affinity of DNA Repair Proteins. *Journal of the American Chemical Society*, *139*(36), 12784–12792. https://doi.org/10.1021/jacs.7b07230

88. O'Brien, E., Holt, M. E., Thompson, M. K., Salay, L. E., Ehlinger, A. C., Chazin, W. J., & Barton, J. K. (2017). The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. *Science*, *355*(6327). https://doi.org/10.1126/science.aag1789

89. Ren, B., Duan, X., & Ding, H. (2009). Redox Control of the DNA Damage-inducible Protein DinG Helicase Activity via Its Iron-Sulfur Cluster*. *Journal of Biological Chemistry*, *284*(8), 4829–4835. https://doi.org/10.1074/jbc.m807943200

90. Merino, E. J., Boal, A. K., & Barton, J. K. (2008). Biological contexts for DNA charge transport chemistry. *Current Opinion in Chemical Biology*, *12*(2), 229–237. https://doi.org/10.1016/j.cbpa.2008.01.046

Chapter 3: The *Mycobacterium tuberculosis* DNA-repair Helicase UvrD1 is Activated by Redox-dependent Dimerization via a 2B domain Cysteine Conserved in other Actinobacteria.

3.1 Abstract

Mycobacterium tuberculosis (*Mtb*) causes Tuberculosis and, during infection, is exposed to reactive oxygen species (ROS) and reactive nitrogen intermediates (RNI) from the host immune response that can cause DNA damage. UvrD-like proteins are involved in DNA repair and replication and belong to the SF1 family of DNA helicases that use ATP hydrolysis to catalyze DNA unwinding. In *Mtb*, there are two UvrD-like enzymes where UvrD1 is most closely related to other family members. Previous studies have suggested that UvrD1 is exclusively monomeric, however it is well-known that *E. coli* UvrD and other UvrD-family members exhibit monomer-dimer equilibria and unwind as dimers in the absence of accessory factors. Here, we reconcile these incongruent studies by showing that *Mtb* UvrD1 exists in monomer, dimer, and tetramer oligomeric forms where dimerization is regulated by redox potential. We identify a 2B domain cysteine, conserved in many Actinobacteria, that underlies this effect. We also show that UvrD1 DNA unwinding activity correlates specifically with the dimer population and is thus titrated directly via increasing positive (i.e., oxidative) redox potential. Consistent with the regulatory role of the 2B domain and the dimerization-based activation of DNA unwinding

in UvrD-family helicases, these results suggest that UvrD1 is activated under oxidizing conditions when it may be needed to respond to DNA damage during infection.

3.2 Significance Statement (50-120 words)

Mtb is an intracellular pathogen that causes Tuberculosis and is exposed to oxidative insults from immune system macrophages. *Mtb* UvrD1 plays a role in DNA repair during infection and has been suggested to function as a monomer. However, we find that UvrD1 can self-assemble, that the balance between monomer and dimer depends on redox potential via a cysteine residue in the regulatory 2B domain, and that unwinding activity is uniquely a property of the dimer. Our results provide direct evidence of the domain interface in these ubiquitous enzymes, reveal a sub-family of UvrD-like enzymes regulated by redox potential, and suggest that *Mt*b UvrD1 is activated by the oxidative conditions imposed during infection.

3.3 Introduction

DNA repair plays an essential role in the ability of organisms to maintain genome integrity in the face of environmental stresses. One particularly flexible and conserved pathway is Nucleotide Excision Repair (NER) which detects, and repairs bulky nucleotide lesions caused by UV light, environmental mutagens, and a subset of oxidative lesions (1–3). In bacteria, global genome NER is initiated when lesions are recognized directly by

UvrA, although an alternative pathway called transcription-coupled NER depends on RNA polymerase stalling as the initiation event (4–7). The removal of the lesion eventually requires the recruitment of a helicase to the site of damage. In Eukaryotes, this function is filled by TFIIH (8–10), while prokaryotes utilize the UvrD-family enzymes (1, 3, 11). In addition to its role in NER, UvrD participates in a range of other pathways of DNA metabolism such as replication (12–15) and recombination (16–18).

UvrD has been well-characterized in many contexts and from model organisms including *E. coli* and *B. subtilis*. It is a superfamily 1A (SF1A) helicase, as defined by core helicase domains 1A and 1B coupled with auxiliary 2A and 2B sub-domains (19, 20). It can both translocate on single-stranded DNA (ssDNA) and unwind double-stranded DNA (dsDNA) under specific conditions. More precisely, while monomers of UvrD-family members (UvrD, Rep and PcrA) are ATP-dependent ssDNA translocases, dimeric forms of these enzymes are required to unwind duplex DNA *in vitro* in the absence of accessory factors or force (21–27). In Rep, this activation is regulated by the mobile 2B domain as both deletion of the 2B domain or a crosslinked 2B domain construct activate the Rep monomer for unwinding (28, 29). Activation of the dimeric UvrD helicase is also accompanied by re-orientation of its 2B sub-domain (30). Additionally, the rotational orientation of the 2B domain regulates the force-dependent unwinding activity of both UvrD and Rep monomers (31, 32). Helicase activation can also occur via binding with accessory factors. For example, *B. stearothermophilus* RepD activates PcrA monomers (31, 32, 25) and the mismatch repair protein MutL activates UvrD monomers (33, 34). Furthermore, these interactions directly affect the orientation of the regulatory 2B domain

(34). In addition to its association with other repair proteins (32), UvrD associates with RNA polymerase through its C-terminal RNAP Interaction Doman (RID) during one mode of transcription coupled NER (35–37). This interaction leads to the stimulation of RNAP backtracking and the recruitment of UvrAB (37).

While many studies have been reported focusing on UvrD-family helicases from model bacteria, less is known about these enzymes in the distantly related human pathogen, *Mycobacterium tuberculosis (Mtb)*. *Mtb* is the causative agent of Tuberculosis and is the leading cause of death worldwide from an infectious agent (38). Although DNA metabolism pathways such as transcription and repair are generally conserved in bacteria, important differences exist (39–42). This appears to be especially true in *Mtb*, perhaps as it is highly evolved for a relatively narrow niche (43). Interestingly, and in contrast to model bacteria, *Mtb* contains two UvrD family enzymes: UvrD1 and UvrD2 (44, 45). UvrD1 has high homology to *E. coli* UvrD including the C-terminal RID (45, 46). Previous work on *Mtb* UvrD1 has shown that it is important for survival after UV and oxidative damage as well as for pathogenesis in mice (47). In stark contrast to other UvrD-family members, UvrD1 has been reported to be monomeric and to either possess helicase activity directly or require activation via the binding of *Mtb* Ku (45, 46, 48).

Here we report that UvrD1 exists in monomer, dimer, and tetrameric forms where dimerization is redox-dependent and is correlated with helicase activity. We identify a 2B domain cysteine that is required for the redox-dependent dimerization, demonstrating that the 2B sub-domain is directly involved in dimerization. Our results explain the function of UvrD1 in the context of the large body of work on UvrD-family proteins and suggest a model where UvrD1 senses the oxidative conditions within human macrophages during

infection through dimerization, resulting in activation of its DNA unwinding activity needed for DNA repair and other DNA metabolic pathways (49–51).

3.4 Results

3.4.1 The oligomeric state of UvrD1 is redox-dependent

Previous studies reported that UvrD1 exists exclusively as a monomer in solution (45, 48). However, upon purifying UvrD1 as described in the Methods, we observed two elution peaks from an S300 size exclusion column run at 4 ºC in Tris pH 8.0 at 25 ºC, 150 mM NaCl, 10% glycerol, and no DTT, consistent with the molecular weights of both monomer (85 kDa) and dimer (170 kDa) species (Fig. S3.1). This result is consistent with studies of *E. coli* UvrD, which exhibits a monomer-dimer-tetramer equilibrium (23, 26). To examine this more quantitatively, we performed analytical ultracentrifugation sedimentation velocity experiments in TRIS pH 8.0 at 25 ºC and 20% glycerol (from here on defined as Buffer A) with 75 mM NaCl and 2.5 M UvrD1. The continuous sedimentation coefficient (c(s)) distribution (52) shows three peaks that we assign to monomer, dimer and higher order oligomers (Fig. 3.1A, Supplemental Table 3.1). The positions of the peaks do not change with UvrD1 concentration indicating that each peak represents a single species; however, the amplitudes of the three peaks change with UvrD1 concentration as expected for a self-assembling monomer-dimer-oligomer system (Fig. S3.2). The oligomeric states of *E. coli* UvrD depend on the salt and glycerol concentrations, with the monomer population favored by higher salt and glycerol concentrations (26). We examined the salt dependence of the UvrD1 oligomeric state by sedimentation velocity at

2.5 M UvrD1 in a range of NaCl concentrations between 75 - 750 mM. Surprisingly, the ratio of monomer to dimer was relatively constant throughout the salt titration apart from the lowest salt concentrations where higher order oligomers were populated at the expense of the monomer population (Fig. 3.1B, S3.3).

In contrast, the addition of 1 mM DTT at 400 mM NaCl in Buffer A shifted the species fraction dramatically to favor the monomer (96%, Fig 3.1C). Even at the lowest NaCl concentration of 75 mM NaCl, the addition of 1 mM DTT resulted in a nearly uniform monomer population (90%, (Fig. S3.4). Thus, we hypothesized that UvrD1 dimerization is dependent on redox potential and reasoned that oxidative conditions should favor dimer formation just as reductive conditions favor the monomer species. To test this, we performed sedimentation velocity experiments in the presence of oxidizing agents after reduction by 1 mM DTT. As predicted, titration of hydrogen peroxide (H_2O_2) resulted in an increase in the dimer population (Fig. 3.1D, S3.5). At 2 mM H_2O_2, the population fraction of dimer saturated at ~25% and the addition of more hydrogen peroxide did not lead to more dimer formation.

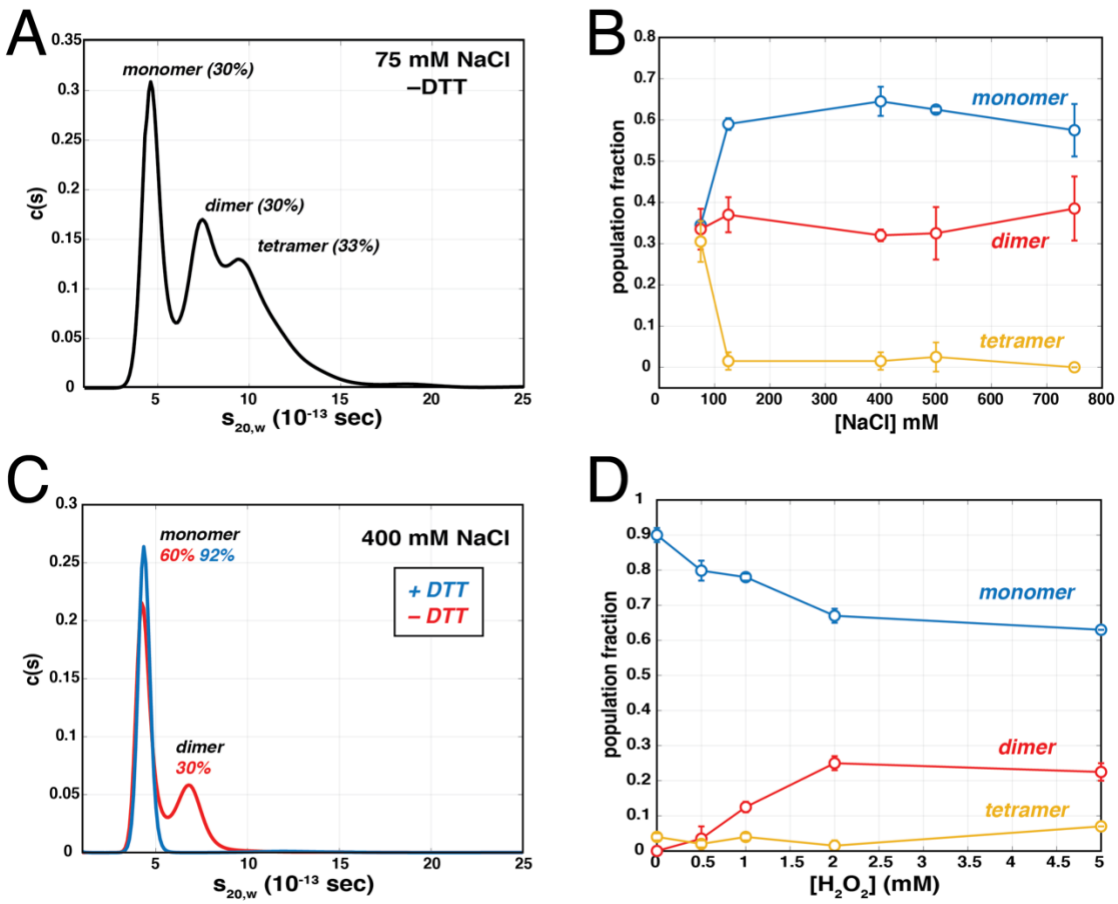


Figure 3.1: Oligomeric states of UvrD1. (A) Sedimentation velocity trace measured at 230 nm in Buffer A plus 75 mM NaCl in the absence of DTT reveals the presence of monomeric, dimeric, and tetrameric species. (B) Summary of results from AUC velocity experiments using 2.5 UvrD1 in the absence of DTT as a function of NaCl concentration (Fig. S3). From 75 mM to 750 mM NaCl, the fraction of monomer present in the monomeric state (blue), dimeric state (red), and tetrameric states (yellow) are shown. (C) The continuous distribution of species from AUC velocity runs measured at 280 nm with 400 mM NaCl in the presence and absence of 1 mM DTT. (D) After treatment of 2.5 mM WT UvrD1 in 75 mM NaCl with 1 mM DTT, a titration series of H_2O_2 from 0 mM to 5 mM was run in AUC velocity experiments (Fig. S5). Increasing concentrations of H_2O_2 result in a decrease in the fraction found in the monomeric state (blue) and an increasing fraction found in the dimeric state (red). Higher order oligomeric states (yellow) represented less than 10% under all conditions.

3.4.2 A 2B domain-2B domain disulfide bond is responsible for redox-dependent dimerization of UvrD1

The dependence of oligomerization on oxidation suggested a role for a thiol-containing amino acid such as methionine or cysteine. In particular, we considered that the potential of cysteines to form disulfide bonds could lead to the formation of dimeric and higher order oligomers. UvrD1 has three cysteine residues for which we estimated their approximate position by generating a threaded homology model of UvrD1 based on the structure of *E. coli* UvrD (PDB:3LFU, PHYRE2, Methods, Fig. 3.2A). This model shows that, while two cysteines appear buried within the 1A and 1B domains (C107 and C269), a third cysteine (C451) is surface exposed within the 2B domain. Surface calculations of our model with Chimera (53) confirmed this as only C451 possesses solvent exposed surface area in both open (based on *E. coli* PDB:3LFU structure) and closed (based on *G. stearothermophilus* PDB:3PJR structure) conformations (Fig. S3.6). We hypothesized that a disulfide bond between the 2B cysteines of two monomers was responsible for the redox-dependent dimerization. To test this, we constructed and purified a C451A mutant (which will we refer to as the 2B mutant) and examined whether it is able to form dimers. During purification of this construct, the S300 elution profile showed only a single peak consistent with a monomer in contrast to WT UvrD1 (Fig. 3.2B). Sedimentation velocity experiments confirmed this result, as the 2B cysteine mutant was monomeric in both the presence and absence of DTT (Fig. 3.2C, Supplemental Table 3.1). In contrast, a double mutant of the 1A and 1B domain cysteines (C107T/C269T, now referred to as the 1A1B double mutant) maintained the ability to form dimers (Fig. 3.2D, Supplemental Table 3.1).

The three cysteine residues found in UvrD1 are not conserved in *E. coli* UvrD, despite the presence of six cysteine residues (Fig. 3.3A), and *E. coli* UvrD does not display a redox-dependent dimerization (Fig. S3.7). However, the 2B domain cysteine identified

here is conserved across various Actinobacterial classes (Fig. 3.3B,C, S3.8). A particularly high, but not universal, conservation was found in the Corynebacteriales order which includes *Mtb* and other pathogenic bacteria (54, 55). In addition, we found the same sequence in the PcrA helicase found in one strain of the Firmicute, *Clostridioides difficile* (NCTC13750), which represents another important human pathogen that interacts with macrophages (56, 57).

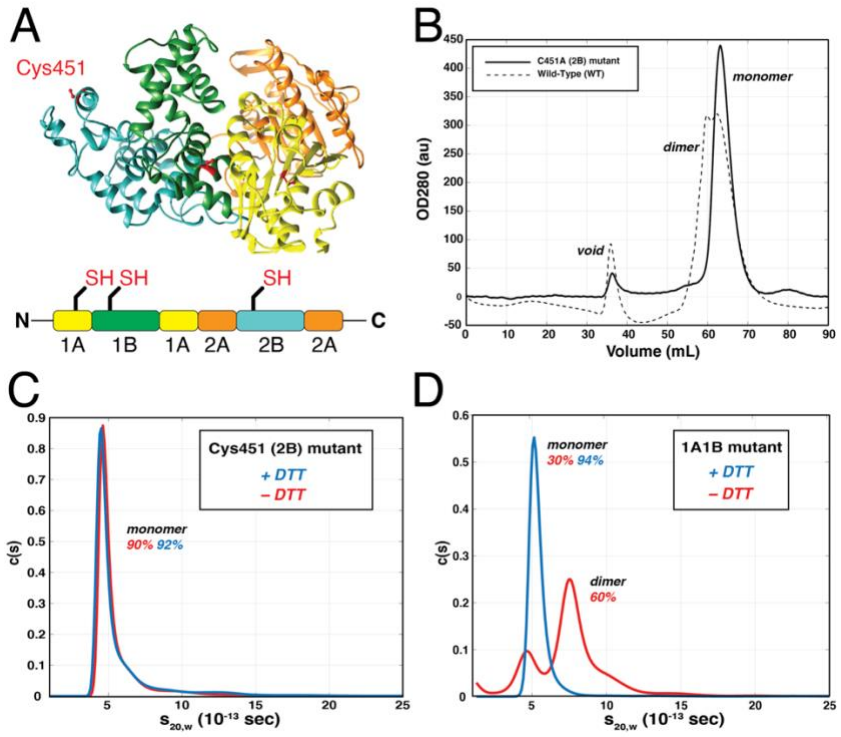


Figure 3.2: Redox-Dependence of UvrD1 dimerization is due to C451 in the 2B domain. (A) Predicted structure of UvrD1 from threading UvrD1 sequence on the E. coli UvrD structure (PDB: 3LFU). Domain organization is indicated as well as the position of the three cysteine residues described in the text. (B) Size exclusion chromatography (S300) of C451A mutant (solid) as compared to WT (dashed). Both constructs were run in Tris pH 8.0, 150 mM NaCl, 10% glycerol, and the absence of DTT as in Fig. S3.1. (C) AUC velocity experiments on the 2B mutant in Buffer A plus 75 mM NaCl and the presence (blue) and absence (red) of DTT indicate that the mutant loses the ability to dimerize. Each trace is an average of two runs and the population fractions for

monomer and dimer are indicated. (D) In contrast, AUC velocity of the 1A1B double mutant in 75 mM NaCl and the presence (blue) and absence (red) of 1 mM DTT indicates that this mutant retains and even enhances dimer formation. Each trace is an average of two runs and population fractions for monomer and dimer are indicated.

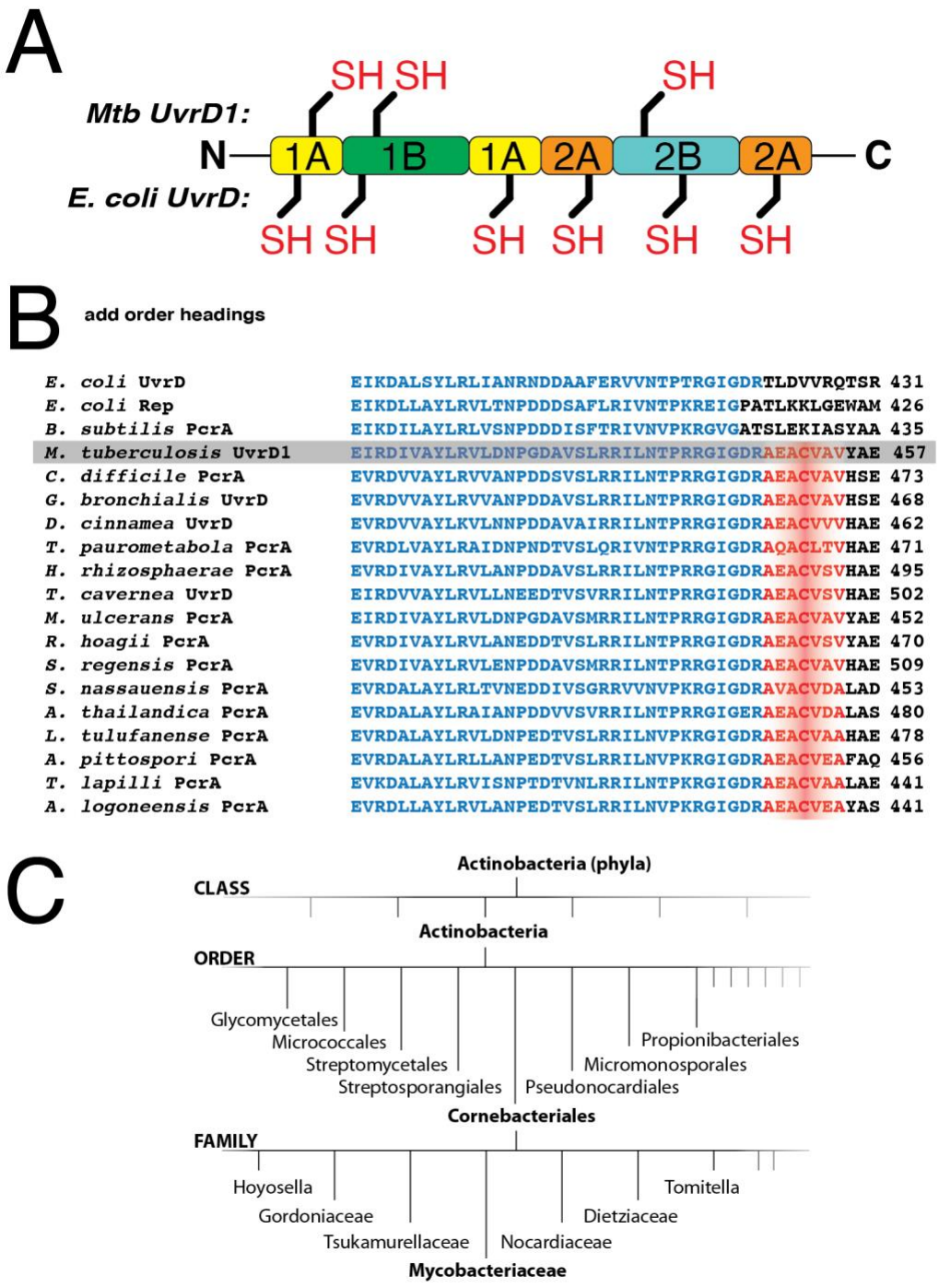


Figure 1.3: Sequence distribution across bacterial species. (A) Distribution of cysteine residues in

Mtb UvrD1 compared to E. coli UvrD. (B) Sequence alignment of the 2B domain region containing C451. The blue sequence is conserved across all UvrD-like family members while the sequence containing 2B cysteine residue (red) is distinct from E. coli and B. subtilis UvrD family enzymes but conserved in many Actinomycetes. See Supplemental information for full species names. (C) Orders and Families of Actinobacteria where the 2B cysteine can be found in available sequence data.

3.4.3 The dimer of UvrD1 is required for DNA unwinding activity.

Previous studies of UvrD1 suggested that the monomer possesses helicase activity (48). However, this conclusion was based on measurements of helicase activity performed in solution conditions distinct from those used to examine its oligomerization state. In particular, the analysis of oligomeric state was performed in the presence of 5 mM DTT while helicase assays were performed in buffer lacking DTT entirely (48). In other studies, UvrD1 was surmised to be a monomer based on sedimentation through a glycerol gradient and was reported to have feeble unwinding activity that was dramatically activated in the presence of *Mtb* Ku (45).

Given the observations of other UvrD-family helicases (23, 25, 26, 28, 58), we hypothesized that only the dimer of UvrD1 would be capable of unwinding DNA. To test this hypothesis, we used a stopped-flow assay to measure the time-dependence of UvrD1-catalyzed DNA unwinding (Fig. 3.4A). Specifically, we used a double-stranded 18 bp DNA with a single-stranded dT_{20} 3' flanking region (tail), with a Cy5 fluorophore on the 5' end of the tailed strand, and a black hole quencher (BHQ2) on the 3' end of the complementary strand as described previously (33, 34). In the double-stranded form, fluorescence from Cy5 is quenched due to the presence of the BHQ2 (59). Upon full unwinding, the strands are separated and the BHQ2 strand is trapped via an excess of

unlabeled complementary "trap" DNA resulting in an increase in Cy5 fluorescence. This excess of trap also serves to bind any UvrD1 that dissociates from the labeled template ensuring single-round turnover conditions (Fig. S3.9A). UvrD1 was pre-bound to the labeled DNA template and was loaded in one syringe. This solution was rapidly mixed with the contents of the other syringe consisting of an excess of trap strand, 5 mM Mg^{+2}, and 1 mM ATP. The fraction of DNA unwound as a function of time was calculated by comparing experimental traces to a positive control consisting of fully single-stranded Cy5-labeled DNA and a negative control in the absence of ATP (Fig. S3.9B).

In the absence of reducing agent, 200 nM WT UvrD1 can unwind ~27% of 2 nM duplex DNA in a single-round reaction (Fig. 3.4B, red) consistent with the fraction dimer in these conditions (Fig. 3.1A). In addition, the kinetics and final percent unwound show a duplex-length-dependence as is expected (Fig. S3.10A). Fits to an n-step model with a non-productive fraction (Fig. S3.10A,B) lead to an average unwinding rate of 64.8 ± 6.4 bp/s (Supplemental Table 2). Similarly, a lag time analysis (60, 61) yielded an estimate of the unwinding rate of 83.3 ± 12.3 bp/s (Fig. S10C,D). In contrast, in the presence of 1 mM DTT, which shifts the UvrD1 population to favor monomers, no unwinding is observed consistent with the UvrD1 dimer being required for unwinding activity (Fig. 3.4B, blue). In addition, the 2B mutant (C451A), which we have shown is an obligate monomer even under oxidative conditions, lacks helicase activity in both the presence and absence of DTT (Fig. 3.4C, cyan and purple). Consistently, the 1A1B double mutant is still able to unwind DNA in the absence of DTT (Fig. 3.4C, pink). In fact, the 1A1B double mutant unwinds a higher fraction of DNA compared to WT (~63%), consistent with the AUC results suggesting that a higher fraction of the enzyme is in the dimeric form in the absence of

DTT (Fig. 3.2D). Thus, the dimer fraction of UvrD1 formed via the 2B domain disulfide bond is required for DNA unwinding activity under single-turnover conditions. Experiments under multiple-turnover conditions (i.e., in the absence of trap) recapitulated these results as no unwinding was observed with monomeric UvrD1 (either WT+DTT or the 2B mutant) (Fig. S3.11). This further suggests that even multiple monomer binding events are not sufficient for unwinding activity.

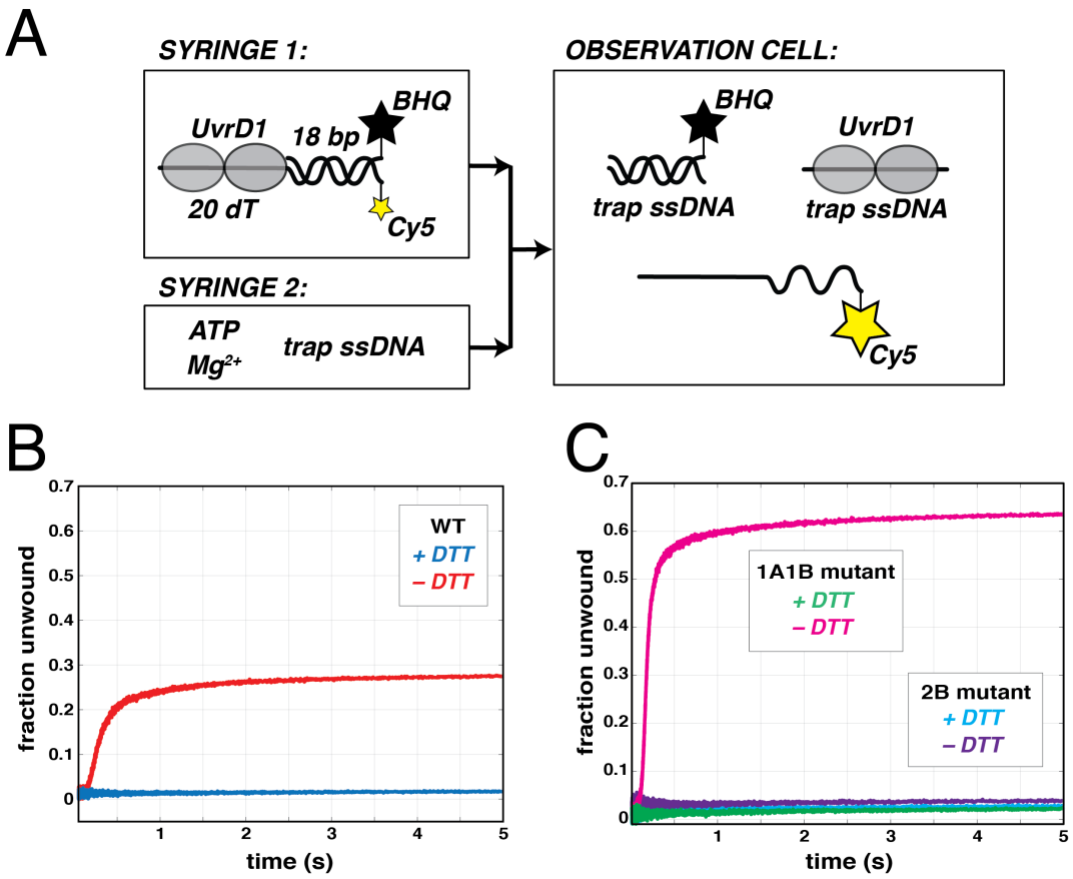


Figure 3.4: Unwinding activity is dependent on redox-dependent dimer. (A) Stopped-flow assay for monitoring DNA unwinding. One syringe is filled with UvrD1 that has been pre-equilibrated with a fluorescently labeled unwinding template consisting of an 18 bp duplex with a 3' single-stranded 20 dT tail. The 5' end of the loading strand is labeled with Cy5 and the 3' end of the other strand is labeled with a black hole quencher (BHQ). A second syringe is filled with ATP, Mg^{2+}, and single-stranded trap DNA. Upon mixing UvrD1 unwinds some fraction of the DNA resulting in a fluorescence enhancement due to the separation of the Cy5 from the BHQ. The excess trap

DNA binds both the BHQ labeled single-strand and any free UvrD1 to establish single-round turnover conditions. (B) Fraction of DNA template unwound with 200 nM wild-type UvrD1 in the presence (blue) and absence (red) of DTT. (C) Fraction of DNA template unwound with 200 nM 1A1B double and 2B mutants of UvrD1 in the presence (green and cyan) and absence (pink and purple) of DTT in Buffer A with 75 mM NaCl.

3.4.4 Both monomers and dimers of UvrD1 bind DNA

When considering why UvrD1 dimerization is required for DNA unwinding, we initially considered two hypotheses. One was that monomers are unable to bind DNA at the concentrations utilized and the second was that dimerization is required for helicase activation. To determine the nature of the DNA-bound species, we used sedimentation velocity to examine both WT UvrD1 and the 2B mutant in the absence of reducing agent and in the presence of fluorescently labeled 18 bp dT_{20} DNA. The results show that both monomers and dimers interact with the DNA and appear to bind the DNA in the same ratio as the free oligomeric forms. Specifically, in the presence of 1.5 uM WT UvrD1 and 1.5 M DNA, 30% of the DNA was bound by monomers and 31% was bound by dimers (Fig. 3.5A, red, Supplemental Table 3.3). In the presence of 1.5 uM of the 2B mutant, only monomers were bound to DNA (Fig. 3.5A, blue). Even at higher molar ratios of monomeric UvrD1 (i.e., the 2B mutant) to DNA, only monomers are observed bound to DNA suggesting that DNA binding alone does not stimulate dimerization (Fig. S3.12). In contrast, when WT UvrD1 in the absence of DTT is mixed with DNA possessing a shorter single stranded extension (dT_{10}), only monomers bind (Fig. 3.5B, Supplemental Table 3.3) and no DNA unwinding is observed (Fig. S3.13). This result is consistent with a model where each individual monomer interacts with the ssDNA in the context of the bound dimer as is seen with *E. coli* UvrD (23). Fluorescent anisotropy experiments with the unwinding

substrate DNA (18 bp dT_{20}) and either WT or 2B mutant UvrD1 yielded similar concentration dependencies of binding, consistent with the idea that the affinities of the monomeric and dimeric species are similar (Fig. 3.5C). Lastly, dissociation kinetics measured via a reduction in protein-stimulated fluorescence enhancement are also similar between WT and the 2B mutant (Fig. S3.14) and indicate that there is no significant difference in residence time at the junction.

The observation of DNA-bound monomers with similar binding properties as the dimer directly eliminates the possibility that monomers do not unwind because they do not bind at the DNA junction and suggests that some other property of the dimer is required for unwinding. This is consistent with other studies of UvrD-family enzymes as described in the Discussion.

At a constant redox potential established by the addition of 2 mM H_2O_2, the expected protein-concentration dependence of DNA unwinding is observed (Fig. 3.5D). Here, the fraction of DNA unwound saturates at ~27% consistent with the fraction of dimers under these conditions (Fig. 3.1D) further suggesting that the monomers and dimers compete approximately equally for this DNA substrate under these conditions.

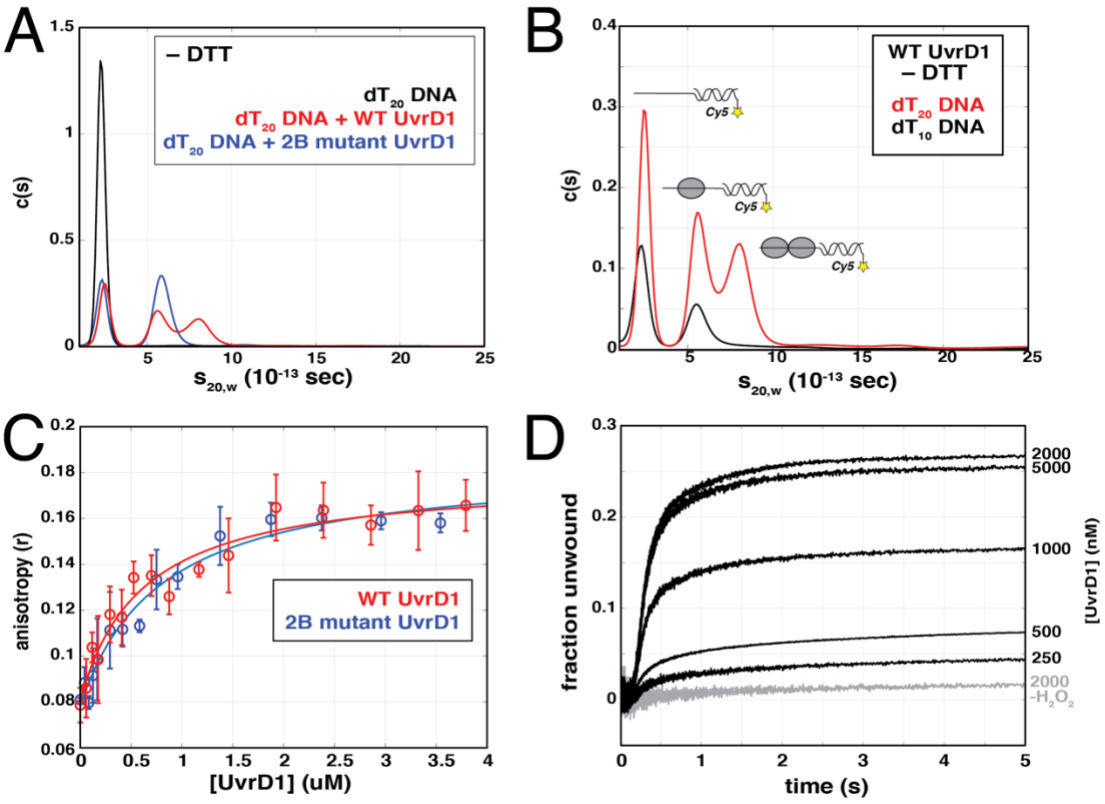


Figure 3.5: Monomer and dimer forms of UvrD1 both bind DNA unwinding template. (A) All experiments were performed in Buffer A with 75 mM NaCl. A Cy5-labeled DNA was used in sedimentation velocity experiments to specifically interrogate DNA-bound species. dT_{20} DNA alone (black), DNA and wild-type UvrD1 in the absence of reducing agent (red), and DNA and the 2B mutant in the absence of reducing agent (blue) are shown along with the bound species represented by each peak. (B) WT UvrD1 in the absence of reducing agent bound to dT_{20} and dT_{10} DNA templates. Both monomers and dimers bind dT_{20} (red) while only monomers bind dT_{10} (black). (C) Fluorescent anisotropy of a FAM labeled helicase template DNA (dT20 ssDNA tail with 18bp duplex) as a function of UvrD1 concentration. In the absence of DTT, both WT UvrD1 (red) and the 2B mutant (blue) displayed similar concentration dependencies of binding suggesting that they bind this template with similar affinity. (D) DNA unwinding traces as a function of UvrD1 concentration first treated with 1 mM DTT, followed by the addition of 2 mM H_2O_2. A control in the absence of oxidizing agent is shown for comparison ($-H_2O_2$).

3.4.5 Both UvrD1 monomers and dimers are single-stranded DNA translocases

As both monomers and dimers can bind to the DNA substrate (Fig. 3.5), we considered our second hypothesis that postulated that monomers are unable to translocate

along single-stranded DNA thus preventing helicase activity. To test this, we examined the translocation kinetics of UvrD1 on ssDNA. Since UvrD1 has a 3’ to 5’ DNA unwinding polarity, the kinetics of translocation were measured via stopped-flow assays (62, 63) by monitoring the arrival of UvrD1 at the 5′-end of a series of Cy3 5’-end-labeled oligodeoxythymidylate ssDNAs of different lengths (L = 20, 35, 45, 75, and 104 nucleotides) (Fig. 6A). Arrival of a translocating protein at the 5′-end results in an enhanced Cy3 fluorescence and subsequent dissociation of UvrD1 leads to a return of the signal to baseline. A heparin concentration of 1 mg/ml was added to prevent rebinding of free UvrD1 to the ssDNA, ensuring single-round conditions (Fig. S3.15A). In addition, negative controls in the absence of ATP showed no change in fluorescence (Fig. S3.15A) and experiments containing ATP and UvrD1 in the presence of 3’-labeled DNA showed only a decay in fluorescence consistent with translocation away from the dye (Fig. S3.15B). In conditions favoring dimeric UvrD1, the presence of ATP resulted in the expected length-dependent peaks of fluorescence indicative of UvrD1 translocation in the 3’ to 5’ direction (62, 63) (Fig. S3.16). However, monomeric UvrD1 generated either by the addition of DTT or the use of the 2B mutant also resulted in DNA-length-dependent changes in the fluorescence signal consistent with 3’ to 5’ ssDNA translocation (Fig. 3.6B,C). Global analysis using an n-step sequential model (Fig. S3.17 (62, 64) produced better fits for the 2B mutant data possibly because the WT UvrD1 still contains trace amounts of dimer. (Methods and Fig. S3.18) However, fits of both data sets yielded consistent estimates of the macroscopic translocation rate (mk_t) of 120 ± 5 nt/sec (WT+DTT) and 130 ± 10 nt/sec (2B-DTT) for monomeric UvrD1 which is very similar to the ssDNA translocation rate measured for E. coli UvrD monomers (62, 63). In addition, the estimated dissociation rate

(k_d) from the fitting analysis (4.1 ± 0.1 s^{-1}) was on the same order of magnitude as that obtained by experimentally measures (8.1 ± 0.2 s^{-1}) (Fig. S3.18). Other fit parameters are listed in Supplemental Table 3.4 (64).

The time-courses exhibited under monomeric UvrD1 conditions were distinct from those collected under oxidative conditions suggesting that the monomer and dimer populations exhibit distinct translocation kinetics (Fig. 3.6B,C, and Fig. S3.16). Fits using the percent fraction of monomeric and dimeric species and the translocation parameters obtained under monomeric conditions resulted in estimates for the ssDNA translocation properties of the UvrD1 dimer (Supplemental Table 3.4). Although analysis of the mixed dimer/monomer population was challenging due to the presence of multiple species and kinetic phases, taken together the data unequivocally show that both monomeric and dimeric UvrD1 translocate with 3’ to 5’ directionality along ssDNA.

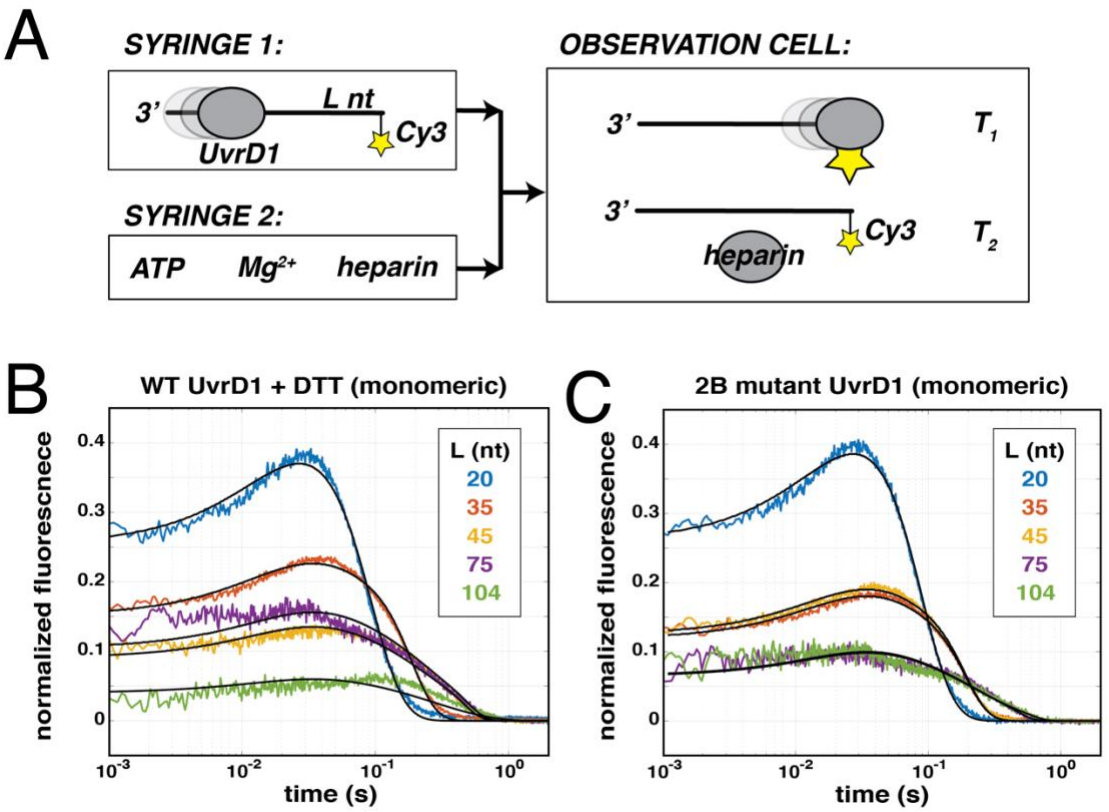


Figure 3.6: Monomeric UvrD1 translocates on single-stranded DNA. (A) The stopped-flow translocation assay. UvrD1 is allowed to equilibrate and bind to ssDNA templates of different lengths (L). These templates are labeled a the 5' end with Cy3. The protein DNA solution is mixed with buffer containing ATP, Mg^{2+}, and heparin to stimulate a single round of translocation. As the protein passes the label, the fluorescence increases (T1). As it dissociates, the fluorescence decreases, and free protein is bound by heparin (T2). (B) WT UvrD1 in the presence of 1 mM DTT produces traces consistent with translocation. The peak broadens and moves to longer times with increasing template lengths. (C) Monomeric UvrD1 generated by use of the 2B mutant, also shows clear peaks consistent with translocation. In both (B) and (C) experiments were conducted in Buffer A with 75 mM NaCl, where the traces were normalized to the average of 10 final plateau values observed in the raw data at each DNA length. Fits to an n-step sequential stepping model are shown in solid lines.

3.4.6 DNA unwinding activity is titrated by redox potential through dimer formation

We have shown that UvrD1 dimerization can be titrated via the addition of an oxidizing agent such as H_2O_2 (Fig. 3.1D). Furthermore, these dimers are, in the absence of other activators, required for DNA unwinding (Fig. 3.4). To determine the quantitative relationship between dimer fraction, DNA unwinding, and redox potential in millivolts

(mV), we performed both sedimentation velocity and helicase assays using H_2O_2 to titrate redox potential. The 1A1B double mutant was used for these titrations to ensure that any effects stemmed directly from the cysteine in the 2B domain. We tested two different concentrations 1 uM and 2 uM 1A1B double mutant in the presence of 1 mM DTT which results in over 95% monomer (Fig. 3.2D) and shows no DNA unwinding (Fig. 3.4C). We then titrated H_2O_2 from 0 – 5 mM which corresponds to redox potentials between -270 to 130 mV. As the redox potential became more positive (oxidizing), the fraction of DNA unwound increased (Fig. 3.7, S3.19 purple). As in the case of the WT UvrD1 (Fig. 3.1D), the fraction of UvrD1 1A1B double mutant dimer also increased with increasing H_2O_2 (Fig. 3.7, orange). In fact, there is a quantitative correlation between the fraction of DNA unwound and the fraction of UvrD1 mutant dimer, consistent with the hypothesis that UvrD1 helicase activity is stimulated via increasing positive redox potentials found under oxidative conditions (Fig. 3.7 and Fig. S3.19).

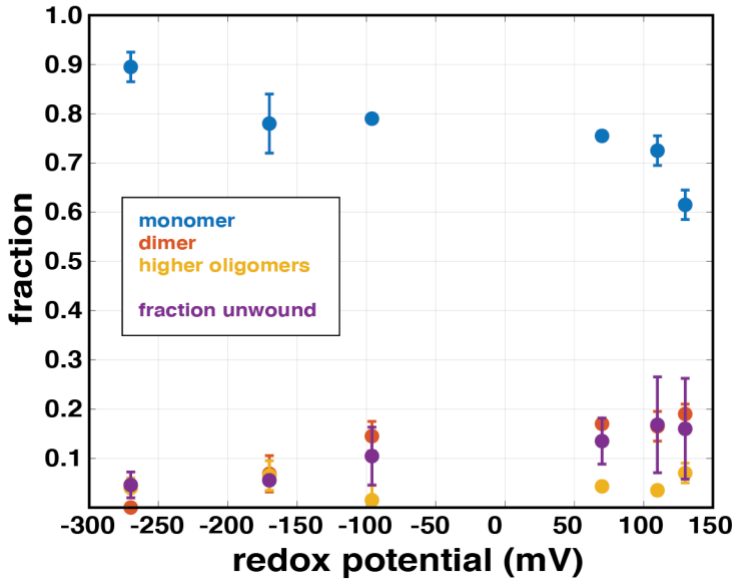


Figure 3.7: Unwinding activity correlates with dimer fraction and is titrated by redox potential. 1 M 1A1B double mutant UvrD1 was dialyzed in Buffer A at 75 mM NaCl, treated with 1 mM DTT, and then incubated with varying concentrations of H_2O_2 (0-5 mM). Redox potential was measured and the samples were subject to AUC and used for DNA unwinding assays. The fraction of monomer present in different oligomeric states and fraction DNA unwound are plotted as a function of redox potential as follows: monomer (blue), dimer (orange), higher oligomers (yellow), and fraction DNA unwound (purple).

3.5 Discussion

Processive DNA helicases are enzymes that couple ATP binding and hydrolysis to translocation on ssDNA and unwinding of double-stranded DNA and are involved in critical pathways throughout nucleic acid metabolism. The SF1 superfamily of helicases is the largest group of known helicases and includes UvrD, Rep and PcrA (19, 20). These helicases consist of two RecA-like domains (1A and 2A) and two accessory sub-domains (1B and 2B) (19, 20). Studies of SF1 helicases have shown that in the absence of accessory

factors or force exerted on the DNA, individual monomers possess ssDNA translocase activity, but not helicase activity; helicase activity requires at least a dimeric form of the enzyme. (21, 22, 27, 28, 30, 32, 33, 62–64). Yet, the dimerization interface is not known as crystal structures of UvrD and PcrA only reveal a monomer bound to DNA (65–68).

Previous studies of *Mtb* UvrD1 concluded that it is monomeric as shown by size exclusion chromatography (SEC) and equilibrium sedimentation and is capable of unwinding dsDNA in this form (45, 48). However, in one case, SEC and equilibrium sedimentation were performed in the presence of the reducing agent DTT, whereas DNA unwinding assays were performed in its absence (48). In the other case, helicase activity of UvrD1 was attributed to monomers and was dependent on its binding partner, Ku (45). When we purified *Mtb* UvrD1 without reducing agent it eluted as two peaks on SEC corresponding to the expected molecular weights of both monomer and dimer (Fig. 1A). AUC experiments show that, while salt concentration had almost no effect on the amount of the dimeric form, the addition of DTT results in a dramatic shift to a single monomeric peak suggesting that *Mtb* UvrD1 undergoes a cysteine dependent dimerization (Fig. 3.1). We also observe a higher oligomeric species increasing in a concentration dependent manner in sedimentation velocity experiments performed under oxidizing conditions. We do not observe *Mtb* UvrD1 tetramer binding to the tailed construct used in our unwinding assays (Fig. 3.5A), but we cannot eliminate the possibility that the tetrameric form could bind to DNA possessing longer single-stranded tails.

We identified a critical cysteine residue in the 2B domain of UvrD1 (C451) that is required for UvrD1 dimerization (Fig. 3.2). More specifically, the C451A mutation abrogated both dimerization and helicase activity pointing to the existence of a 2B-2B

disulfide bonded dimer. The 2B domain of UvrD-like helicases has been described as an auto-inhibitory domain (28, 32) that can adopt a range of rotational conformational states relative to the rest of the protein (30, 31, 34, 36, 69). The different 2B domain conformational states of the monomer are influenced by salt concentration, DNA binding, enzyme dimerization and the binding of accessory protein factors (26, 34, 69, 70). For example, crystal structures of Rep bound to ssDNA showed Rep monomer bound to DNA in two different conformations (66). In one of the conformations the 2B domain is in an "open" conformation, while in the other, the 2B domain is reoriented by a 130° swivel motion around a hinge region to contact the 1B domain. This swiveling motion closes the binding groove located between 1A, 1B, and 2A domains around the DNA template. Consistently, single molecule and ensemble FRET studies have also shown that the 2B domain of a monomer can be in closed or open conformations (31, 34, 69–71). Furthermore, removing the 2B domain from Rep causes the Rep monomer to gain helicase activity (32). These observations and others have led to the hypothesis that the 2B domain is auto-inhibitory for monomer helicase activity but serves as the interface between subunits within the functional Rep and UvrD dimers (19, 28, 30, 72). In this context, dimerization reorients the 2B domain to relieve its inhibitory properties resulting in active dimers.

The formation of a 2B-2B disulfide bond in the case of *Mtb* UvrD1 fits well in this model where 2B-2B driven dimerization results in an active helicase conformation. Further support for this idea can be found in the specific location of C451. Mutations to 2B domain threonine residues such as T426 of *B. stearothermophilus* PcrA and T422 of *E. coli* UvrD disrupt helicase activity (73). This threonine residue is conserved in non-Actinobacteria

UvrDs, is absent from UvrD-like proteins containing the cysteine residue described here and is located 2 amino acids upstream to the C451 of *Mtb* UvrD1. The similar location of these residues suggests that UvrD-like helicases share a common 2B-2B dimerization interface and that mutating this threonine may destabilize the dimeric form of the enzyme. Interestingly, the 2B domains of the SF1 family members RecB and RecC interact in a RecBCD helicase complex (74). We aligned the UvrD1 2B domain sequence with the 2B domain sequences of *E. coli* RecB and RecC which lack a 2B cysteine equivalent to C451. When we mapped residues that aligned close to the 2B cysteine onto the structure of RecBCD (PDB:5LD2) (74, 75), we observed that these regions in RecB and RecC are only 10-15Å apart. Again, this is consistent with our identification of the region surrounding the 2B cysteine of UvrD1 as an interface for 2B-2B domain-based activation of UvrD1-like helicases.

As noted, the 2B cysteine we have identified in *Mtb* UvrD1 appears unique to certain classes of Actinobacteria (Fig. 3). The 2B domain of *E. coli* UvrD does contain a cysteine residue at a different location (C441). However, experiments with *E. coli* UvrD showed no effect on oligomeric state with changes in redox potential (Fig. S7).

Another pathway for stimulation of unwinding by UvrD-family enzymes requires the binding of activating partners. The mismatch-repair protein MutL can activate the monomer helicase activity of *E. coli* UvrD as well as stimulate the activity of UvrD dimers and activation is accompanied by a change in the rotational state of the 2B domain (33). This interaction also leads to enhanced processivity which enables UvrD to unwind longer stretches of DNA when functioning with MutL (32). Similarly, the accessory factor PriC can activate the Rep monomer helicase and stimulate the Rep dimer helicase (76). By

analogy, the NHEJ factor, *Mtb* Ku, has been reported to bind to the C-terminal region of UvrD1 (45) suggesting a role for UvrD1 in double-strand-break repair (DSB). As the results from our study show that UvrD1 can exist as a monomer or dimer depending on the redox potential, we are currently investigating whether Ku can activate the UvrD1 monomer helicase and/or stimulate dimer activity and whether it does this via modulation of the 2B domain conformation.

Mtb actually has two UvrD-family members in its genome, UvrD1 and UvrD2, with UvrD1 being the homologue to *E. coli* UvrD (45–47, 77) and UvrD2 consisting of an N-terminal SF1 helicase motor linked to an HRDC (helicase and Rnase D C-terminal) domain and tetracysteine motif domains (44, 78). The cysteines within the tetracysteine motif bind zinc and the domain is required for helicase activity of UvrD2 *in vitro* (44). However, in this case the activity appears to be dependent on the presence of the domain and not the presence of the individual cysteines themselves. Interestingly, while WT UvrD2 does not show a dependence on Ku, a truncated construct lacking the tetracysteine domain can be activated by Ku (44). Therefore, while both UvrD1 and UvrD2 utilize cysteine residues for helicase activity, they do so via distinct biochemical mechanisms.

During infection, *Mtb* resides within alveolar macrophages and neutrophils where it is exposed to reactive oxygen species (ROI) and reactive nitrogen intermediates (RNI) that cause DNA damage (49, 79, 80). *Mtb* lacks mismatch repair (81) and the response of *Mtb* to these insults likely involves NER pathways (79). During NER, UvrD1 unwinds the damaged DNA strand to remove bulky lesions that have been recognized by UvrA/UvrB and excised by UvrC (2). Analysis of gene expression data have shown UvrA to be highly expressed during oxidative stress (82). In addition, both *uvrD1/uvrB* and *uvrD1/uvrA*

double mutants in *M. smegmatis* have been shown to be more sensitive to tertiary butyl hydroperoxide and acidified nitrite than wild-type strains (46, 83). All of these observations suggest a role of NER enzymes, including UvrD1, during oxidative stress in *Mtb*. This is distinct from *E. coli*, in which repair of RNI and ROI-induced DNA damage is accomplished by base-excision repair and homologous recombination (84). Thus, the redox dependent dimerization of UvrD1 we report here may represent an important mechanism in *Mtb* underlying the repair of oxidative-dependent DNA damage during infection.

3.6 Materials and Methods

Cloning, overexpression, and purification of *Mycobacterium tuberculosis* UvrD1: *Mtb* UvrD1 (Rv0949) from H37Rv was cloned in the expression vector pET with SUMO-His tag at the N-terminus and Kannamycin resistance. It was PCR amplified with BamH1 at the 5' and HindIII at the 3'end. UvrD1 2B domain cysteine to alanine substitution and 1A1B double cysteine to threonine substitution mutations were introduced into the *Mtb* UvrD1 plasmid by PCR amplification using primers and site directed mutagenesis kit Agilent product 200521. Sequences for all primers and a list of plasmids can be found in Supplemental Table 5. The inserts of all UvrD1 plasmids were sequenced to exclude the acquisition of unwanted coding changes during amplification or cloning. The pET-*Mtb* UvrD1 plasmids were transformed into *Escherichia coli* BL21(DE3). Cultures (3 L) were grown at 37 °C in a Luria-Bertani medium containing 50 mg/ml kanamycin until the A_{600} reached ~0.5. The cultures were chilled on ice for about 1 hour, and the expression of recombinant protein was induced around 0.55 OD with 0.25 mM isopropyl-β-D-

thiogalactopyranoside (IPTG), followed by incubation at 16 °C for 16 h with constant shaking. The cells were harvested by centrifugation, and the pellets were either stored at -80 °C or used for subsequent procedures that were performed at 4 °C. The bacterial cells from the 3-liter culture were resuspended in 50-75 ml of lysis buffer (50 mM Tris-HCl, pH 7.5, 0.25 M NaCl, 10% sucrose). Lysozyme and Triton X-100 were added to final concentrations of 1 mg/ml and 0.1%, respectively. At the time of lysis, a complete EDTA-free protease inhibitor cocktail (Sigma: product 118735800) was added to the lysate. Next the lysates were sonicated and insoluble material was removed by centrifugation at 14K for 45 minutes. The soluble extracts were applied to 2-ml columns of nickel-nitrilotriacetic acid-agarose (Ni-NTA) (QIAGEN catalog no. 30210) that had been equilibrated with lysis buffer without protease inhibitors. The columns were washed with 10X column volume of wash buffer (50 mM Tris-HCl, pH 8.0, 0.25 M NaCl, 0.05% Triton X-100, 10% glycerol) and then eluted stepwise with wash buffer containing 50, 100, 200, 500, and 1000 mM imidazole. The polypeptide compositions of the column fractions were monitored by SDS-PAGE. The his-SUMO-tagged UvrD1 polypeptides were recovered predominantly in the 100- and 200-mM imidazole eluates. Fractions containing the UvrD1 protein were pooled and His-tagged Ulp1 protease was added (at a ratio of 1:500 wt/wt protease per protein) and dialyzed against dialysis buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.1% Triton X-100, 10% glycerol) containing 150 mM NaCl overnight. The SUMO-cut UvrD1 was then incubated with Ni-NTA agarose for about 3 hours and the untagged UvrD1 (cleaved) was recovered in the flow-through and wash fractions. After pooling and concentrating the fractions by VIVASPIN centrifugal filters (30 kDa cutoff: product Sartorius VS2021), the protein was loaded on heparin HiTRAP column (Cytiva: product 17040701) (5 ml x 2) pre-

equilibrated with dialysis buffer containing 150mM NaCl. Upon running a linear gradient from 200 to 800 mM NaCl, the protein eluted at approximately 400 mM NaCl. The fractions with a single band on a reducing SDS-PAGE corresponding to the molecular weight of a monomer were pooled together and concentrated to load on the S300 sizing column (HiPrep 16/60 Sephacryl S-300 HR column, Cytvia, product 17116701) in buffer 150mM NaCl, 10% glycerol and TRIS pH 8.0 at 25 °C without DTT. The peak fractions were pooled and stored in -80 °C at concentrations about 15-20 mM. Both the 2B mutant and 1A1B double mutant were overexpressed and purified in a manner similar to the wild-type UvrD1 protein.

Homology Modeling of *Mtb* UvrD1: The predicted structure of *Mtb* UvrD1 was obtained using PHYRE2 (85) by submitting the amino acid sequence of *Mtb* UvrD1. The PDB reference used for modeling open structure of *Mtb* UvrD1 was *E. coli* UvrD was 3LFU and for modeling closed structure is *B. stearothermophilus* PcrA helicase complex 3PJR and methods for SASA calculations done in Chimera (53).

Phylogeny analysis of *Mtb* UvrD1: The protein sequence of *Mtb* UvrD1 obtained from uniprot (https://www.uniprot.org/uniprot/P9WMQ1.fasta) was used to Blast against various genus and species of classes of Phylum Actinobacteria (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Two-three genus of orders in which the cysteine in 2B domain is conserved were chosen and aligned in Clustal omega (86) and the phylogenetic tree was plotted using NJ plot (87).

Analytical Size Exclusion Chromatography: 1 mL of *Mtb* UvrD1 (30 mM) was injected in a S300 gel filtration column at a flow rate of 0.25 mL/min, monitoring absorbance at 280 nm for a measure of the elution volume (*V*) in buffer 150mM NaCl, 10% glycerol and TRIS pH 8.0. The value of V_e/V_o was interpolated using the generated standard curve (Bio-Rad gel filtration standard; product #1511901) to yield the estimated molecular weight of *Mtb* UvrD1 monomer 90 kDa and dimer 170 kDa fractions.

DNA Substrates: Single stranded DNA which are either labeled with Cy5, Cy3, FAM or BHQ2 were ordered from IDT. For annealing of the oligos the Cy5 labeled at 5'end of the single stranded DNA was mixed with an equimolar concentration of unlabeled complementary strand or complementary strand labeled with BHQ2 in 10 mM Tris pH 8.0, 50 mM NaCl, followed by heating to 95 °C for five minutes and slow cooling to room temperature.

Synthesis of poly-dT: The homodeoxypolynucleotide, poly-dT substrate was used to measure dissociation rate from internal sites of ssDNA. Since the poly-dT from commercial sources is polydisperse, we prepared samples using enzyme terminal deoxynucleotidyl transferase (TdTase) from calf thymus gland to catalyze polymerization of deoxynucleotide triphosphate into poly-dT of more well-defined lengths (88). The protocol includes mixing dT(100) with potassium cacodylate buffer, potassium chloride, Cobalt chloride, Inorganic pyrophosphatase (ThermoFisher Scientific #EF0221), deoxythymidine-5'-triphosphate dTTP (ThermoFisher Scientific #R0171) and terminal deoxynucleotidyl transferase (TdTase ThermoFisher Scientific #10533065). The reaction was kept at room temperature for 3-4 days and poly-dT was purified using phenol

chloroform extraction and suspending the air-dried pellet in water. The weight average length of the poly-dT was determined by measuring the weight average sedimentation coefficient by boundary sedimentation velocity experiments using AUC. The weight average length determined from this method is 964 nucleotides and was determined using method from the measured weight average sedimentation coefficient on poly-U (89).

Analytical ultracentrifugation: The analytical ultracentrifugation sedimentation velocity experiments were performed using a Proteome Lab XL-A analytical ultracentrifuge equipped with an An50Ti rotor (Beckman Coulter, Fullerton, CA). The sample (380 µl) and buffer (410 µl) were loaded into each sector of an Epon charcoal-filled two-sector centerpiece. All experiments were performed at 25 °C and 42,000 rpm. Absorbance data were collected by scanning the sample cells at intervals of 0.003 cm, monitoring either at 230 nm or 280 nm depending on protein concentration to maintain an absorbance signal between 0.1 and 1. Both the DNA and protein samples were dialyzed in buffer 75 mM NaCl, 20% glycerol and 10 mM TRIS pH 8.0 except the salt titration where the protein was dialyzed in different salts at 20% glycerol at 10mM TRIS pH 8.0.

Continuous sedimentation coefficient distributions, $c(s)$, were calculated using SEDFIT(6), truncating the fit at 7.0 radial position to avoid contributions of glycerol buildup (90). This analysis yielded individual sedimentation coefficients for each monomer, dimer, and tetramer species as well as a weighted average frictional coefficient (f/f_o) for the entire distribution (Table S1). Calculated sedimentation coefficients were converted to 20 ^{o}C water conditions ($s_{20,w}$) according to:

$$s_{20,w} = s_{\exp} \frac{\eta_{\exp}}{\eta_{20,w}} \left(\frac{1-\bar{v}_{20}\,\rho_{20,w}}{1-\bar{v}_{\exp}\,\rho_{\exp}} \right) \qquad (3.1)$$

where $\rho_{20,w}$ and $\eta_{20,w}$ are density and viscosity of water at 20 °C, $\rho_{\exp}$ and $\eta_{\exp}$ are density and viscosity of the buffer at the experimental temperature of 25 °C , and $\bar{v}_{20}$ and $\bar{v}_{\exp}$ are partial specific volumes of the protein at 20 °C and at 25 °C. Buffer densities, ($\rho_{\exp}$) and viscosities ($\eta_{\exp}$) were calculated from buffer composition using SEDNTERP (91). Partial specific volumes ($\bar{v}_{\exp}$) for *Mtb* UvrD1 and point mutations were calculated in SEDNTERP using the amino acid composition. Integration of the entire c(s) distribution vs. the integration of an individual sedimentation species was performed and used to calculate the population fraction (92).

For AUC experiments done in the presence of Cy5 labeled DNA, the absorbance signal was collected by scanning the sample cells at 650 nm. Partial specific volumes ($\bar{v}_{\exp}$) for labeled DNA and the UvrD1-DNA complex were calculated according to:

$$\bar{v} = \frac{(\sum_{i=1}^{n} n_i M_i \bar{v}_i)}{\sum_{i=1}^{n} n_i M_i} \qquad (3.2)$$

For AUC experiments conducted at different redox potentials, the protein was first dialyzed in Buffer A with 75 mM NaCl and then 1 mM DTT was added at a respective concentration and then titrated with a range of H_2O_2 from 0.5 to 2 mM. Redox potential was measured via Metler Toledo Redox micro electrode product #UX-35902-33. After H_2O_2 treatment protein was incubated at room temperature for two hours before performing AUC.

Stopped-flow double-stranded (ds) DNA unwinding assay: All stopped-flow experiments were carried out at 25°C using an Applied Photo physics instrument SX-20, total shot volume 100 µl, dead time 1 ms. All experiments were carried out in Buffer with TRIS pH 8.0, 75 mM NaCl, and 20% glycerol in the absence or presence of 1 mM DTT. Cy5 fluorophore was excited using 625 nm LED (Applied Photo physics Ltd., Leatherhead, UK) and its fluorescence emission was monitored at wavelengths >665 nm using a long-pass filter (Newport Optics). The traces represent the average of 5 independent shots and at least two different protein purifications. In this assay, a double-stranded 18-basepair DNA with a T20/T10 tail with Cy5 fluorophore is attached to the long strand and the black hole quencher (BHQ_2) attached to the short strand (Supplemental Table 6). The concentrations mentioned are the final after mixing the contents of both syringes. UvrD1 (200 nM) is incubated with 2 nM DNA in one syringe which is then rapidly mixed with the contents in another syringe that consists of (TRAP a complementary DNA strand in excess protein (25X, 5M), Mg^{+2} (5 mM), and ATP (1 mM). The concentrations of ATP and Mg^{+2} used were determined to be optimal for *Mtb* UvrD1 unwinding assays (48). DNA strand separation is accompanied by an increase in the fluorescence signal. The unwinding signal is normalized to the signal from positive and negative control to get fraction of DNA unwound. The positive control is 2 nM double stranded DNA with BHQ2 on the short and Cy5 on the long strand that is denatured in the presence of TRAP and 200 nM UvrD1 is added at room temperature to get maximum fluorescence signal. The negative control is the average fluorescence value recorded for the fully annealed DNA (2 nM) with UvrD1 (200 nM) shot against buffer alone. To get DNA unwinding with a change in redox potential the protein was dialyzed treated with DTT and titrated with H_2O_2 in similar way

as done for AUC experiments and incubated with DNA. This protein DNA mix was then used for unwinding experiments.

Stopped flow single stranded (ss) DNA translocation assays: The kinetics of UvrD1 monomer and dimer translocation was examined in a stopped-flow experiment by monitoring the arrival of UvrD1 of a series of oligodeoxythymidylates length (L= 20, 35, 45, 64, 75, 94, and 104) nucleotides labeled at the 5′-end with Cy3 (Supplemental Table 6) (62). Cy3 fluorescence was excited using a 535 nm LED with a 550 nm short-pass cut-off filter and emission was monitored at >570 nm using a long-pass filter (Newport Optics). UvrD1 was pre-incubated with ssDNA in one syringe and reactions were initiated by 1:1 mixing with 1 mM ATP, 5 mM $MgCl_2$ and heparin at a concentration of 1 mg/ml to prevent rebinding of UvrD1 (50 nM) to DNA (100 nM). Excess DNA to UvrD1 ratio was used to prevent binding of more than one UvrD1 monomer on DNA (62). Global analysis of time courses using the n-step sequential model was done to calculate various translocation parameters. For making heparin solution to be used as a TRAP for protein ensuring single round conditions, heparin sodium salt (porcine intestinal mucosa, Millipore Sigma #H3393) was dialyzed into Buffer A plus 75 mM NaCl and concentrations were determined by an Azure A standard curve (93).

Tryptophan fluorescence-based dissociation kinetics: Dissociation kinetics of UvrD1 were monitored by the increase in UvrD1 tryptophan fluorescence, excited using a 290 nm LED (Applied Photophysics Ltd., Leatherhead, UK) and monitoring emission at >305 nm using a long-pass filter (check if this is from Newport or AP). The observed dissociation rate from internal ssDNA sites for UvrD1 monomer (with 1 mM DTT) and

dimer (no DTT) was measured using $(dT)_{100}$ and poly(dT) oligos (average length of 964 nucleotides; see synthesis of poly-dT section), respectively. In one syringe UvrD1 (100 nM) was added with DNA (50 nM) (concentrations listed are after equal volume mixing). In another syringe ATP, $MgCl_2$ and heparin were added at the same concentrations as translocation assays and the observed dissociation kinetic traces were best fit to a single exponential using ProData Viewer (Applied Photophysics). All experiments were performed at 25 ^oC in Buffer A with 75 mM NaCl and represent the average of 5 independent shots (88).

ssDNA translocation and unwinding fitting analysis: The translocation and the unwinding data was fit using the code from https://github.com/ordabayev/global-fit and unwinding and translocation rates are calculated. Globalfit is a wrapper around lmfit https://lmfit.github.io/lmfit-py/ providing an interface for multiple curves fitting with global parameters. Python 3 is installed via Anaconda along with modules like numpy, scipy, matpotlib, lmfit, emcee, corner, os and pandas and then globalfit model is used to fit the data for unwinding using n-step unwinding model and translocation using a two-step dissociation model (64).

Steady-State Anisotropy Measurements: All fluorescence titrations were performed using a spectrofluorometer (ISS, Champaign, IL) equipped with Glan-Thompson polarizers. Measurements of the anisotropy and total fluorescence intensity of FAM-labeled double stranded DNA 18bp with T20 single stranded tail were recorded using excitation and emission wavelengths of 490 and 522 nm, respectively, using

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}} \qquad (3.3)$$

where $I_{TOT} = I_{VV} + 2GI_{VH}$ and G is the G factor (94). The recorded value of G factor remained between 0.85 to .9 throughout the titrations. Titrations were performed using a Starna cells cuvette catalog number 16.100F-Q-10/Z15 with dimensions 12.5 x 12.5 x 45 mm and pathlength 1cm. The protein was mixed 3-4 times with DNA during titrations and let it sit for 5 minutes before recording anisotropy values. The total volume of added protein was 30% of the initial volume and as a control the dilution with buffer of up-to 30% for DNA alone sample did not change the anisotropy value of the DNA. All titrations were conducted in Buffer A at 25 °C in 75 mM NaCl. The data represents average from three independent experiments for WT-DTT and 2 independent experiments for the 2B mutant.

Acknowledgments

This work was supported by NIH R01 GM134362 to EAG, NIH R35 GM136632 to TML, and by NIH T32 AI007172 to AC. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

3.7 References

1. A. Sancar, DNA Excision Repair. *Annu Rev Biochem* **65**, 43–81 (1996).

2. Y. Peng, H. Wang, L. Santana-Santos, C. Kisker, B. van Houten, "Nucleotide Excision Repair from Bacteria to Humans: Structure–Function Studies" in Chemical Carcinogenesis., (2011), pp. 267–296.

3. C. Kisker, J. Kuper, B. van Houten, Prokaryotic nucleotide excision repair. *Csh Perspect Biol* **5**, a012591–a012591 (2013).

4. K. Howan, *et al.*, Initiation of transcription-coupled repair characterized at single-molecule resolution. *Nature* **490**, 1–4 (2012).

5. A. J. Smith, N. J. Savery, RNA polymerase mutants defective in the initiation of transcription-coupled DNA repair. *Nucleic Acids Res* **33**, 755–764 (2005).

6. B. Pani, E. Nudler, Mechanistic insights into transcription coupled DNA repair. *Dna Repair* **56**, 42–50 (2017).

7. V. Kamarthapu, E. Nudler, Rethinking transcription coupled DNA repair. *Curr Opin Microbiol* **24**, 15–20 (2015).

8. G. Kokic, *et al.*, Structural basis of TFIIH activation for nucleotide excision repair. *Nat Commun* **10**, 2885 (2019).

9. V. Oksenych, B. B. de Jesus, A. Zhovmer, J. M. Egly, F. Coin, Molecular insights into the recruitment of TFIIH to sites of DNA damage. *Embo J* **28**, 2971–2980 (2009).

10. E. Compe, J. M. Egly, TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Bio* **13**, 343–354 (2012).

11. R. R. Iyer, A. Pluciennik, V. Burdett, P. L. Modrich, DNA Mismatch Repair: Functions and Mechanisms. *Chem Rev* **106**, 302–323 (2006).

12. C. Bruand, S. D. Ehrlich, UvrD-dependent replication of rolling-circle plasmids in Escherichia coli. *Mol Microbiol* **35**, 204–210 (2000).

13. G. F. Moolenaar, C. Moorman, N. Goosen, Role of the Escherichia coli Nucleotide Excision Repair Proteins in DNA Replication. *J Bacteriol* **182**, 5706–5714 (2000).

14. M. Florés, N. Sanchez, B. Michel, A fork-clearing role for UvrD. *Mol Microbiol* **57**, 1664–1675 (2005).

15. R. C. Heller, K. J. Marians, Non-replicative helicases at the replication fork. *Dna Repair* **6**, 945–952 (2007).

16. H. M. Arthur, R. G. Lloyd, Hyper-recombination in uvrD mutants of Escherichia coli K-12. *Mol Gen Genetics Mgg* **180**, 185–191 (1980).

17. X. Veaute, *et al.*, UvrD helicase, unlike Rep helicase, dismantles RecA nucleoprotein filaments in Escherichia coli. *Embo J* **24**, 180–189 (2005).

18. V. Petrova, *et al.*, Active displacement of RecA filaments by UvrD translocase activity. *Nucleic Acids Res* **43**, 4133–4149 (2015).

19. T. M. Lohman, E. J. Tomko, C. G. Wu, Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nat Rev Mol Cell Bio* **9**, 391–401 (2008).

20. K. D. Raney, A. K. Byrd, S. Aarattuthodiyil, DNA Helicases and DNA Motor Proteins. *Adv Exp Med Biol* **767**, 17–46 (2012).

21. M. S. Dillingham, D. B. Wigley, M. R. Webb, Demonstration of Unidirectional Single-Stranded DNA Translocation by PcrA Helicase: Measurement of Step Size and Translocation Speed. *Biochemistry-us* **39**, 205–212 (1999).

22. M. S. Dillingham, D. B. Wigley, M. R. Webb, Direct Measurement of Single-Stranded DNA Translocation by PcrA Helicase Using the Fluorescent Base Analogue 2-Aminopurine. *Biochemistry-us* **41**, 643–651 (2001).

23. N. K. Maluf, C. J. Fischer, T. M. Lohman, A Dimer of Escherichia coli UvrD is the active form of the helicase in vitro. *J Mol Biol* **325**, 913–935 (2003).

24. E. J. Tomko, C. J. Fischer, A. Niedziela-Majka, T. M. Lohman, A nonuniform stepping mechanism for E. coli UvrD monomer translocation along single-stranded DNA. *Mol Cell* **26**, 335–347 (2007).

25. A. Niedziela-Majka, M. A. Chesnik, E. J. Tomko, T. M. Lohman, Bacillus stearothermophilus PcrA monomer is a single-stranded DNA translocase but not a processive helicase in vitro. *J Biol Chem* **282**, 27076–27085 (2007).

26. N. K. Maluf, T. M. Lohman, Self-association Equilibria of Escherichia coli UvrD Helicase Studied by Analytical Ultracentrifugation. *J Mol Biol* **325**, 889–912 (2003).

27. K. S. Lee, H. Balci, H. Jia, T. M. Lohman, T. Ha, Direct imaging of single UvrD helicase dynamics on long single-stranded DNA. *Nat Commun* **4**, 1878 (2013).

28. K. M. Brendza, *et al.*, Autoinhibition of Escherichia coli Rep monomer helicase activity by its 2B subdomain. *P Natl Acad Sci Usa* **102**, 10076–10081 (2005).

29. S. Arslan, R. Khafizov, C. D. Thomas, Y. R. Chemla, T. Ha, Engineering of a superhelicase through conformational control. *Science* **348**, 344–347 (2015).

30. B. Nguyen, Y. Ordabayev, J. E. Sokoloski, E. Weiland, T. M. Lohman, Large domain movements upon UvrD dimerization and helicase activation. *Proc National Acad Sci* **114**, 12178–12183 (2017).

31. M. J. Comstock, *et al.*, Direct observation of structure-function relationship in a nucleic acid–processing enzyme. *Science* **348**, 352–354 (2015).

32. M. A. Makurath, K. D. Whitley, B. Nguyen, T. M. Lohman, Y. R. Chemla, Regulation of Rep helicase unwinding by an auto-inhibitory subdomain. *Nucleic Acids Res* **47**, 2523–2532 (2019).

33. Y. A. Ordabayev, B. Nguyen, A. Niedziela-Majka, T. M. Lohman, Regulation of UvrD Helicase Activity by MutL. *J Mol Biol* **430**, 4260–4274 (2018).

34. Y. A. Ordabayev, B. Nguyen, A. G. Kozlov, H. Jia, T. M. Lohman, UvrD helicase activation by MutL involves rotation of its 2B subdomain. *Proc National Acad Sci* **116**, 16320–16325 (2019).

35. E. J. Gwynn, *et al.*, The conserved C-terminus of the PcrA/UvrD helicase interacts directly with RNA polymerase. *PloS one* **8**, e78141 (2013).

36. K. Sanders, *et al.*, The structure and function of an RNA polymerase interaction domain in the PcrA/UvrD helicase. *Nucleic Acids Res* **45** (2017).

37. V. Epshtein, *et al.*, UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* **505**, 372 (2014).

38. W. H. Organization, Global TB Report 2020 (2020).

39. D. Jensen, A. R. Manzano, J. Rammohan, C. L. Stallings, E. A. Galburt, CarD and RbpA modify the kinetics of initial transcription and slow promoter escape of the Mycobacterium tuberculosis RNA polymerase. *Nucleic Acids Research* **47**, 6685–6698 (2019).

40. H. Boyaci, R. M. Saecker, E. A. Campbell, Transcription initiation in mycobacteria: a biophysical perspective. *Biochem Soc Symp*, 1–13 (2019).

41. C. Bertrand, A. Thibessard, C. Bruand, F. Lecointe, P. Leblond, Bacterial NHEJ: a never ending story. *Mol Microbiol* **111**, 1139–1151 (2019).

42. J. Chen, H. Boyaci, E. A. Campbell, Diverse and unified mechanisms of transcription initiation in bacteria. *Nat Rev Microbiol* **19**, 95–109 (2021).

43. R. Hershberg, *et al.*, High Functional Diversity in Mycobacterium tuberculosis Driven by Genetic Drift and Human Demography. *Plos Biol* **6**, e311 (2008).

44. K. M. Sinha, N. C. Stephanou, M.-C. Unciuleac, M. S. Glickman, S. Shuman, Domain requirements for DNA unwinding by mycobacterial UvrD2, an essential DNA helicase. *Biochemistry-us* **47**, 9355–9364 (2008).

45. K. M. Sinha, N. C. Stephanou, F. Gao, M. S. Glickman, S. Shuman, Mycobacterial UvrD1 is a Ku-dependent DNA helicase that plays a role in multiple DNA repair events, including double-strand break repair. *J Biol Chem* **282**, 15114–15125 (2007).

46. K. M. Sinha, M. S. Glickman, S. Shuman, Mutational analysis of Mycobacterium UvrD1 identifies functional groups required for ATP hydrolysis, DNA unwinding, and chemomechanical coupling. *Biochemistry* **48**, 4019–4030 (2009).

47. J. Houghton, *et al.*, Important role for Mycobacterium tuberculosis UvrD1 in pathogenesis and persistence apart from its function in nucleotide excision repair. *J Bacteriol* **194**, 2916–2923 (2012).

48. E. Curti, S. J. Smerdon, E. O. Davis, Characterization of the Helicase Activity and Substrate Specificity of Mycobacterium tuberculosis UvrD ▿ . *J Bacteriol* **189**, 1542–1555 (2007).

49. H. T. Pacl, V. P. Reddy, V. Saini, K. C. Chinta, A. J. C. Steyn, Host-pathogen redox dynamics modulate Mycobacterium tuberculosis pathogenesis. *Pathog Dis* **76** (2018).

50. M. Mehta, A. Singh, Mycobacterium tuberculosis WhiB3 maintains redox homeostasis and survival in response to reactive oxygen and nitrogen species. *Free Radical Bio Med* **131**, 50–58 (2018).

51. S. Upadhyay, E. Mittal, J. A. Philips, Tuberculosis and the art of macrophage manipulation. *Pathog Dis* **76** (2018).

52. P. Schuck, Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys J* **78**, 1606–1619 (2000).

53. E. F. Pettersen, *et al.*, UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).

54. H. Sowani, M. Kulkarni, S. Zinjarde, V. Javdekar, The Microbiology of Skin, Soft Tissue, Bone and Joint Infections. 105–121 (2017).

55. M. Goodfellow, A. L. Jones, Bergey's Manual of Systematics of Archaea and Bacteria. 1–14 (2020).

56. A. Gupta, A. N. Ananthakrishnan, Economic burden and cost-effectiveness of therapies for Clostridiodes difficile infection: a narrative review. *Ther Adv Gastroenter* **14**, 175628482110186 (2021).

57. P.-J. Chiu, *et al.*, Clostridioides difficile spores stimulate inflammatory cytokine responses and induce cytotoxicity in macrophages. *Anaerobe* **70**, 102381 (2021).

58. W. Cheng, J. Hsieh, K. M. Brendza, T. M. Lohman, E. coli Rep oligomers are required to initiate DNA unwinding in vitro11Edited by D. Draper. *J Mol Biol* **310**, 327–350 (2001).

59. S. A. E. Marras, F. R. Kramer, S. Tyagi, Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res* **30**, e122–e122 (2002).

60. J. A. Ali, T. M. Lohman, Kinetic Measurement of the Step Size of DNA Unwinding by Escherichia coli UvrD Helicase. *Science* **275**, 377–380 (1997).

61. C. G. Wu, C. Bradford, T. M. Lohman, Escherichia coli RecBC helicase has two translocase activities controlled by a single ATPase motor. *Nat Struct Mol Biol* **17**, 1210–1217 (2010).

62. C. J. Fischer, N. K. Maluf, T. M. Lohman, Mechanism of ATP-dependent Translocation of E.coli UvrD Monomers Along Single-stranded DNA. *J Mol Biol* **344**, 1287–1309 (2004).

63. E. J. Tomko, C. J. Fischer, T. M. Lohman, Single-Stranded DNA Translocation of E. coli UvrD Monomer Is Tightly Coupled to ATP Hydrolysis. *J Mol Biol* **418**, 32–46 (2012).

64. A. L. Lucius, N. K. Maluf, C. J. Fischer, T. M. Lohman, General Methods for Analysis of Sequential "n-step" Kinetic Mechanisms: Application to Single Turnover Kinetics of Helicase-Catalyzed DNA Unwinding. *Biophys J* **85**, 2224–2239 (2003).

65. H. S. Subramanya, L. E. Bird, J. A. Brannigan, D. B. Wigley, Crystal structure of a DExx box DNA helicase. *Nature* **384**, 379–383 (1996).

66. S. Korolev, J. Hsieh, G. H. Gauss, T. M. Lohman, G. Waksman, Major Domain Swiveling Revealed by the Crystal Structures of Complexes of E. coli Rep Helicase Bound to Single-Stranded DNA and ADP. *Cell* **90**, 635–647 (1997).

67. S. S. Velankar, P. Soultanas, M. S. Dillingham, H. S. Subramanya, D. B. Wigley, Crystal Structures of Complexes of PcrA DNA Helicase with a DNA Substrate Indicate an Inchworm Mechanism. *Cell* **97**, 75–84 (1999).

68. J. Y. Lee, W. Yang, UvrD Helicase Unwinds DNA One Base Pair at a Time by a Two-Part Power Stroke. *Cell* **127**, 1349–1360 (2006).

69. H. Jia, *et al.*, Rotations of the 2B Sub-domain of E. coli UvrD Helicase/Translocase Coupled to Nucleotide and DNA Binding. *J Mol Biol* **411**, 633–648 (2011).

70. H. Yokota, Y. A. Chujo, Y. Harada, Single-Molecule Imaging of the Oligomer Formation of the Nonhexameric Escherichia coli UvrD Helicase. *Biophys J* **104**, 924–933 (2013).

71. J. Park, *et al.*, PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. *Cell* **142**, 544–555 (2010).

72. B. Sun, *et al.*, Impediment of E. coli UvrD by DNA-destabilizing force reveals a strained-inchworm mechanism of DNA unwinding - SI. *Embo J* **27**, 3279–3287 (2008).

73. P. Soultanas, M. S. Dillingham, P. Wiley, M. R. Webb, D. B. Wigley, Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism. *Embo J* **19**, 3799–3810 (2000).

74. M. R. Singleton, M. S. Dillingham, M. Gaudier, S. C. Kowalczykowski, D. B. Wigley, Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432**, 187–193 (2004).

75. M. Wilkinson, Y. Chaban, D. B. Wigley, Mechanism for nuclease regulation in RecBCD. *Elife* **5**, e18227 (2016).

76. B. Nguyen, M. K. Shinn, E. Weiland, T. M. Lohman, Regulation of E. coli Rep Helicase Activity by PriC. *J Mol Biol* **433**, 167072 (2021).

77. P. Singh, *et al.*, Mycobacterium tuberculosis UvrD1 and UvrA proteins suppress DNA strand exchange promoted by cognate and noncognate RecA proteins. *Biochemistry-us* **49**, 4872–4883 (2010).

78. A. Williams, *et al.*, UvrD2 is essential in Mycobacterium tuberculosis, but its helicase activity is not required. *J Bacteriol* **193**, 4487–4494 (2011).

79. K. H. Darwin, C. F. Nathan, Role for Nucleotide Excision Repair in Virulence of Mycobacterium tuberculosis. *Infect Immun* **73**, 4581–4587 (2005).

80. S. K. Matta, D. Kumar, Hypoxia and classical activation limits Mycobacterium tuberculosis survival by Akt-dependent glycolytic shift in macrophages. *Cell Death Discov* **2**, 16022 (2016).

81. B. Springer, *et al.*, Lack of mismatch correction facilitates genome evolution in mycobacteria. *Mol Microbiol* **53**, 1601–1609 (2004).

82. L. Cabusora, E. Sutton, A. Fulmer, C. V. Forst, Differential network expression during drug and stress response. *Bioinformatics* **21**, 2898–2905 (2005).

83. C. Güthlein, *et al.*, Characterization of the Mycobacterial NER System Reveals Novel Functions of the uvrD1 Helicase ▿ . *J Bacteriol* **191**, 555–562 (2008).

84. J. A. Imlay, S. Linn, Mutagenesis and stress responses induced in Escherichia coli by hydrogen peroxide. *J Bacteriol* **169**, 2967–2976 (1987).

85. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845–858 (2015).

86. F. Sievers, *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).

87. G. Perrière, M. Gouy, WWW-query: An on-line retrieval system for biological sequence banks. *Biochimie* **78**, 364–369 (1996).

88. E. Tomko, "Transient-state kinetic studies of Escherichia Coli UvrD monomer translocation along single-stranded DNA." (2010).

89. L. D. Inners, G. Felsenfeld, Conformation of polyribouridylic acid in solution. *J Mol Biol* **50**, 373–389 (1970).

90. J. P. Gabrielson, K. K. Arthur, B. S. Kendrick, T. W. Randolph, M. R. Stoner, Common excipients impair detection of protein aggregates during sedimentation velocity analytical ultracentrifugation. *J Pharm Sci* **98**, 50–62 (2009).

91. T. M. Laue, B. D. Shah, T. M. Ridgeway, S. L. Pelletier, "Computer-aided interpretation of analytical sedimentation data for proteins" in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, S. Harding, A. Rowe, J. Horton, Eds. (Royal Society of Chemistry, 1992), pp. 90–125.

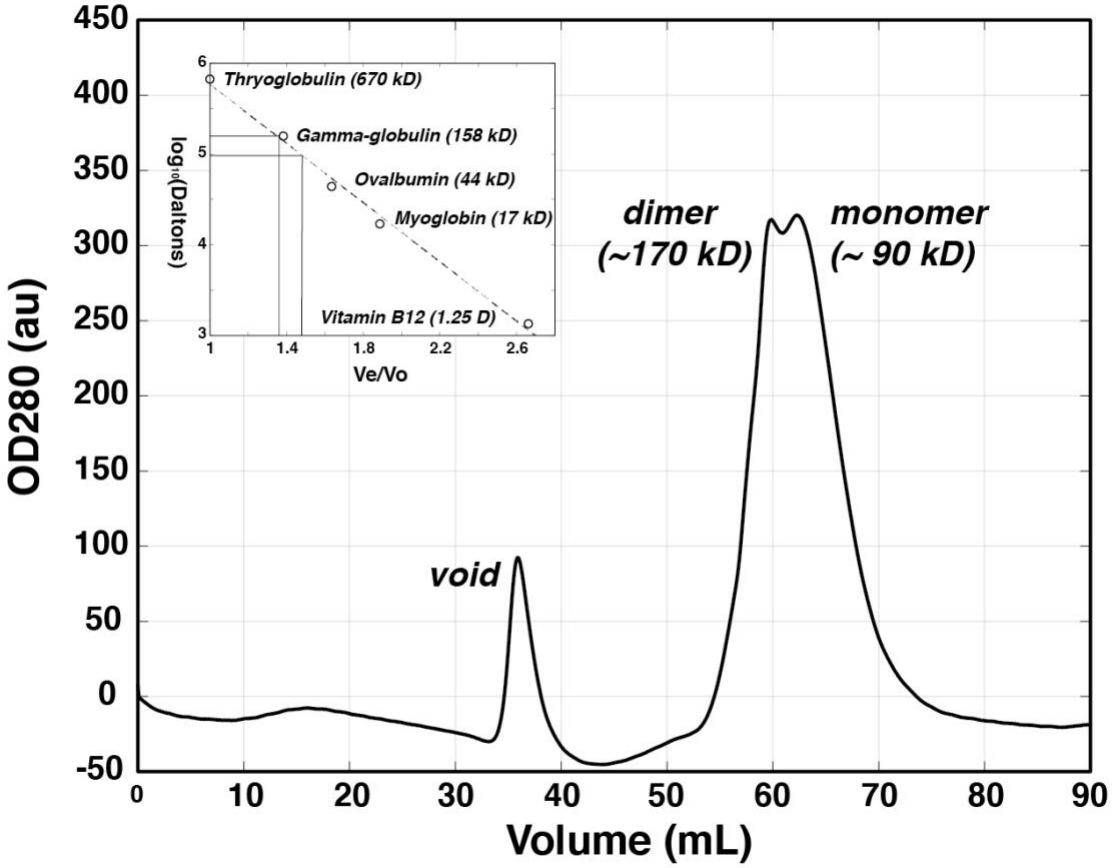
92. J. Dam, P. Schuck, Calculating Sedimentation Coefficient Distributions by Direct Modeling of Sedimentation Velocity Concentration Profiles. *Methods Enzymol* **384**, 185–212 (2004).

93. D. P. Mascotti, T. M. Lohman, Thermodynamics of Charged Oligopeptide-Heparin Interactions. *Biochemistry-us* **34**, 2908–2915 (1995).

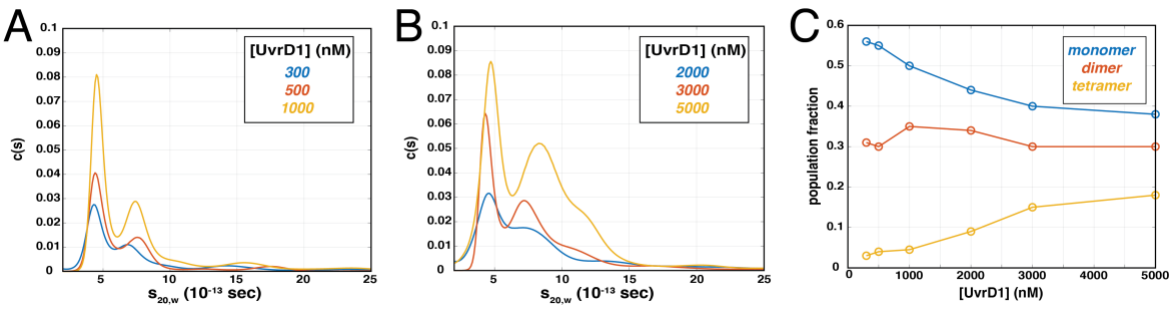
94. J. R. Lakowicz, "Fluorescence Anisotropy" in Principles of Fluorescence Spectroscopy., (Springer US, 1999), pp. 291–319.

Supplemental Information

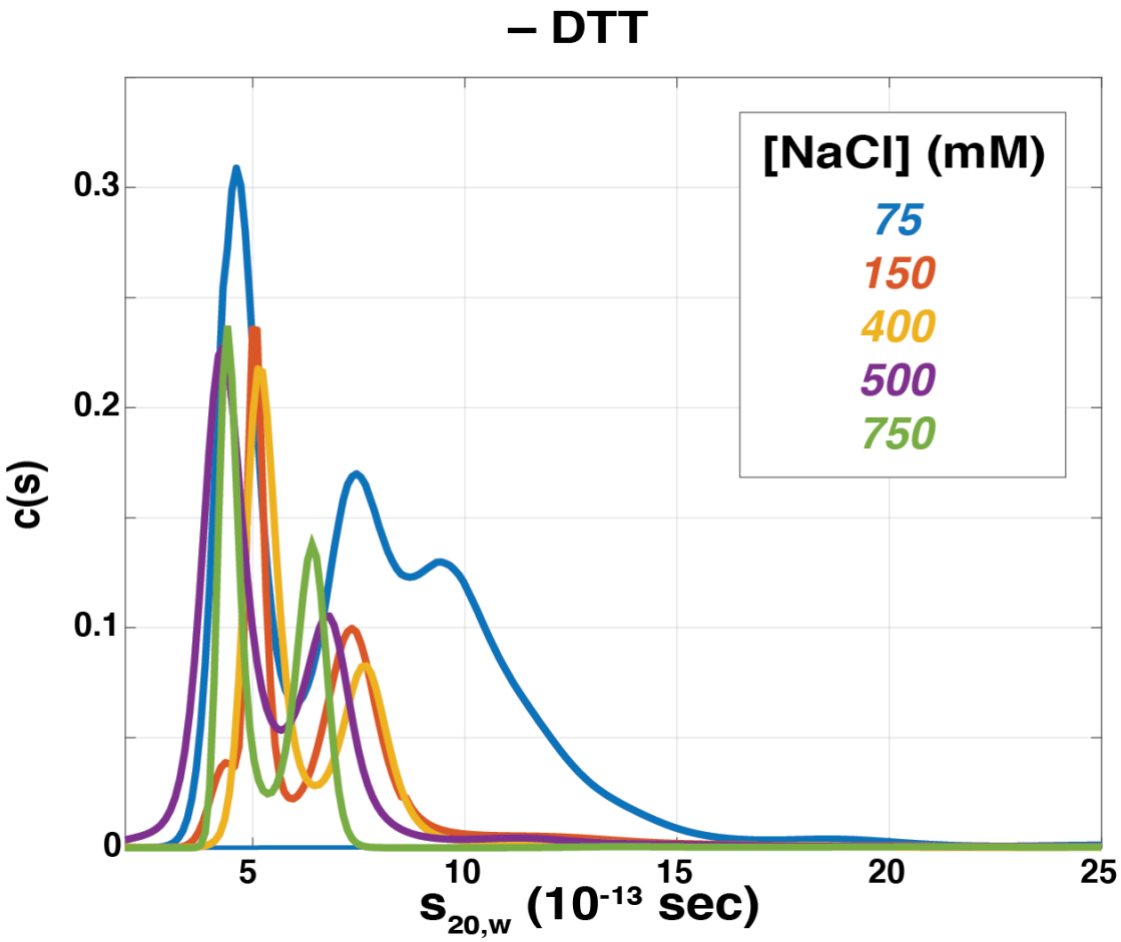
Chapter 3: The *Mycobacterium tuberculosis* DNA-repair Helicase UvrD1 is Activated by Redox-dependent Dimerization via a 2B domain Cysteine Conserved in other Actinobacteria.



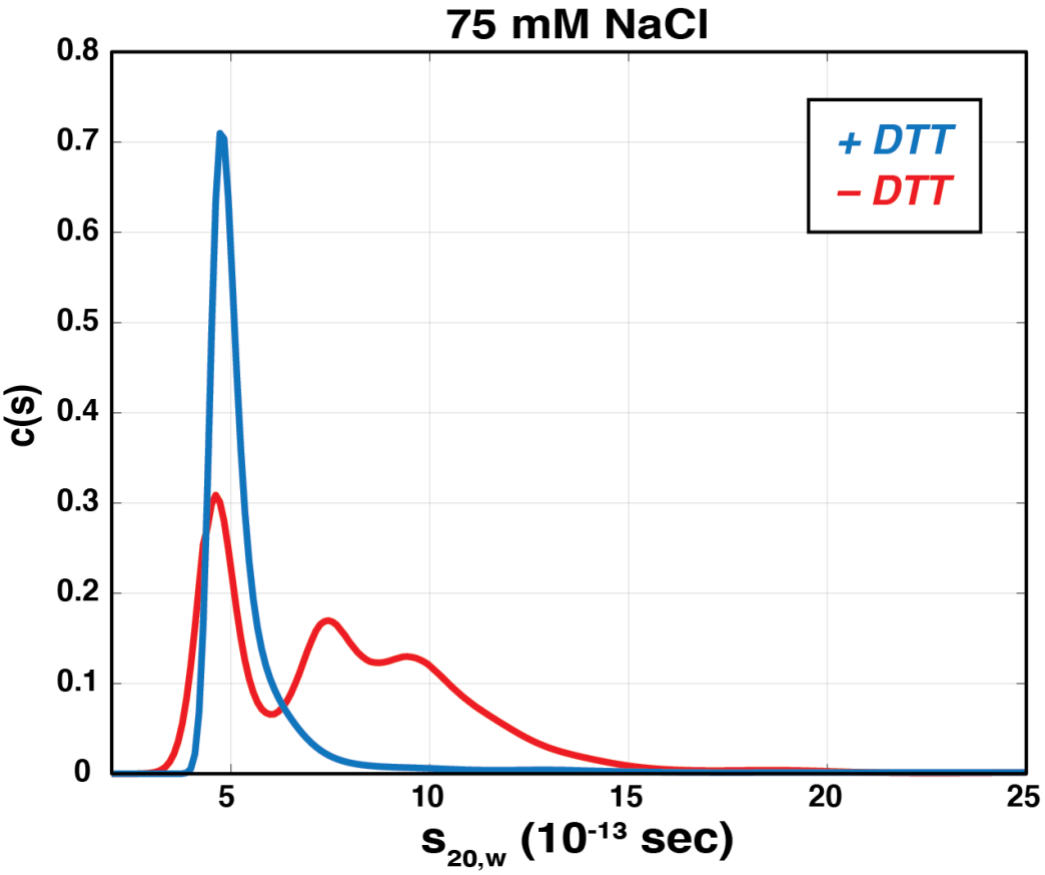
Supplemental Figure 3.1: Size exclusion chromatogram indicating the presence of multiple oligomeric forms of UvrD1. A S300 size exclusion column was run in 150 mM NaCl, 10% (v/v) glycerol, 10mM TRIS pH 8.0 at 25 ºC in the absence of DTT at 4 ºC. Standards run on the same column were used to estimate the molecular weight of the two peaks (see Methods) (insert). The first peak is consistent with the presence of UvrD1 dimers (170,000 Da measured compared to 170,000 Da theoretical) while the second is consistent with the presence of UvrD1 monomers (90,000 Da compared to 85,000Da theoretical). We note here that depending on the resolution of individual sizing runs, this doublet peak is not observed upon purification; instead, a long tail at larger molecular weights is observed.



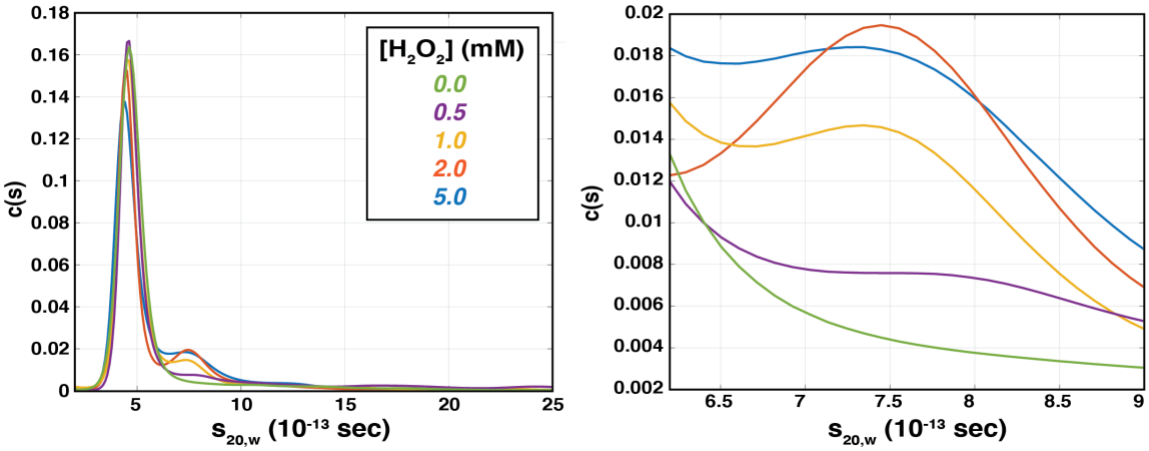
Supplemental Figure 3.2: AUC sedimentation velocity c(s) distributions of UvrD1 as a function of protein concentration. All experiments were performed in Buffer A with 75 mM NaCl and 1mM DTT at respective concentrations and treated with 2mM H2O2. (A) 300 nM (blue), 500 nM (red), and 1000 nM (yellow) runs were monitored at 230 nm. (B) 2000 nM (blue), 3000 nM (red), and 5000 nM (yellow) runs were monitored at 280 nm. (C) The population fraction of monomer (blue), dimer (red), and tetramer (yellow) were calculated by integrating the data from A and B and are plotted as a function of protein concentration.



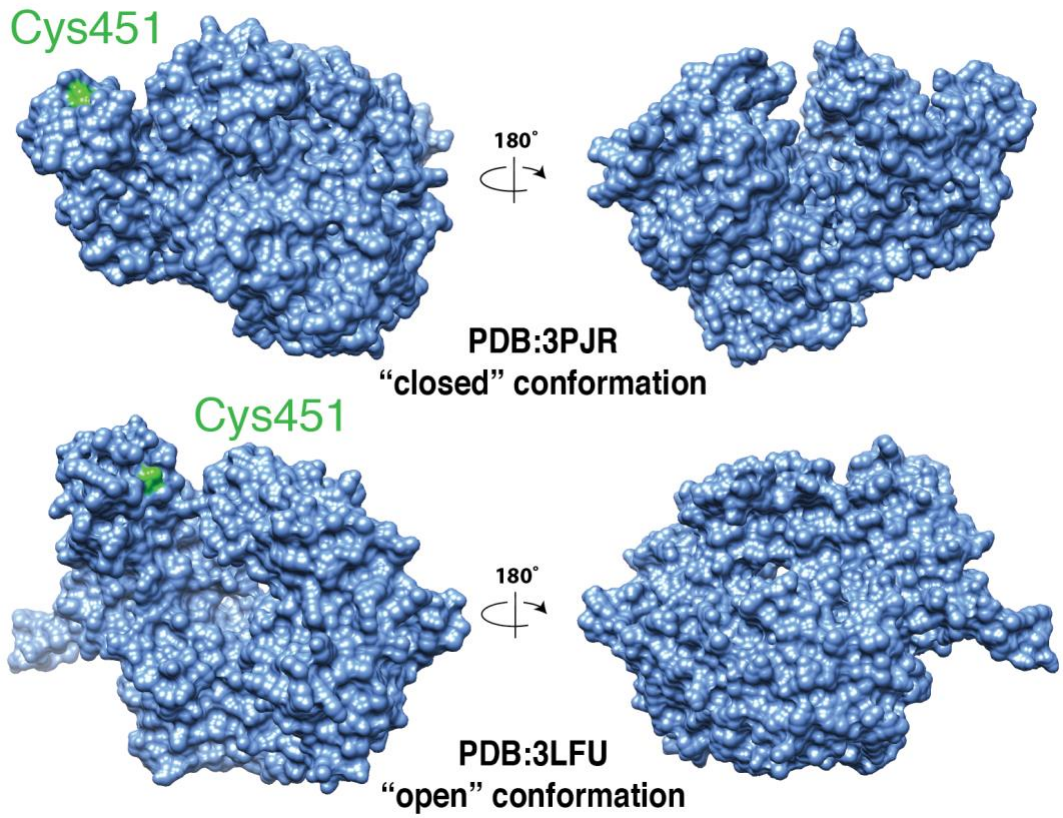
Supplemental Figure 3.3: AUC sedimentation velocity c(s) distributions at different NaCl concentrations. Sedimentation velocity traces used to measure fractions of monomer, dimer, and higher order oligomers presented in the main text (Fig. 1B). Experiments were performed at a constant 2.5 μM protein concentration in Buffer A (TRIS pH 8.0 at 25 C, 20% glycerol) without DTT at the NaCl concentration indicated in the legend. All traces are averages of two separate experiments and were monitored at 280 nm except for 75 mM NaCl which was monitored at 230 nm. Therefore, the amplitudes of 75 mM condition should not be compared to the other traces.



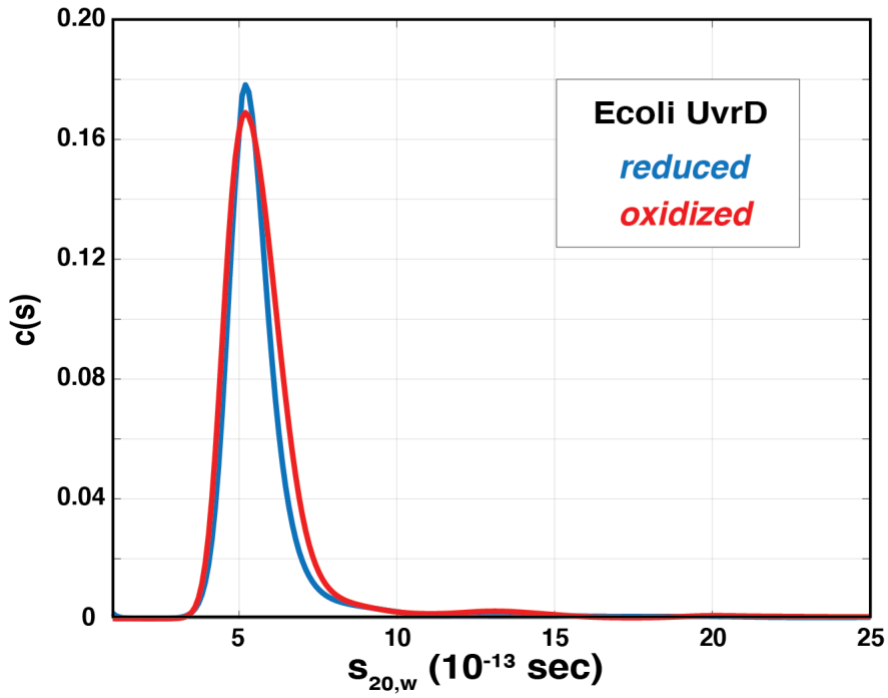
Supplemental Figure 3.4: DTT dependence of dimerization at 75 mM NaCl. AUC sedimentation velocity c(s) distributions display a similar effect of DTT in Buffer A with 75 mM NaCl compared to the results in the main text at 400 mM NaCl (Fig. 1C). Each trace is an average of two separate experiments monitored at 230 nm at 2.5 μM protein concentration.



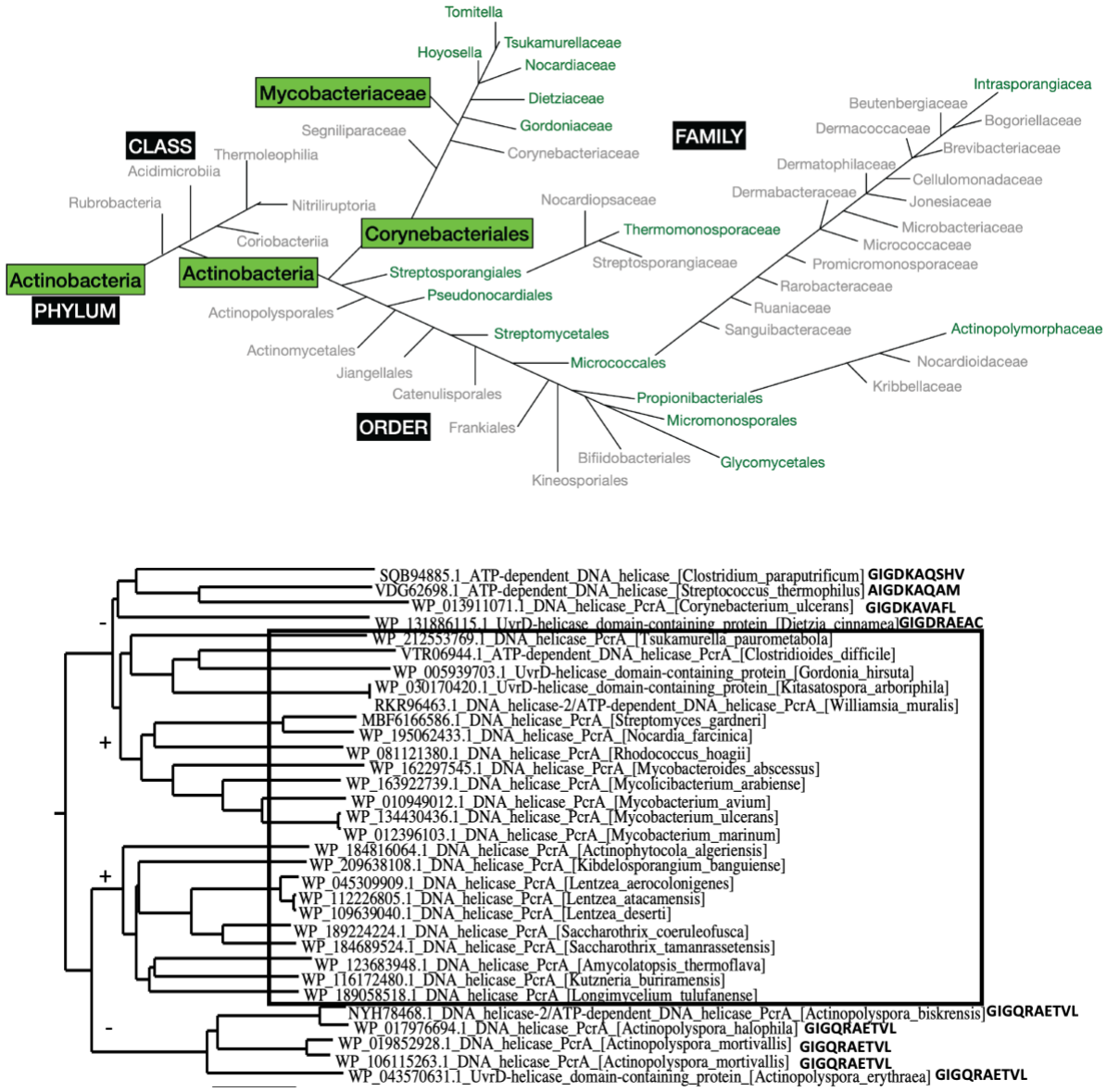
Supplemental Figure 3.5: AUC sedimentation velocity c(s) distributions at different hydrogen peroxide concentrations. 2.5 µM UvrD1 was dialyzed in Buffer A with 75 mM NaCl and 1 mM DTT and then subjected to H_2O_2 addition (see Methods). The dimer fraction was found to increase as a function of H_2O_2 concentration (quantified in Fig. 1D). (Right) A zoom into the region representative of the dimer is shown for clarity. Each trace is an average of two separate experiments, monitored at 280 nm.



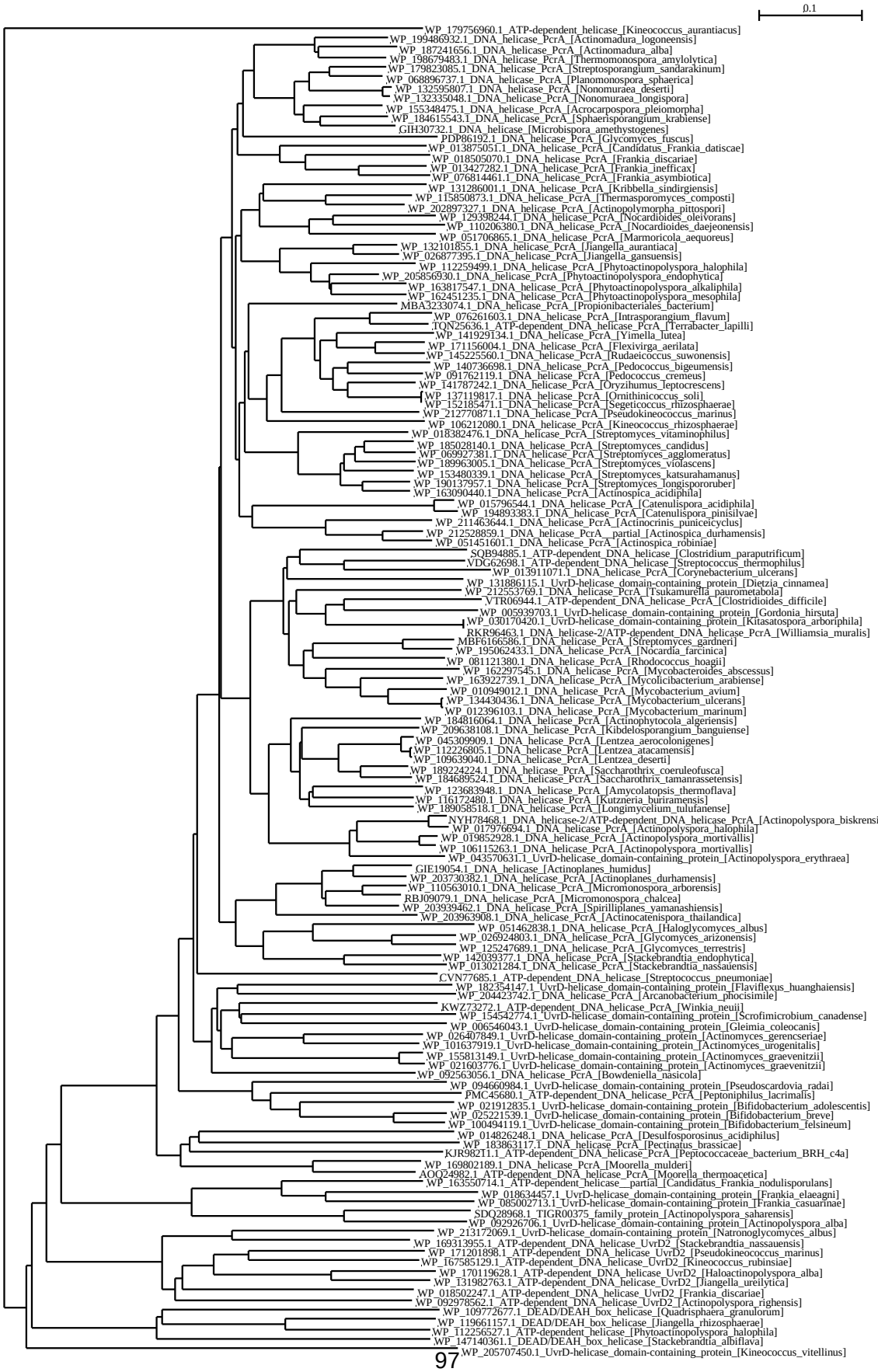
Supplemental Figure 3.6: (A) A surface view of the threaded *Mtb* UvrD1 structure based on the *G. stearothermophilus* UvrD "closed" structure (PDB:3PJR). (B) A surface view of the threaded *Mtb* UvrD1 structure based on the *E. coli* "open" UvrD structure (PDB:3LFU). In both cases, of the three total cysteines in the *Mtb* UvrD1 sequence, only C451 in the 2B domain is surfaced exposed (green).

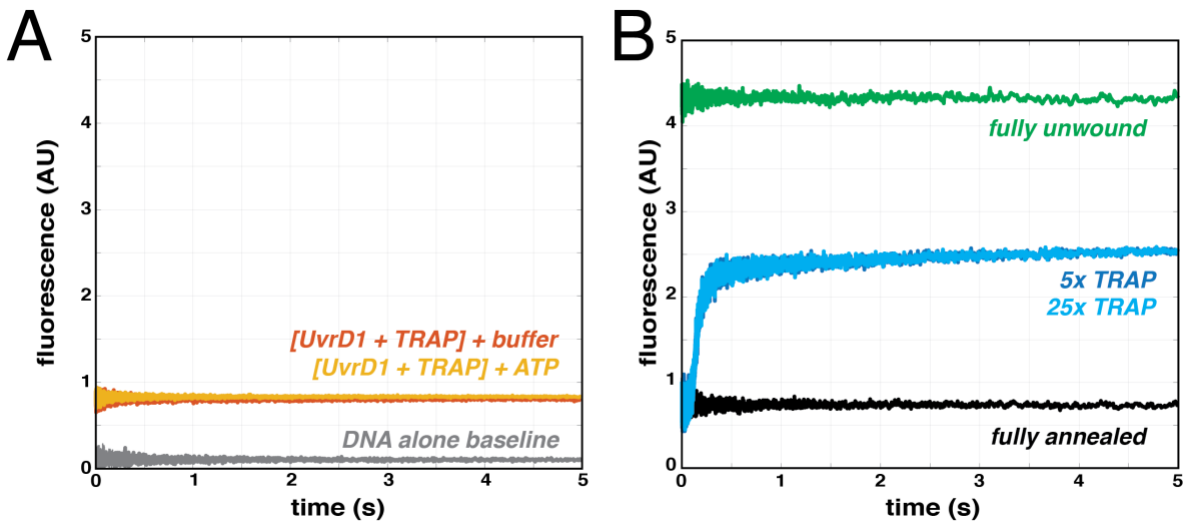


Supplemental Figure 3.7: *E. coli* UvrD does not dimerize in a redox-dependent manner. Sedimentation velocity was run with *E. coli* UvrD in both reducing and oxidizing conditions. Briefly, 4 µM UvrD stored in 200 mM NaCl, 50% (v/v) glycerol,10 mM TRIS pH 8.3 25 ºC and 25 mM beta-mercaptoethanol was first dialyzed in Buffer A with 75 mM NaCl and 2 µM protein was incubated with either 1 mM DTT (blue) or 1 mM DTT with 2 mM H_2O_2 (red). The resulting c(s) distributions indicated a single species observed in both redox conditions, suggesting that *E. coli* UvrD does not dimerize in response to oxidative conditions.

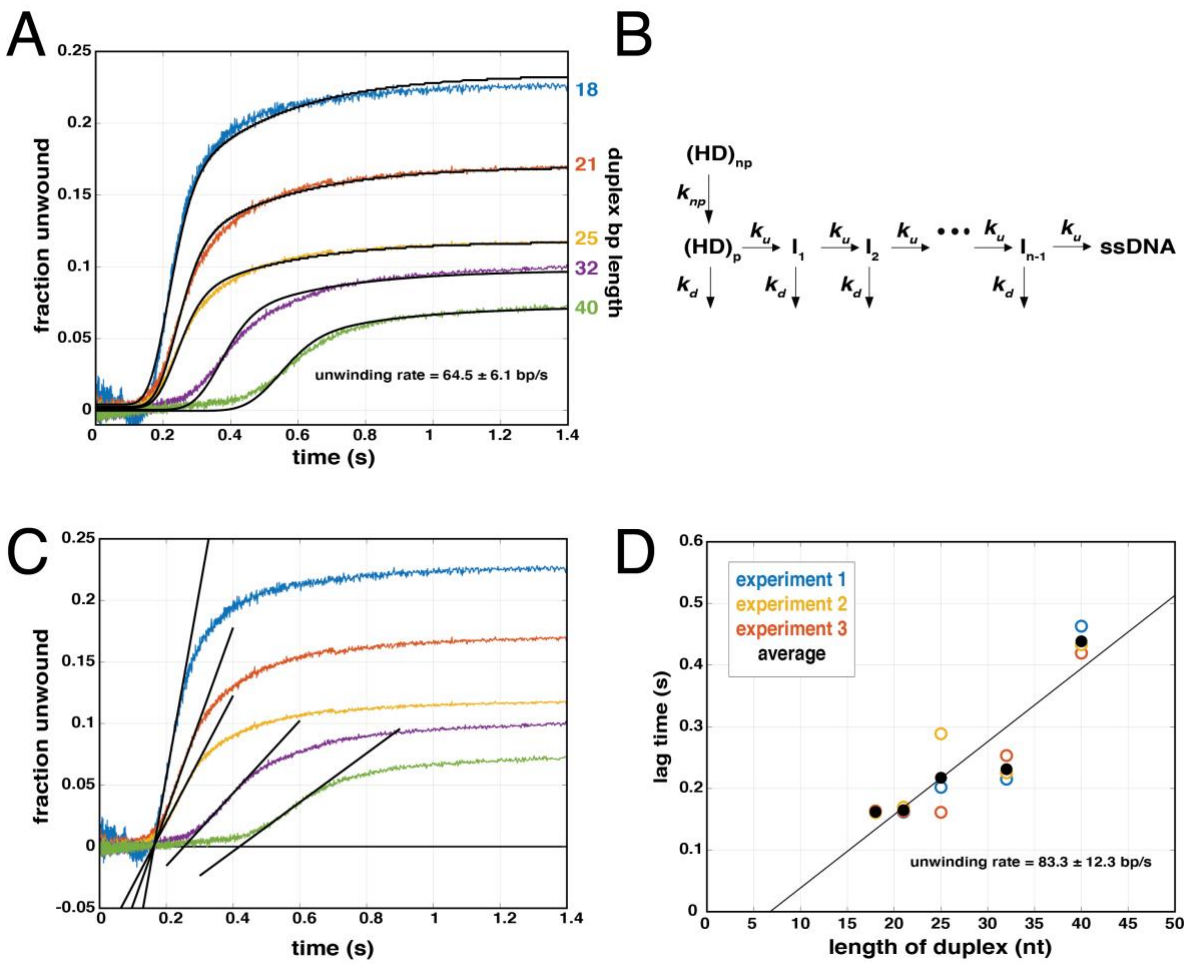


Supplemental Figure 3.8: A more detailed look at the distribution of the 2B cysteine across the orders and families of the Actinobacteria class. In the top panel, groups containing at least one organism with the 2B cysteine in the same position as UvrD1 are indicated in green. The boxed green labels refer to the family in which *Mtb* is found. Those that do not contain the 2B cysteine are indicated in gray. In the bottom panel a section of a larger phylogenic tree constructed using the UvrD1 sequence in different groups of actinobacteria and firmicutes. The boxed species all contain the 2B cysteine and the surrounding conserved residues as in Fig. 3B. The full tree can be seen on the next page.

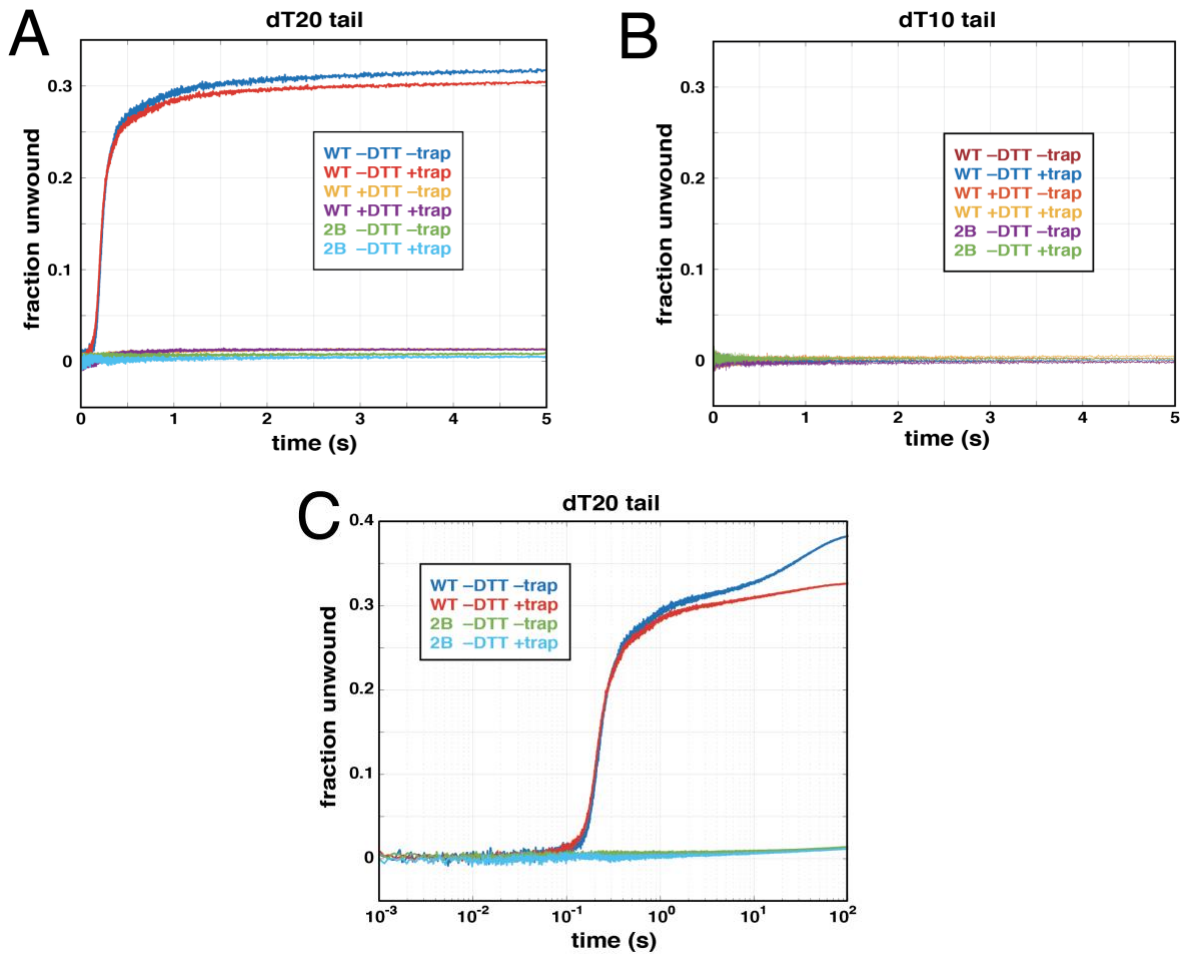

0.1
WP_179756960.1_ATP-dependent_helicase_[Kineococcus_aurantiacus]
WP_199486932.1_DNA_helicase_PcrA_[Actinomadura_logoneensis]
WP_187241656.1_DNA_helicase_PcrA_[Actinomadura_alba]
WP_198679483.1_DNA_helicase_PcrA_[Thermomonospora_amylolytica]
WP_179823085.1_DNA_helicase_PcrA_[Streptosporangium_sandarakinum]
WP_068896737.1_DNA_helicase_PcrA_[Planomonospora_sphaerica]
WP_132595807.1_DNA_helicase_PcrA_[Nonomuraea_deserti]
WP_132335048.1_DNA_helicase_PcrA_[Nonomuraea_longispora]
WP_155348475.1_DNA_helicase_PcrA_[Acrocarpospora_pleiomorpha]
WP_184615543.1_DNA_helicase_PcrA_[Sphaerisporangium_krabiense]
GIH30732.1_DNA_helicase_[Microbispora_amethystogenes]
PDP86192.1_DNA_helicase_PcrA_[Glycomyces_fuscus]
WP_013875051.1_DNA_helicase_PcrA_[Candidatus_Frankia_datiscae]
WP_018505070.1_DNA_helicase_PcrA_[Frankia_discariae]
WP_013427282.1_DNA_helicase_PcrA_[Frankia_inefficax]
WP_076814461.1_DNA_helicase_PcrA_[Frankia_asymbiotica]
WP_131286001.1_DNA_helicase_PcrA_[Kribbella_sindirgiensis]
WP_115850873.1_DNA_helicase_PcrA_[Thermasporomyces_composti]
WP_202897327.1_DNA_helicase_PcrA_[Actinopolymorpha_pittospori]
WP_129398244.1_DNA_helicase_PcrA_[Nocardioides_oleivorans]
WP_110206380.1_DNA_helicase_PcrA_[Nocardioides_daejeonensis]
WP_051706865.1_DNA_helicase_PcrA_[Marmoricola_aequoreus]
WP_132101855.1_DNA_helicase_PcrA_[Jiangella_aurantiaca]
WP_026877395.1_DNA_helicase_PcrA_[Jiangella_gansuensis]
WP_112259499.1_DNA_helicase_PcrA_[Phytoactinopolyspora_halophila]
WP_205856930.1_DNA_helicase_PcrA_[Phytoactinopolyspora_endophytica]
WP_163817547.1_DNA_helicase_PcrA_[Phytoactinopolyspora_alkaliphila]
WP_162451235.1_DNA_helicase_PcrA_[Phytoactinopolyspora_mesophila]
MBA3233074.1_DNA_helicase_PcrA_[Propionibacteriales_bacterium]
WP_076261603.1_DNA_helicase_PcrA_[Intrasporangium_flavum]
TQN25636.1_ATP-dependent_DNA_helicase_PcrA_[Terrabacter_lapilli]
WP_141929134.1_DNA_helicase_PcrA_[Yimella_lutea]
WP_171156004.1_DNA_helicase_PcrA_[Flexivirga_aerilata]
WP_145225560.1_DNA_helicase_PcrA_[Rudaeicoccus_suwonensis]
WP_140736698.1_DNA_helicase_PcrA_[Pedococcus_bigeumensis]
WP_091762119.1_DNA_helicase_PcrA_[Pedococcus_cremeus]
WP_141787242.1_DNA_helicase_PcrA_[Oryzihumus_leptocrescens]
WP_137119817.1_DNA_helicase_PcrA_[Ornithinicoccus_soli]
WP_152185471.1_DNA_helicase_PcrA_[Segeticoccus_rhizosphaerae]
WP_212770871.1_DNA_helicase_PcrA_[Pseudokineococcus_marinus]
WP_106212080.1_DNA_helicase_PcrA_[Kineococcus_rhizosphaerae]
WP_018382476.1_DNA_helicase_PcrA_[Streptomyces_vitaminophilus]
WP_185028140.1_DNA_helicase_PcrA_[Streptomyces_candidus]
WP_069927381.1_DNA_helicase_PcrA_[Streptomyces_agglomeratus]
WP_189963005.1_DNA_helicase_PcrA_[Streptomyces_violascens]
WP_153480339.1_DNA_helicase_PcrA_[Streptomyces_katsurahamanus]
WP_190137957.1_DNA_helicase_PcrA_[Streptomyces_longispororuber]
WP_163090440.1_DNA_helicase_PcrA_[Actinospica_acidiphila]
WP_015796544.1_DNA_helicase_PcrA_[Catenulispora_acidiphila]
WP_194893383.1_DNA_helicase_PcrA_[Catenulispora_pinisilvae]
WP_211463644.1_DNA_helicase_PcrA_[Actinocrinis_puniceicyclus]
WP_212528859.1_DNA_helicase_PcrA__partial_[Actinospica_durhamensis]
WP_051451601.1_DNA_helicase_PcrA_[Actinospica_robiniae]
SQB94885.1_ATP-dependent_DNA_helicase_[Clostridium_paraputrificum]
VDG62698.1_ATP-dependent_DNA_helicase_[Streptococcus_thermophilus]
WP_013911071.1_DNA_helicase_PcrA_[Corynebacterium_ulcerans]
WP_131886115.1_UvrD-helicase_domain-containing_protein_[Dietzia_cinnamea]
WP_212553769.1_DNA_helicase_PcrA_[Tsukamurella_pauurometabola]
VTR06944.1_ATP-dependent_DNA_helicase_PcrA_[Clostridioides_difficile]
WP_005939703.1_UvrD-helicase_domain-containing_protein_[Gordonia_hirsuta]
WP_030170420.1_UvrD-helicase_domain-containing_protein_[Kitasatospora_arboriphila]
RKR96463.1_DNA_helicase-2/ATP-dependent_DNA_helicase_PcrA_[Williamsia_muralis]
MBF6166586.1_DNA_helicase_PcrA_[Streptomyces_gardneri]
WP_195062433.1_DNA_helicase_PcrA_[Nocardia_farcinica]
WP_081121380.1_DNA_helicase_PcrA_[Rhodococcus_hoagii]
WP_162297545.1_DNA_helicase_PcrA_[Mycobacteroides_abscessus]
WP_163922739.1_DNA_helicase_PcrA_[Mycolicibacterium_arabiense]
WP_010949012.1_DNA_helicase_PcrA_[Mycobacterium_avium]
WP_134430436.1_DNA_helicase_PcrA_[Mycobacterium_ulcerans]
WP_012396103.1_DNA_helicase_PcrA_[Mycobacterium_marinum]
WP_184816064.1_DNA_helicase_PcrA_[Actinophytocola_algeriensis]
WP_209638108.1_DNA_helicase_PcrA_[Kibdelosporangium_banguiense]
WP_045309909.1_DNA_helicase_PcrA_[Lentzea_aerocolonigenes]
WP_112226805.1_DNA_helicase_PcrA_[Lentzea_atacamensis]
WP_109639040.1_DNA_helicase_PcrA_[Lentzea_deserti]
WP_189224224.1_DNA_helicase_PcrA_[Saccharothrix_coeruleofusca]
WP_184689524.1_DNA_helicase_PcrA_[Saccharothrix_tamanrassetensis]
WP_123683948.1_DNA_helicase_PcrA_[Amycolatopsis_thermoflava]
WP_161724480.1_DNA_helicase_PcrA_[Kutzneria_buriramensis]
WP_189058518.1_DNA_helicase_PcrA_[Longimycelium_tulufanense]
NYH78468.1_DNA_helicase-2/ATP-dependent_DNA_helicase_PcrA_[Actinopolyspora_biskrensis]
WP_017976694.1_DNA_helicase_PcrA_[Actinopolyspora_halophila]
WP_019852928.1_DNA_helicase_PcrA_[Actinopolyspora_mortivallis]
WP_106115263.1_DNA_helicase_PcrA_[Actinopolyspora_mortivallis]
WP_043570631.1_UvrD-helicase_domain-containing_protein_[Actinopolyspora_erythraea]
GIE19054.1_DNA_helicase_[Actinoplanes_humidus]
WP_203730382.1_DNA_helicase_PcrA_[Actinoplanes_durhamensis]
WP_110563010.1_DNA_helicase_PcrA_[Micromonospora_arborensis]
RBJ09079.1_DNA_helicase_PcrA_[Micromonospora_chalcea]
WP_203939462.1_DNA_helicase_PcrA_[Spirilliplanes_yamanashiensis]
WP_203963908.1_DNA_helicase_PcrA_[Actinocatenispora_thailandica]
WP_051462838.1_DNA_helicase_PcrA_[Haloglycomyces_albus]
WP_026924803.1_DNA_helicase_PcrA_[Glycomyces_arizonensis]
WP_125247689.1_DNA_helicase_PcrA_[Glycomyces_terrestris]
WP_142039377.1_DNA_helicase_PcrA_[Stackebrandtia_endophytica]
WP_013021284.1_DNA_helicase_PcrA_[Stackebrandtia_nassauensis]
CVN77685.1_ATP-dependent_DNA_helicase_[Streptococcus_pneumoniae]
WP_182354147.1_UvrD-helicase_domain-containing_protein_[Flaviflexus_huanghaiensis]
WP_204423742.1_DNA_helicase_PcrA_[Arcanobacterium_phocisimile]
KWZ73272.1_ATP-dependent_DNA_helicase_PcrA_[Winkia_neuii]
WP_154542774.1_UvrD-helicase_domain-containing_protein_[Scrofimicrobium_canadense]
WP_006546043.1_UvrD-helicase_domain-containing_protein_[Gleimia_coleocanis]
WP_026407849.1_UvrD-helicase_domain-containing_protein_[Actinomyces_gerencseriae]
WP_101637919.1_UvrD-helicase_domain-containing_protein_[Actinomyces_urogenitalis]
WP_155813149.1_UvrD-helicase_domain-containing_protein_[Actinomyces_graevenitzii]
WP_021603776.1_UvrD-helicase_domain-containing_protein_[Actinomyces_graevenitzii]
WP_092563056.1_DNA_helicase_PcrA_[Bowdeniella_nasicola]
WP_094660984.1_UvrD-helicase_domain-containing_protein_[Pseudoscardovia_radai]
PMC45680.1_ATP-dependent_DNA_helicase_PcrA_[Peptoniphilus_lacrimalis]
WP_021912835.1_UvrD-helicase_domain-containing_protein_[Bifidobacterium_adolescentis]
WP_025221539.1_UvrD-helicase_domain-containing_protein_[Bifidobacterium_breve]
WP_100494119.1_UvrD-helicase_domain-containing_protein_[Bifidobacterium_felsineum]
WP_014826248.1_DNA_helicase_PcrA_[Desulfosporosinus_acidiphilus]
WP_183863117.1_DNA_helicase_PcrA_[Pectinatus_brassicae]
KJR98211.1_ATP-dependent_DNA_helicase_PcrA_[Peptococcaceae_bacterium_BRH_c4a]
WP_169802189.1_DNA_helicase_PcrA_[Moorella_mulderi]
AOQ24982.1_ATP-dependent_DNA_helicase_PcrA_[Moorella_thermoacetica]
WP_163550714.1_ATP-dependent_helicase__partial_[Candidatus_Frankia_nodulisporulans]
WP_018634457.1_UvrD-helicase_domain-containing_protein_[Frankia_elaeagni]
WP_085002713.1_UvrD-helicase_domain-containing_protein_[Frankia_casuarinae]
SDQ28968.1_TIGR00375_family_protein_[Actinopolyspora_saharensis]
WP_092926706.1_UvrD-helicase_domain-containing_protein_[Actinopolyspora_alba]
WP_213172069.1_UvrD-helicase_domain-containing_protein_[Natronoglycomyces_albus]
WP_169313955.1_ATP-dependent_DNA_helicase_UvrD2_[Stackebrandtia_nassauensis]
WP_171201898.1_ATP-dependent_DNA_helicase_UvrD2_[Pseudokineococcus_marinus]
WP_167585129.1_ATP-dependent_DNA_helicase_UvrD2_[Kineococcus_rubinsiae]
WP_170119628.1_ATP-dependent_DNA_helicase_UvrD2_[Haloactinopolyspora_alba]
WP_131982763.1_ATP-dependent_DNA_helicase_UvrD2_[Jiangella_ureilytica]
WP_018502247.1_ATP-dependent_DNA_helicase_UvrD2_[Frankia_discariae]
WP_092978562.1_ATP-dependent_DNA_helicase_UvrD2_[Actinopolyspora_righensis]
WP_109772677.1_DEAD/DEAH_box_helicase_[Quadrisphaera_granulorum]
WP_119661157.1_DEAD/DEAH_box_helicase_[Jiangella_rhizosphaerae]
WP_T12256527.1_ATP-dependent_helicase_[Phytoactinopolyspora_halophila]
WP_147140361.1_DEAD/DEAH_box_helicase_[Stackebrandtia_albiflava]
WP_205707450.1_UvrD-helicase_domain-containing_protein_[Kineococcus_vitellinus]



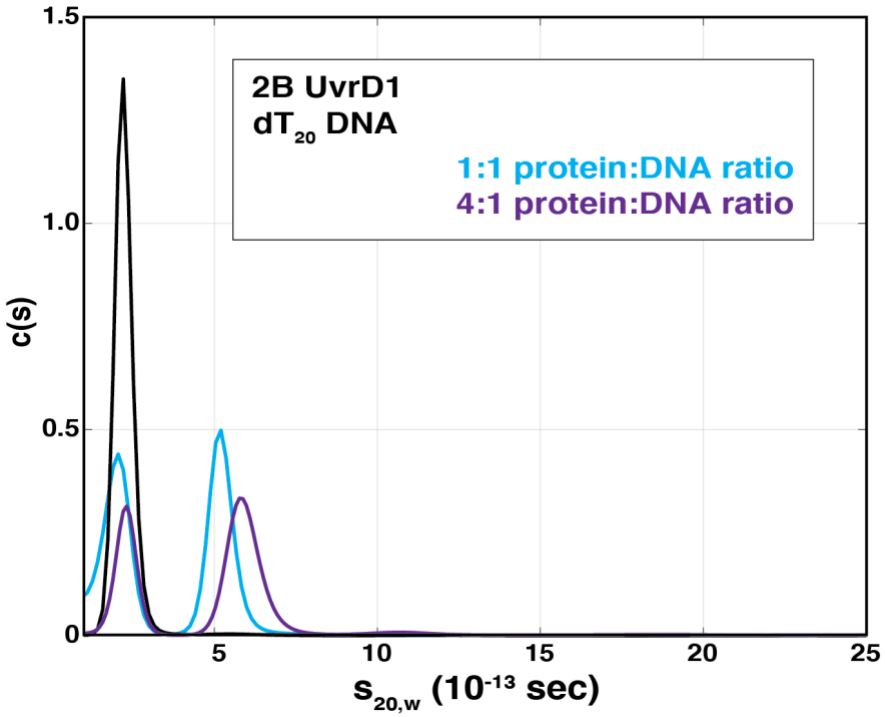
Supplemental Figure 3.9: DNA unwinding controls and calculation of fraction unwound. (A) When TRAP DNA is equilibrated with UvrD1 and the DNA unwinding template and shot against ATP (yellow), the signal is constant and identical to that in the absence of ATP (orange). Note that the addition of UvrD1 does lead to a slight increase in the fluorescence signal compared to DNA alone (gray). (B) The fluorescence signals for traces in the presence of the same UvrD1 concentration including the fully unwound template (green), the fully annealed template (black), and two experimental traces with different molar ratios of TRAP (5X (blue) and 25X (cyan) relative to 200 nM protein). The fraction unwound was calculated as (signal – negative control) / (positive control – negative control) as described in the Methods.



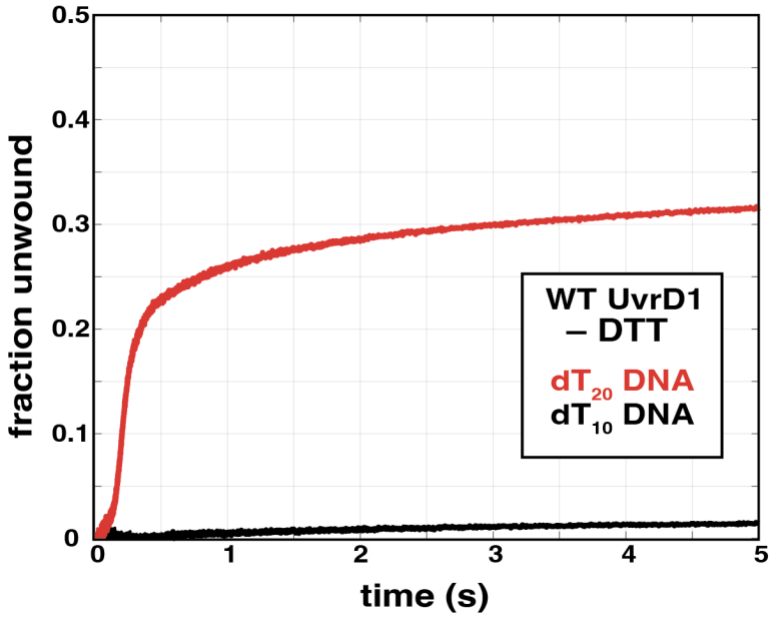
Supplemental Figure 3.10: DNA unwinding as a function of duplex length. DNA templates of different duplex base-pair lengths (18, 21, 25, 32, 40) were used in the unwinding assay under identical conditions to those in the main text (Fig. 4A) for *Mtb* UvrD1 in the absence of DTT. (A) As expected for a processive unwinding mechanism, the kinetics and the final fraction unwound in a single turnover reaction decreased with increasing length of DNA duplex. Examples fits of to an n-step kinetic model including a non-productive fraction are shown (black) and lead to an estimate of the unwinding rate of 64.5 ± 6.1 bp/s. The fit parameters for three separate experiments can be seen in Supplemental Table 2. (B) The kinetic scheme used to fit the traces in (A) is shown. A productive complex $(HD)_p$ can take an unwinding step (k_u) or dissociate (k_d) at each position until arriving at a completely unwound template (ssDNA). A non-productive complex $(HD)_{np}$ must first isomerize to a productive complex (k_{np}) before beginning to unwind. (C) The same data as in (A) showing the estimation of lag times. Lag times were taken as the intercept of the linear portions of the traces with the X-axis as shown. (D) The lag times were measured for each length in each of the three experiments (blue, yellow, and orange circles). The average of these lag times (black circles) was used to fit a line who's inverse gives an estimate of the unwinding rate of 83.3 ± 12.3 bp/s.



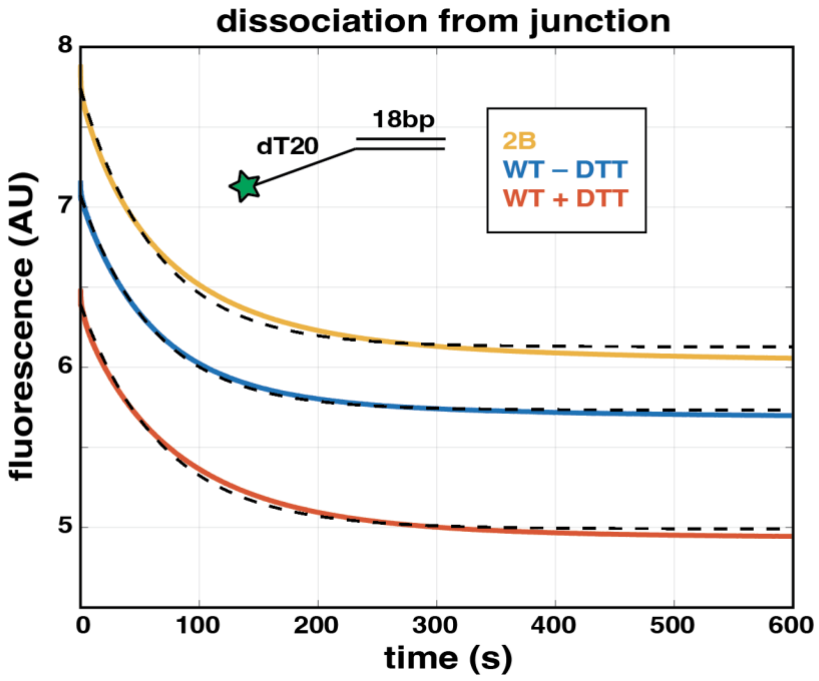
Supplemental Figure 3.11: Stopped-flow DNA unwinding assays under multiple turnover conditions. (A) Using the dT20 template as in the main text, similar amounts of DNA unwinding were observed in the presence and absence of trap. (B) Using the dT10 template, no DNA unwidning was observed regardless of the presence or absence of trap. (C) At much longer time scales (e.g., 100x slower), the lack of trap results in second phase of unwinding, but only under conditions where dimeric UvrD1 is present (blue and red). The monomeric 2B mutant is still unable to unwind, even at these longer time-scales.



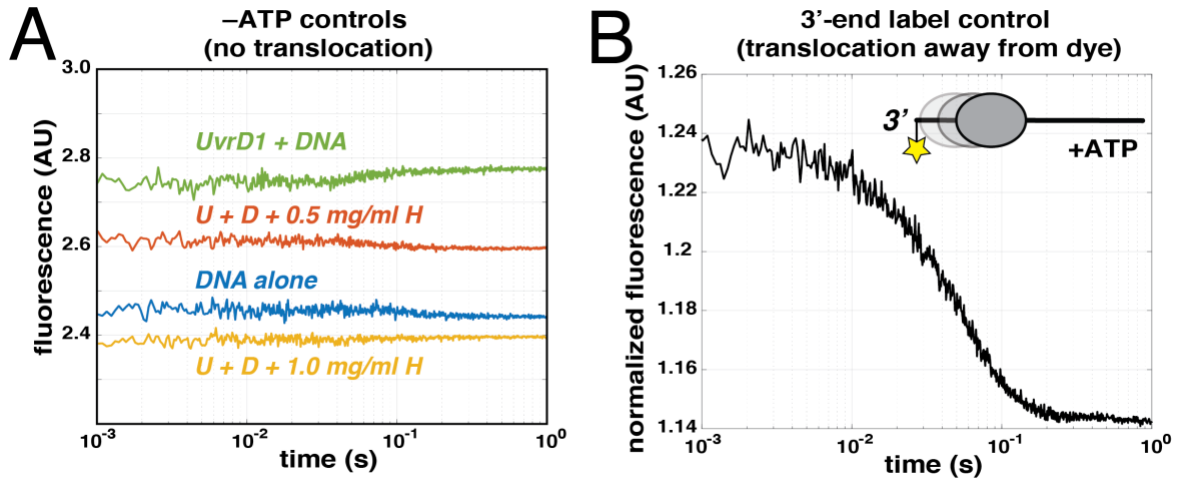
Supplemental Figure 3.12: Excess monomer does not result in dimers bound to DNA. Sedimentation velocity c(s) distributions obtained by monitoring Cy5 fluorescence in Buffer A at 75 mM NaCl in the absence of DTT for dT_{20} DNA alone (black) and the 2B mutant either at a 4X molar excess of protein (purple) or where protein and DNA are at a 1:1 molar stoichiometry (cyan). No sedimenting species near an $s_{20,w}$ value of 8, as measured for the WT dimer-DNA bound complex (Fig. 5B, Supp. Table 3), was detected for either protein:DNA ratio.



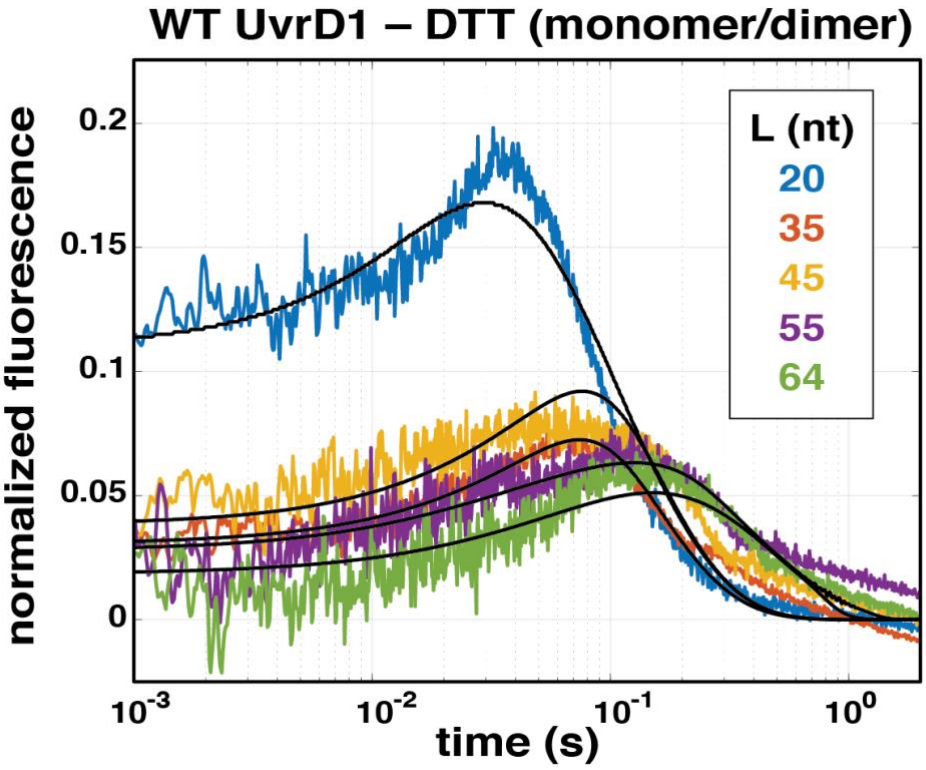
Supplemental Figure 3.13: Comparison of fraction DNA unwound with different lengths of ssDNA tails as a function of time. Experiments here are performed under identical conditions to those in Fig. 4B for WT UvrD1 in the absence of DTT. In contrast to the case with a dT_{20} tail (red) that permits both the monomer and dimer species to bind the unwinding template (Fig. 5B), minimal DNA unwinding is observed with a dT_{10} tail (black), where only the monomeric fraction can bind (Fig. 5B). These results indicate that the UvrD1 dimer is the unwinding-competent species.



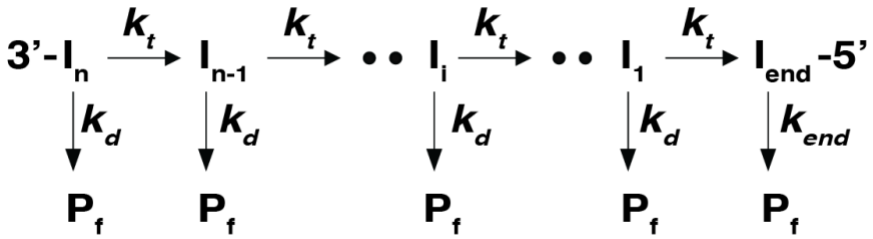
Supplemental Figure 3.14: Dissociation kinetics from DNA junction. Binding of UvrD1 results in a fluorescence enhancement of the template shown above. Mixing of pre-bound protein-DNA complexes with buffer in the presence of TRAP allows for the measurement of the dissociation kinetics of UvrD1 in the absence of ATP. Data using the 2B mutant (yellow), WT –DTT (blue), and WT +DTT (orange) are shown. Fits to each curve with single exponentials are shown in dotted lines and result in estimates of the dissociation rates of 0.016 s^{-1} (2B), 0.016 $^{-1}$ (WT – DTT), and 0.014 s^{-1} (WT + DTT), respectively.



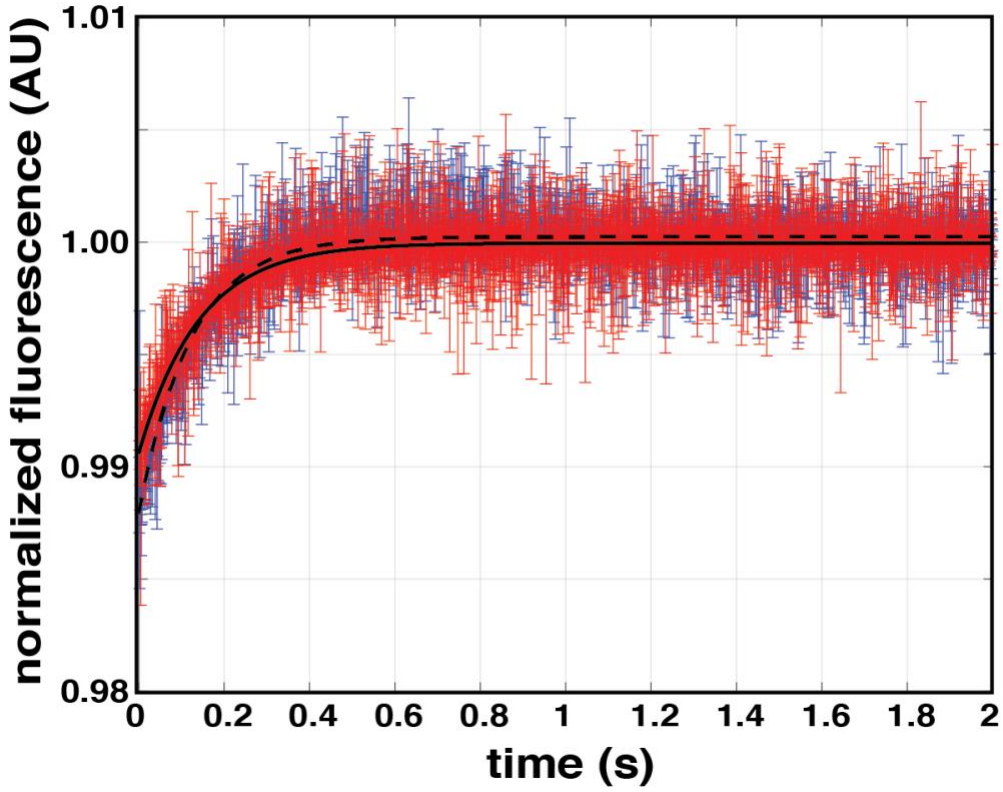
Supplemental Figure 3.15: Heparin controls for establishing single round conditions in translocation assays. Experiments were conducted in Buffer A with 75 mM NaCl at 25 ^{o}C in the absence of DTT. (A) Preincubation of UvrD1 and DNA (green) in the absence of ATP and Mg^{2+} leads to a small increase in fluorescence relative to DNA alone (blue). Addition of 0.5 mg/ml heparin during the preincubation lowers this signal and 1.0 mg/ml heparin (yellow) leads to a fluorescence similar to DNA alone suggesting it completely prevents UvrD1-DNA binding at this concentration. (B) Using a 3'-end labeled DNA template instead of a 5'-end labeled template results in a decrease in fluorescence upon mixing pre-bound UvrD1 with ATP and Mg^{2+}. Fluorescence has been normalized by dividing by the signal from a DNA only trace.



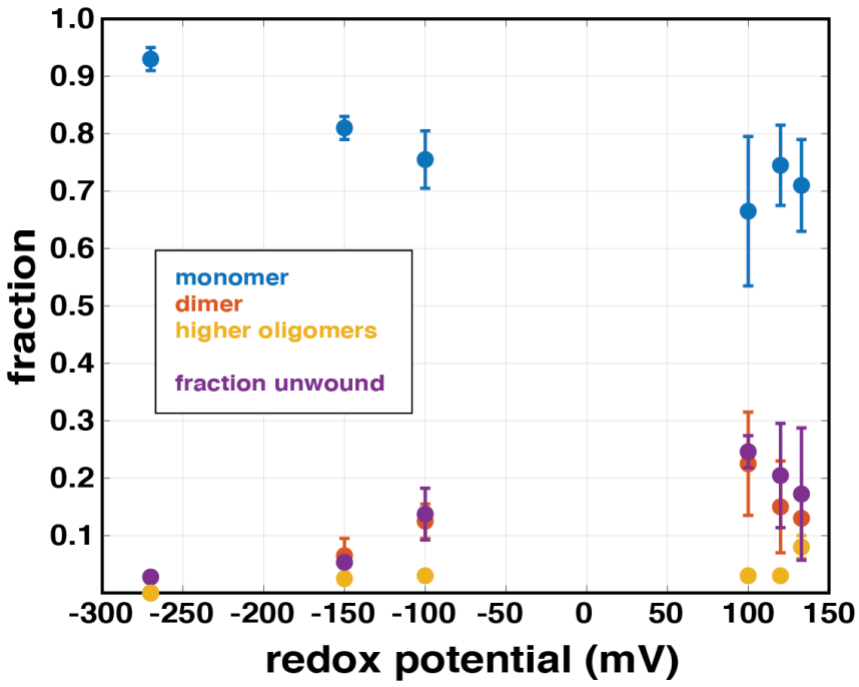
Supplementary Figure 3.16: Translocation traces from UvrD1 in the absence of reducing agent. The presence of both monomeric and dimeric species results in translocation kinetics distinct from that of monomer alone (see main text Fig. 6). Fit parameters to an n-step sequential model with two populations of translocating species can be found in Supplemental Table 4. These traces were normalized to the average of last ten plateau values observed in the raw data at each DNA length, for ease in fitting and visualization.



Supplemental Figure 3.17: N-step kinetic scheme for fitting translocation data. Translocation steps forward occurs with a rate k_t while dissociation from internal sites occurs with a rate k_d and dissociation from the 5'-end of the template occurs with a rate k_{end}.



Supplemental Figure 3.18: Dissociation kinetics of WT UvrD1 in the presence and absence of DTT measured with tryptophan fluorescence. UvrD1 dissociation kinetics was measured from $(dT)_{100}$ for WT+DTT (blue) and $poly(dT)_{964}$ for WT-DTT (red) (Methods). Excess UvrD1 was prebound to the dT DNA oligos and mixed with identical concentrations of ATP, heparin, and Mg^{2+} as those in the translocation assays. Mixing led to UvrD1 dissociation as monitored by an increase in tryptophan fluorescence measured using 290nm excitation and 305nm long pass emission filter. Fitting both traces with a single exponential function (black lines) yielded similar estimates for the observed dissociation rates for both the +DTT condition ($8.1 \pm 0.2\ s^{-1}$) and the -DTT condition ($7.1 \pm 0.2\ s^{-1}$). These values were used to constrain the observed dissociation rates for the monomer and dimer, respectively for the global fitting of translocation traces.



Supplemental Figure 3.19: Dimer fraction correlates with fraction of DNA unwound. The equivalent of Fig. 7 in the main text but using 2 μM UvrD1 1A1B double mutant.

Full Genus names for species listed in Fig. 3B

Escherichia coli
Escherichia coli
Bacillus subtilis
Mycobacterium tuberculosis
Clostridioides difficile
Gordonia bronchialis
Dietzia cinname
Tsukamurella paurometabola
Hoyosella rhizosphaerae
Tomotella caverna
Mycobacterium ulcerans
Rhodococccus hoagie
Streptomyces regensis
Stackebrandtia nassauensis
Actinocatenispora thailandica
Longimycelium tulufanens
Actinopolymorpha pittospori
Terrabacter lapilli
Actinomadura logoneensis

Supplemental Table 3.1: Summary of sedimentation velocity results obtained in the absence of DNA. The hydrodynamic properties of each oligomeric form for both wild-type and cysteine mutant constructs of *Mtb* UvrD1 at 2.5 μM protein concentration. Unless specified, all quantifications presented were obtained in Buffer A with 75 mM NaCl at 25 ºC.

condition	MW (Da) calculated	MW (Da) measured	$s_{20,w}$ measured	f/f_0 measured	fraction oligomer
Monomers					
WT –DTT	85,049	77,000	4.72	1.17	0.3
WT +DTT	85,049	82,600	4.72	1.17	0.9
WT 400 mM NaCl –DTT	85,049	68,100	4.22	1.2	0.6
WT 400 mM NaCl +DTT	85,049	74,600	4.34	1.19	0.92
2B –DTT	85,017	88,100	4.61	1.17	0.9
2B +DTT	85,017	83,000	4.51	1.19	0.92
1A1B +DTT	85,146	86,200	4.72	1.2	0.94
1A1B –DTT	85,146	81,000	5.14	1.21	0.3
Dimers					
WT –DTT	170,098	180,020	7.55	1.17	0.3
1A1B –DTT	170,292	185,000	7.55	1.17	0.6
WT 400mM NaCl –DTT	170,098	131,000	6.79	1.2	0.3
Tetramers					
WT –DTT	340,196	367,000	9.54	1.17	0.33
1A1B +DTT	340,584	373,000	14.5	1.2	0.015
1A1B –DTT	340,584	333,000	14.5	1.17	0.05
WT 400 mM NaCl –DTT	340,196	298,000	14.3	1.2	0.03
2B –DTT	340,068	379,000	13.63	1.17	0.03
2B +DTT	340,068	380,000	13.62	1.19	0.028

Supplemental Table 3.2: Parameter estimates from fits to unwinding data. Unwinding parameters obtained from fits of data shown in Fig. S10A to the model in Fig. S10B. Unwinding rate (k_u), kinetic step-size (m), overall unwinding rate (mk_u), dissociation rate (k_d), non-productive isomerization rate (k_{np}), fraction productive complexes (x_p), processivity (p) and base-pairs unwound per binding event are shown as independently obtained from three replicate experiments along with estimated fit errors. The average of each parameter is shown in the last column along with the standard deviation of the three measurements.

Parameter replicate:	**first**	**second**	**third**	**average**
k_u (s^{-1})	58.5 ± 1.4	76.0 ± 1.1	82.6 ± 1.3	72.4 ± 12.5
m (bp/step)	1.01 ± 0.06	0.83 ± 0.18	0.84 ± 0.38	0.89 ± 0.10
mk_u (bp-s^{-1})	59.1 ± 3.8	63.1 ± 13.7	69.4 ± 28.6	63.9 ± 5.2
k_d (s^{-1})	2.3 ± 1.6	4.0 ± 1.4	3.4 ± 1.9	3.2 ± 0.9
k_{np} (s^{-1})	3.70 ± 0.02	3.30 ± 0.03	3.60 ± 0.05	3.53 ± 0.21
x_p	0.58 ± 0.002	0.66 ± 0.003	0.56 ± 0.004	0.60 ± 0.05
Processivity (P)	0.95 ± 0.02	0.95 ± 0.007	0.95 ± 0.002	0.95 ± 0.01
bp unwound per binding event	20 ± 10	20 ± 3	20 ± 1	20.0 ± 0.01
Chi^2	0.15	0.27	0.23	

Supplemental Table 3.3: Summary of sedimentation velocity results obtained in the presence of DNA. The hydrodynamic properties for labeled DNA alone and DNA complexes with either monomer or dimer WT and 2B mutant UvrD1 at 1.5 or 4 μM protein concentration in case of 2B mutant. All quantifications presented were obtained in Buffer A with 75 mM NaCl at 25 ºC in the absence of DTT.

Species	MW (Da) calculated	MW (Da) measured	$s_{20,w}$ measured	f/f0 measured	fraction DNA/ oligomer	v (ml/g) calculated
3'-dT(10)-ds18-cy5	14500	10600	2.41	1.21	0.97	0.563
3'-dT(20)-ds18-cy5	19000	12800	2.2	1.47	0.92	0.563
$(2B\text{-}UvrD1)_1$ + 3'-dT(20)-ds18-cy5	104008	90800	5.2	1.19	0.47	0.708
$(WT\text{-}UvrD1)_1$ + 3'-dT(10)-ds18-cy5	99508	104000	5.55	1.21	0.35	0.709
$(WT\text{-}UvrD1)_1$ + 3'-dT(20)-ds18-cy5	104008	96300	5.55	1.18	0.31	0.725
$(WT\text{-}UvrD1)_2$ + 3'-dT(20)-ds18-cy5	199000	162000	8.07	1.18	0.31	0.725

Supplemental Table 3.4: Parameter estimates from fits of translocation data. Parameters are as follows: k_t is translocation rate, k_d is dissociation rate from internal sites, k_{end} is the dissociation rate from the ends, r is the binding probability ratio (random/end), m is kinetic step size, d_{app} is the apparent contact size, mk_t is the macroscopic translocation rate, f_m is the fraction monomer and f_d is the fraction dimer population for the mixed fits where both monomer and dimer populations are present, k_{t2} is the translocation rate for the dimer, k_{d2} is the dissociation rate for the dimer, k_{end2} is the rate of dissociation from the ends, r2 is the probability ratio (random/end) for the dimer, Chi^2 is the goodness of fit, and processivity is calculated from the rates of translocation and dissociation and is the average number of steps along the template taken prior to dissociation.

parameter	WT +DTT (average)	2B –DTT (average)	WT –DTT (average)
k_t (s^{-1})	43 ± 6	43 ± 1	40
k_d (s^{-1})	3.8 ± 0.3	4.1 ± 0.1	4.1
k_{end} (s^{-1})	63 ± 3	61 ± 4	61
r	2.4 ± 0.5	2.6 ± 0.2	2.6
m (nt)	3.0 ± 0.5	3.0 ± 0.7	0.8 ± 0.1
d_{app} (nt)	13 ± 1	12 ± 1	23 ± 5
mk_t (nt/s)	120 ± 5	130 ± 10	71 ± 15
f_m	-	-	0.3
f_d	-	-	0.7
k_{t2} (s^{-1})	-	-	88 ± 6
$k_{d2,obs}$ (s^{-1})	-	-	2 ± 1
$k_{end2,obs}$ (s^{-1})	-	-	7 ± 3
r2	-	-	0.5 ± 0.1
processivity	32 ± 1	32 ± 5	53 ± 43
Chi^2	0.4 ± 0.1	0.2 ± 0.1	1 ± 0.3

Supplementary Table 3.5: DNA primers used for creating mutant UvrD1 plasmids.

Name	Sequence	Construct used for
BamH1 forward WT UvrD1	5'CCCGGATCCATGAGTGT GCACGCGACCG 3'	Wild type UvrD1
Hind III reverse WT UvrD1	5'CGCAAGCTTTCAGAGCT TGGTGACAGGGGCGTGGT TG 3'	
451 cysteine to alanine forward	5'GATCGTGCCGAGGCGGC AGTGGCGGTGTACGCCGA GAAC 3'	2B mutant UvrD1 (C451A)
451 cysteine to alanine reverse	5'GTTCTCGGCGTACACCG CCACTGCCGCCTCGGCAC GATC 3'	
269 cysteine to threonine forward	5'CCCGGCGAGTTGACCGT CGTCGGGGATGCCG 3'	1A1B double mutant UvrD1 (C107T/C269T)
269 cysteine to threonine reverse	5'CGGCATCCCCGACGACG GTCAACTCGCCGGG 3'	
107 cysteine to threonine forward	5'GACGTTTCACTCCACCA CGGTGCGTTATCCTTGCGC AACC 3'	
107 cysteine to threonine reverse	5'GGTTGCGCAAGGATAAC GCACCGTGGTGGAGTGAA ACGTC 3'	

Supplemental Table 3.6: Sequences of single-stranded DNA oligomers used for translocation and DNA unwinding studies.

Length (nt)	Sequence	Sequences used for
18	5'-GTT GGT CGG CAG CAG GGC-3'	Trap
18	5'-GCC CTG CTG CCG ACC AAC -3'	Anisotropy assays
38	5'-56-FAM GTT GGT CGG CAG CAG GGC-dT(20)3'	
18	5'-GCC CTG CTG CCG ACC AAC-BHQ_2 -3'	Double-stranded DNA unwinding assays
38	5'-cy5 GTT GGT CGG CAG CAG GGC-dT(20)3'	
21	5′-GCC CTG CTG CCG ACC AAC GAT- (BHQ_2)-3′	
41	5'-cy5 ATC GTT GGT CGG CAG CAG GGC-dT(20)3'	
25	5′-GCC CTG CTG CCG ACC AAC GAT GGT T- (BHQ_2)-3′	
45	5'-cy5 AA CC ATC GTT GGT CGG CAG CAG GGC-dT(20)3'	
32	5′-GCC CTG CTG CCG ACC AAC GAT GGT TAC ATT CC (BHQ_2)-3′	
52	5'-cy5 GG AAT GTA ACC ATC GTT GGT CGG CAG CAG GGC-dT(20)3'	
40	5′-GCC CTG CTG CCG ACC AAC GAT GGT TAC ATT CCC GCT GCT G (BHQ_2)-3′	
60	5'-cy5 C AGC AGC GGG AAT GTA ACC ATC GTT GGT CGG CAG CAG GGC-dT(20)3'	
20	5'-cy3 dT(20)-3'	Single-stranded DNA translocation assays
35	5'-cy3 dT(35)-3'	
45	5'-cy3 dT(45)-3'	
75	5'-cy3 dT(75)-3'	
104	5'-cy3 dT(104)-3'	

Chapter 4: *Mycobacterium tuberculosis* Ku Stimulates Multiround DNA Unwinding by UvrD1 Monomers.

4.1 Abstract

Mycobacterium tuberculosis is the causative agent of Tuberculosis. During the host response to infection, the bacterium is exposed to both reactive oxygen species and nitrogen intermediates that can cause DNA damage. It is becoming clearer that the DNA damage response in *Mtb* and related actinobacteria function via distinct pathways as compared to well-studied model bacteria. For example, we have previously shown that the DNA repair helicase UvrD1 is activated for unwinding via a redox-dependent dimer. In addition, mycobacteria contain a homodimeric Ku protein, homologous to the eukaryotic Ku70/Ku80 dimer, that plays roles in double-stranded break repair via non-homologous end-joining. Independent studies have shown that Ku stimulates the helicase activity of UvrD1, but the molecular mechanism, as well as which redox form of UvrD1 is activated, is unknown. Although UvrD1 monomers are not able to processively unwind DNA under single-round conditions, we show here that these monomers can slowly and inefficiently unwind DNA (*i.e.*, 100x slower than the dimer) under multiround conditions and that Ku specifically enhances this activity to unwind ~15% of duplex DNA templates. We also demonstrate that the UvrD1 C-terminal Tudor domain is required for the formation of a Ku-UvrD1 protein complex and activation. We also observe that *Mtb* Ku coats the duplex region of the template cooperatively and that, while one Ku bound activates unwinding, maximal activation is only observed when the template is saturated with Ku molecules. Our observations reveal aspects

of the interactions between DNA, *Mtb* Ku, and UvrD1 and highlight the potential role of UvrD1 in multiple DNA repair pathways through different mechanisms of activation.

4.2 Introduction

About one-third of the human population is estimated to be infected with *Mycobacterium tuberculosis* (*Mtb*), the causative agent of Tuberculosis Click or tap here to enter text.[1]Click or tap here to enter text.. During infection, the host immune response exposes *Mtb* to reactive oxygen species and reactive nitrogen intermediates that can lead to DNA damage Click or tap here to enter text.[2,3]Click or tap here to enter text.. This damage occurs in the form of base modifications, single-strand breaks, and double-strand breaks (DSB) Click or tap here to enter text.[4,5]Click or tap here to enter text.. While single-stranded breaks can be bypassed by replication and transcription machinery, specialized enzymes are required to repair DSBs Click or tap here to enter text.[6,7]Click or tap here to enter text.. Mycobacteria has three known DSB repair pathways: single-strand annealing (SSA), homologous recombination (HR), and nonhomologous end joining (NHEJ) [8,9]. Repair can proceed through SSA if repeating homologous sequences are available on each side of the break. Resection leads to complementary single strands that can anneal and subsequent flap removal and end-joining lead to repair with deletion of the resected sequences [8]. HR is a high-fidelity DNA repair pathway but may only function when a second intact DNA copy is present; for example, during or after replication and before division [10,11]. NHEJ can either be error-free (if the ends are sealed directly) or mutagenic depending on how the DNA ends are processed prior to ligation [9]. Since NHEJ acts as the major pathway in non-replicating cells, it may be of significant relevance in persistence and pathogenesis of *Mtb* [12–14].

Prokaryotic NHEJ is accomplished through the action of two proteins: Ku and Ligase D (LigD). Bacterial Ku proteins are ~30 kDa, form stable homodimers, bind double-stranded DNA, and are responsible for identifying, binding, and protecting DSBs [15,16]. The core domain of bacterial Ku is homologous to that of the eukaryotic Ku70:Ku80 heterodimer and is highly conserved among bacterial Ku's from different species. Each core domain forms one-half of a dimeric ring-like structure within which DNA binds [17–19]. In contrast, the C-terminal domain (CTD) is unique to bacterial Ku and consists of 20-25 conserved amino acids and an extended region of basic residues of varying lengths (i.e., 14 in *Mtb* to 40 in *M. smegmatis*). The conserved sequence is important for recruiting LigD to DNA ends and the extended basic region has also been implicated in DNA binding and threading [20].

DNA helicases play varied roles in DNA repair including directly unwinding duplex DNA, removing damaged ssDNA, stimulating the resection of blunt-ended substrates, and removing DNA-bound proteins [21–23]. Many bacterial helicases like UvrD, UvrD1, Rep, and PcrA, and eukaryotic helicases and helicase complexes like TFIIH, FANCJ, and WRN are involved in DNA repair pathways [24–29]. *Mtb* UvrD1 belongs to the UvrD/PcrA subgroup of SF1 super-family helicases and consists of two RecA-like domains (1A and 2A) with two accessory domains (1B and 2B). Members of this subgroup have been shown to unwind DNA as dimers or as monomers in the presence of accessory factors [23,30–34]. In addition, they sometimes contain a C-terminal Tudor domain which has been implicated in protein-protein interactions with binding partners [35,36]. We previously showed that *Mtb* UvrD1 exists as a mixture of monomers and dimers depending on redox potential [37]. Although both monomers and dimers of UvrD1 can bind and translocate on ssDNA, only the dimer formed in oxidative conditions via a specific disulfide bond between 2B domains can unwind DNA [37].

Studies performed prior to the elucidation of the redox-dependent dimerization and activation of UvrD1 showed that *Mtb* Ku stimulates the helicase activity of UvrD1 and

provided evidence for protein-protein interaction between the Ku and the CTD of UvrD1 [38–40]. Here we show that Ku specifically stimulates the helicase activity of UvrD1 monomers that can only be observed under multiround binding conditions. This activity is slow (~100x slower compared to the dimer) and inefficient (~15% fraction of templates unwound in the presence of saturating Ku). Furthermore, removing the C-terminal Tudor domain of UvrD1 abrogates the Ku-dependent activation. Lastly, *Mtb* Ku coats the dsDNA region of our unwinding templates even in the presence of 3' single-stranded overhangs. This binding is dependent on magnesium, is cooperative, and Ku-stimulated UvrD1 unwinding increases as the stoichiometry of bound Ku increases achieving a maximum of activation with a fully coated DNA duplex. These data serve to define the molecular mechanism for Ku-based activation of UvrD1 while quantitatively comparing the slow Ku-stimulated activity of monomeric UvrD1 in multiround conditions to the rapid and efficient single-round (*i.e.*, processive) activity of oxidatively formed dimers. That said, our measurements highlight a possible role of UvrD1 monomers in remodeling and processing Ku-bound templates during DSB repair.

4.3 Material and methods

Protein purification

The open reading frame encoding *M. tuberculosis* Ku (Rv0937c) was cloned in pET45b vector with ampicillin resistance using a BamH1 site at the start codon and a HindIII site 3′ of the stop codon. The plasmids were confirmed by sequencing to exclude the acquisition of unwanted coding changes. The pET-Mtb Ku plasmid was transformed into *E. coli* BL21(DE3). Ku was overexpressed by induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at OD of 0.5, followed by incubation at 23°C for 16 h with constant shaking overnight. The cells were harvested by centrifugation, and the pellets were

either stored at -80 °C or used for subsequent procedures that were performed at 4 °C. An ~12 g cell pellet obtained from 1 L of liquid culture was resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 0.25 M NaCl, 10% sucrose). The lysates were sonicated, and insoluble material was removed by centrifugation at 14K for 45 minutes. The soluble extracts were applied to 2 ml columns of nickel-nitrilotriacetic acid agarose (Ni-NTA) (QIAGEN catalogue no. 30210) that had been equilibrated with lysis buffer. The columns were washed with 10X column volume of wash buffer (50 mM Tris-HCl, pH 8.0, 0.25 M NaCl, 10% glycerol) and then eluted stepwise with wash buffer containing 50-, 100-, 200-, 500-, and 1000-mM imidazole. The polypeptide compositions of the column fractions were monitored by SDS-PAGE. Ku was recovered in 100- and 200 mM imidazole fractions. After overnight dialysis in (50 mM TRIS-HCL pH 8.0, 60 mM NaCl, 10% glycerol) the protein was passed through 200 mM NaCl in DEAE-Sephacryl chromatography followed by size exclusion chromatography where it eluted as a dimer. *Mtb* UvrD1 proteins both wild type and C-terminal deletion mutant were expressed and purified as described previously [37].

DNA substrates

Single-stranded DNA, which is either labelled with Cy5, Cy3, FAM or BHQ2 were ordered from Integrated DNA technologies (IDT). Duplex DNA substrates were prepared by annealing ssDNA oligos with the fluorescent label at 5'end of the single strand was mixed with an equimolar concentration of unlabelled, BHQ2 or FAM labelled complementary strand in 10 mM Tris pH 8.0, 50 mM NaCl, followed by heating to 95 °C for five minutes and slow cooling to room temperature.

Stopped flow unwinding assays

All stopped-flow unwinding experiments were carried out at 25°C using an Applied Photophysics instrument SX-20, minimum total shot volume 100 µl, dead time 2 ms. The

experiments were carried out in Buffer with TRIS pH 8.0, 75 mM NaCl, 5mM $MgCl_2$, 20% glycerol and with 1 mM DTT for reducing conditions. The DNA substrates used in the assay are double-stranded 18-basepair or 32-basepair DNA with a dT20, dT40, or dT60 tail where Cy5 fluorophore is attached to the 5' end of the long strand and the black hole quencher (BHQ_3) is attached to the 3' end of the short strand. Cy5 fluorophore was excited using 625 nm LED (Applied Photophysics Ltd., Leatherhead, UK) and its fluorescence emission was monitored at wavelengths >665 nm using a long-pass filter (Newport Optics). DNA unwinding was monitored as the increase in Cy5 fluorescence upon DNA strand separation. The traces represent the average of 5 independent shots and at least two different protein purifications. UvrD1 (400 nM) is incubated with or without Ku (1600 nM) and 20 nM DNA in one syringe and is rapidly mixed with the contents of the other syringe: containing 2 mM ATP, 10 mM Mg^{+2}, and either no TRAP for multiround experiments or one of three single-stranded DNA TRAPs (an 18bp-dT20 partial duplex or a dT40-10bp hairpin TRAP) in excess of protein (25X, 5 μM). In the case of Ku-stimulated unwinding, the slow rate of DNA unwinding is monitored over 2400 seconds. The unwinding signal is normalized to the signal from positive and negative controls to calculate the fraction of DNA unwound. The signal for fully unwound DNA is alternatively obtained using 20 or 50 nM of the labeled partial duplex denatured in the presence of TRAP or unannealed 20 or 50 nM single-stranded Cy5 labeled DNA in presence of UvrD1 and Ku. The signal for fully annealed DNA is obtained from 20 nM duplex DNA with Cy5 and the BHQ_3 in the presence of UvrD1 and Ku but the absence of ATP.

Fluorescence intensity measurements and binding density analysis

All fluorescence titrations were performed using a spectrofluorometer (ISS, Champaign, IL). Titrations were performed in buffer containing TRIS pH 8.0, 75 mM NaCl, 20% glycerol, 5 mM $MgCl_2$, 2 mM DTT and 9 μM BSA. Measurements of total fluorescence intensity of FAM-labelled 32bp double-stranded DNA at 20, 50 and 300 nM concentrations with T20

single-stranded tail or dT20 ssDNA alone at 50 nM concentrations were recorded using excitation and emission wavelengths of 490 and 520 nm, respectively. After correcting for dilution, model independent binding density analysis was performed to calculate number of Ku dimers bound to DNA substrate [41,42].

For binding density calculations, we performed titrations of Ku dimers (ligand, X) using fluorescein labelled (FAM) DNA referred to here as the macromolecule (M). S is the quenching signal linked to binding.

If M_f = free M, M_t= total M, X_t = ligand and *n* is the binding stoichiometry, then

$$nX + M \leftrightarrow MX_n \qquad (4.1)$$

The total macromolecular concentration is equal to the free concentration plus the sum of the concentrations of each bound species (i = 1 to n):

$$M_t = M_f + \sum_{i=1}^{n} MX_i \qquad (4.2)$$

The observed signal (S_{obs}) will be equal to a sum of the signals from each species:

$$S_{obs} = S_f M_t + \sum_{i=1}^{n} S_i\, MX_i \qquad (4.3)$$

The binding density (<$\bar{x}$>) is defined as the average number of sites bound such that

$$\langle \bar{x} \rangle = \sum_{i=1}^{n} i\, MX_i / M_t \qquad (4.4)$$

Substituting equation 4 and 2 into 3 and rearranging, one obtains

$$\Delta S_{obs} = \left(S_{obs} - S_f M_t\right)/S_f M_t = \sum_{i=1}^{n} (\Delta S_i / i)\, \langle \bar{x} \rangle \qquad (4.5)$$

$S_f M_t$ represents the initial signal from M_t before the start of the titration, ΔS_{obs} is the experimentally determined fractional molar signal change, and $\Delta S_i = (S_i - S_f)/S_f$. $(\Delta S_i/i)$ is then the signal change per bound ligand in a complex containing i ligands.

Thus, ΔS_{obs} is only a function of binding density distribution since $(\Delta S_i/i)$ is an intrinsic property of the system. Therefore, at constant ΔS_{obs}, the binding density ($\langle\bar{x}\rangle$) is also constant. Since $\langle\bar{x}\rangle$ depends thermodynamically only on the free ligand concentration X_f at equilibrium, at constant ΔS_{obs} also implies constant X_f.

To calculate binding density, titrations at different total macromolecule concentrations are performed such that system is in equilibrium and the assembly state of ligand does not change. At each pair of total concentrations $\{M_t, X_t\}$, there then exist corresponding values for $\langle\bar{x}\rangle$ and X_f and at each $\langle\bar{x}\rangle$, the experimental signal (ΔS_{obs}) is constant.

$$X_t = X_f + \langle\bar{x}\rangle M_t \qquad (4.6)$$

By plotting X_t versus M_t at constant ΔS_{obs}, one obtains X_f as the y-intercept and $\langle\bar{x}\rangle$ as the slope. This then allows for a conversion of the experimental signal (ΔS_{obs}) to binding density $\langle\bar{x}\rangle$, and an identification of the number of ligand molecules bound at saturation.

Analytical ultracentrifugation

Analytical ultracentrifugation sedimentation velocity experiments were performed using a Proteome Lab XL-A analytical ultracentrifuge equipped with an An50Ti rotor (Beckman Coulter, Fullerton, CA). The sample (380 µl) and buffer (410 µl) were loaded into each sector of an Epon charcoal-filled two-sector centerpiece. All sedimentation velocity experiments were performed at 25 °C and 42,000 rpm. Absorbance data were collected by scanning the sample cells at intervals of 0.003 cm, monitoring either at 280 nm for protein

absorbance or 650nm for Cy5 absorbance. Absorbance signal of protein alone or DNA-protein complexes were maintained between 0.1 and 1.

Continuous sedimentation coefficient distributions, *c*(*s*), were calculated using SEDFIT[43,44]. This analysis yielded individual sedimentation coefficients for each monomer, dimer, and higher oligomer species of proteins as well as a weighted average frictional coefficient (f/f_o) for the entire distribution. Calculated sedimentation coefficients were converted to 20 °C water conditions ($s_{20,w}$) according to:

$$s_{20,w} = s_{\exp} \frac{\eta_{\exp}}{\eta_{20,w}} \left(\frac{1-\bar{v}_{20}\,\rho_{20,w}}{1-\bar{v}_{\exp}\,\rho_{\exp}} \right) \qquad (4.7)$$

)where $\rho_{20,w}$ and $\eta_{20,w}$ are density and viscosity of water at 20 °C, $\rho_{\exp}$ and $\eta_{\exp}$ are density and viscosity of the buffer at the experimental temperature of 25 °C , and $\bar{v}_{20}$ and $\bar{v}_{\exp}$ are partial specific volumes of the protein at 20 °C and at 25 °C. Buffer densities, ($\rho_{\exp}$) and viscosities ($\eta_{\exp}$) were calculated from buffer composition using SEDNTERP[45]. Partial specific volumes ($\bar{v}_{\exp}$) for *Mtb* UvrD1 and Ku were calculated in SEDNTERP using the amino acid composition. Integration of the entire c(s) distribution vs. the integration of an individual sedimentation species was performed and used to calculate the population fraction [43].

For AUC experiments done in the presence of Cy5 labeled DNA, the absorbance signal was collected by scanning the sample cells at 650 nm. Partial specific volumes ($\bar{v}_{\exp}$) for labeled DNA and the UvrD1-DNA complex were calculated according to

$$\bar{v} = \frac{(\sum_{i=1}^{n} n_i M_i \bar{v}_i)}{\sum_{i=1}^{n} n_i M_i} \qquad (4.8)$$

4.4 Results

4.4.1 Ku does not stimulate UvrD1 dimer- or monomer-catalyzed DNA unwinding on short timescales.

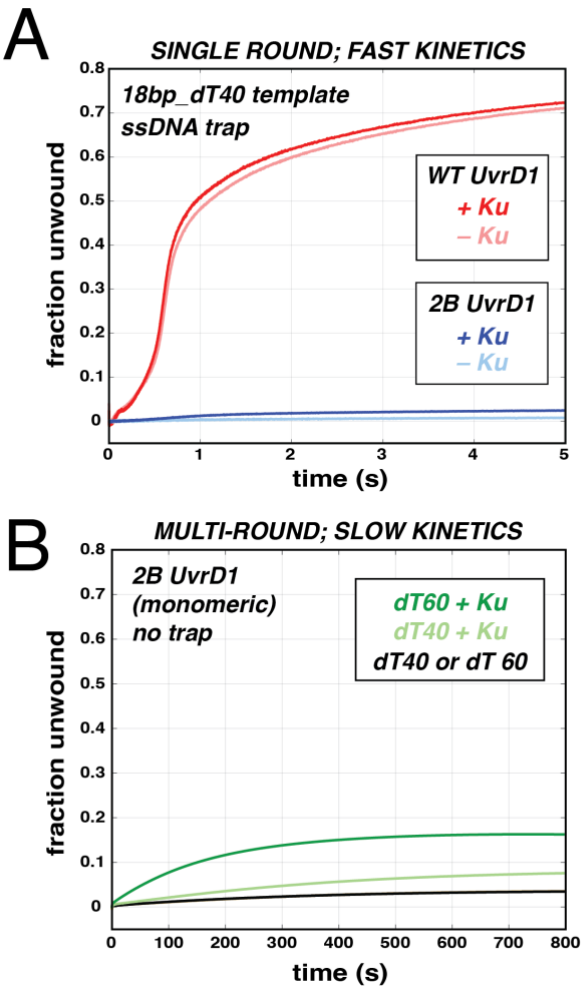


Figure 4.1: Ku stimulates unwinding of 2BUvrD1 on longer timescales and in multiround conditions. (A) Single-round unwinding of an 18bp-dT40 DNA with WTUvrD1 (red) and 2BUvrD1 (C451A, blue) in the presence and absence of 800 nM Ku at short timescales. (B) Multiround unwinding of 18bp duplex with either dT40 or dT60 ssDNA tails with monomeric UvrD1 (2BUvrD1) in the presence (green) or absence (black) of Ku over long timescales. In these experiments, the final concentrations of components are 10 nM DNA substrate, 200 nM UvrD1 and 800 nM Ku.

We have previously shown that *Mtb* UvrD1 forms a disulfide-bonded dimer under oxidative conditions via a 2B domain cysteine and that the dimeric form is an active helicase [37]. The monomeric form of UvrD1 – generated either via reducing conditions or by mutation of cysteine 451 in the 2B domain to alanine – can bind and translocate on single-stranded DNA but lacks helicase activity under previously tested conditions [37]. As it has been reported that Ku stimulates UvrD1 helicase activity [38], we performed DNA unwinding experiments with WT UvrD1 under conditions where dimer formation is favored in the presence and absence of Ku. Specifically, we ran stopped-flow kinetic assays using a 18bp-dT40 template labeled with a Cy5 dye and a black hole quencher (BHQ) on the 5' and 3' termini of the blunt end of the duplex respectively. After rapidly mixing a pre-incubated protein-DNA complex with a solution containing 2 mM ATP, 10 mM $MgCl_2$, and 5 µM ssDNA trap, we observed an increase in Cy5 fluorescence due to unwinding of the 18 bp duplex. However, contrary to our expectations, we did not observe an increase in either the rate or extent of unwinding in the presence of 800 nM concentrations of Ku (Figure 4.1A, red curves). We hypothesized that perhaps Ku specifically activates the monomeric form of UvrD1 that we previously observed to completely lack helicase activity under these conditions.

Stopped-flow experiments were performed with the 2B domain mutant enzyme (C451A, an obligate monomer) to ask whether monomeric UvrD1 could be activated by Ku. Consistent with our previous work, monomer UvrD1 alone displayed no unwinding activity. In addition, we observed no activity upon the addition of 800 nM Ku (Figure 4.1A, blue curves). This was true under both single-round and multiround conditions at these timescales (Supplemental Figure S4.1A). Since the previous Ku-dependent study presented end point measurements taken after 10 minutes of reaction time, we next asked whether Ku-dependent activation may take place on longer timescales compared to the rapid unwinding exhibited by dimeric UvrD1 [37].

4.4.2 Slow unwinding via multiple binding events is catalyzed by UvrD1 monomers and enhanced by Ku.

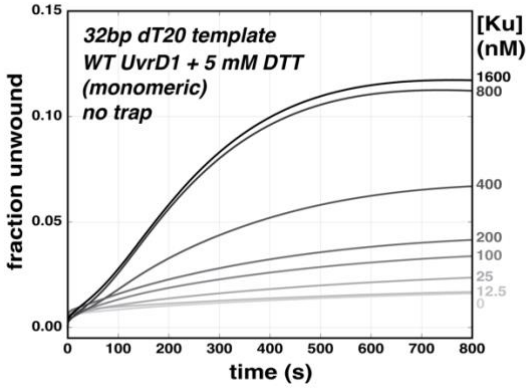


Figure 4.2: Ku stimulates unwinding by UvrD1 monomers in a concentration dependent manner. Plots in the curve show unwinding by a monomer. Increasing Ku concentrations with constant concentrations of UvrD1 monomer (200 nM) and DNA (50 nM) increases unwinding by UvrD1 monomer.

Dimeric UvrD1 unwinds 18 bp duplexes with a dT20 tail in just over a second (Figure 1A, red) [37]. To test the idea that Ku exhibits its effect on longer timescales, we monitored unwinding reactions for over 12 minutes. Consistent with previous observations, unwinding was only observed for UvrD1 dimers under single-round conditions even at these longer timescales (Supplemental Figure S4.1B). However, upon the removal of ssDNA trap, a small amount (~3%) of multiround unwinding was observed with the monomeric 2B mutant enzyme (Figure 4.1B, black). The dependence on the removal of trap indicates that this activity is catalyzed by multiple binding events of monomeric UvrD1 in contrast to the processive unwinding displayed by a single dimer. On dT40 and dT60 tailed templates, the monomer exhibited very similar activities (Figure 4.1B, overlapping black curves). We next asked whether Ku might specifically potentiate this activity of the monomer under multiround conditions. Indeed, in the presence of 800 nM Ku, we observed higher degrees of unwinding and this activation increased with increasing tail lengths (Figure 4.1B, green curves). The presence of Ku alone under these conditions does not recapitulate this effect indicating that the effect is due to an enhancement of UvrD1-dependent activity (Supplemental Figure S4.2).

As would be expected, this Ku dependent activation of monomer unwinding is not unique to the 2B mutant as it is also observed with WT monomers generated by reducing the disulfide bond between 2B domains with 5 mM DTT (Figure 4.2). In addition, multi-turnover unwinding on these time scales can be observed on longer duplexes (*e.g.*, 32 bp) and shorter ssDNA tails (*e.g.*, 20 nt), and the activation increases with increasing concentration of Ku (Figure 4.2). Taken together, these experiments indicate that Ku specifically promotes DNA unwinding by monomers of UvrD1 under multiround conditions over a time course of 10 minutes.

4.4.3 Ku-dependent activation of unwinding depends on multiple Ku binding events.

The experiments above show that multiple DNA binding events by monomeric UvrD1 are required for the unwinding observed at long timescales. However, they don't speak to whether activation depends on multiple Ku binding events. We aimed to perform unwinding assays using different traps that would effectively trap Ku. Using a DNA unwinding template (32bp-dT20) labeled with fluorescein on the 5' terminus of the blunt end, titration of Ku results in an increasing amount of fluorescence quenching (Figure 4.3A, black). However, Ku does not bind a dT20 ssDNA template (Figure 4.3A, red). These data are consistent with studies of Ku across biology indicating that Ku primarily interacts with duplex regions of DNA [46–48]. Unwinding assays were thus performed in the presence of UvrD1, Ku, and three distinct nucleic acid traps to assess whether Ku activates UvrD1 in the context of a single Ku binding event or multiple rounds of dsDNA binding.

On the dT60-18bp unwinding templates used here, ssDNA dT20 trap reduces the amount of unwinding observed relative to no trap (Figure 3B, red compared to black). However, in the presence of mixed ssDNA/dsDNA traps (*i.e.*, 18bp-dT20 or 10bp hairpin-dT40), the unwinding activity was eliminated (Figure 4.3B, blue and cyan). Given that ssDNA robustly

traps UvrD1 in the absence of Ku and Ku does not bind ssDNA alone, these results suggest the following. First, the presence of Ku reduces the preference of UvrD1 for ssDNA vs the ssDNA/dsDNA junction present in the unwinding template. Second, Ku-dependent activation depends on multiple binding events between Ku and the unwinding template.

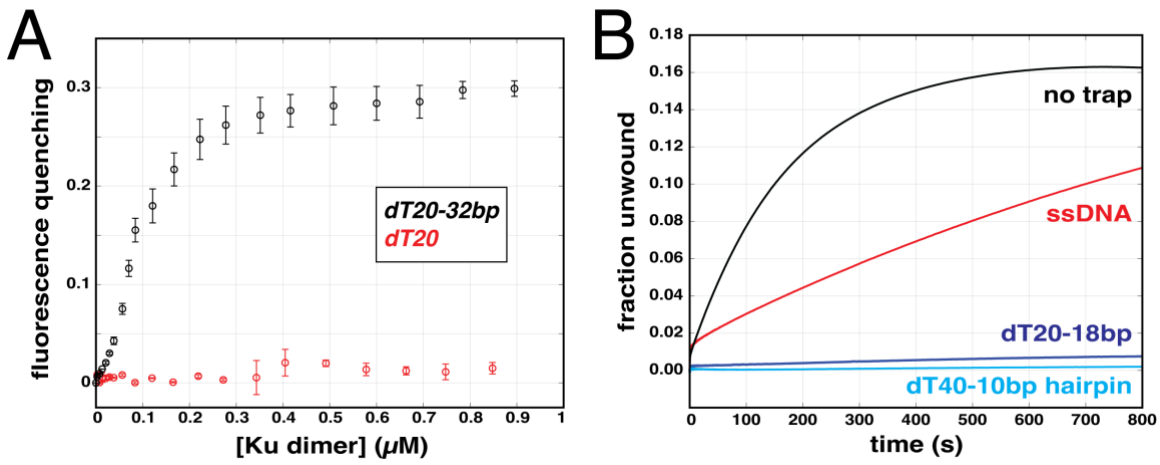


Figure 4.3: Ku-stimulated UvrD1 monomer unwinding depends on multiple Ku binding events. (A) Fluorescence quenching of 50 nM labeled DNA templates as a function of Ku dimer concentration (dT20-32bp (black), and ssDNA (red). (B) Fraction unwound of a dT60-32bp template (10 nM) as a function of time in the presence of UvrD1 (200 nM), Ku (400 nM), and no trap (black), ssDNA trap (red), dT20-18bp trap (blue), or dT40-10bp hairpin trap (cyan) 5μM each.

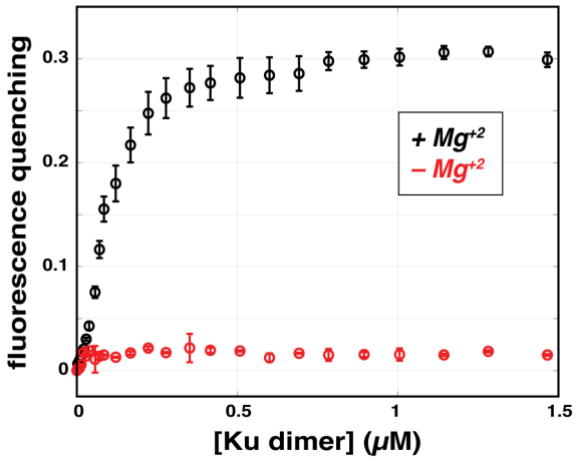


Figure 4.4: Ku-DNA interaction requires magnesium. Fluorescence quenching of a DNA unwinding DNA template labeled with FAM at the blunt end as a function of Ku dimer concentration in the presence (black) and absence (red) of 5 mM magnesium.

4.4.4 Ku binding to DNA depends on the presence of magnesium.

During our experiments to quantitate the *Mtb* Ku-DNA interaction we unexpectedly uncovered a previously unreported and dramatic dependence on magnesium. The buffer used for unwinding experiments contains 5 mM Mg^{+2} along with 1 mM ATP since both are needed for the catalytic activity of UvrD1. However, these constituents were left out of our initial attempts at measuring the DNA binding properties of Ku. Using the fluorescence quenching assay as above in absence of magnesium resulted in the complete lack of an interaction (Figure 4.4, red). Upon the inclusion of 5 mM Mg^{+2} to the binding buffer, robust DNA binding was observed (Figure 4.4, black). Since Ku is a dimer in solution (Supplemental Figure S4.3A), we asked whether the presence of Mg^{+2} influenced the distribution of Ku oligomers in the absence of DNA. AUC velocity sedimentation of 6 mM Ku in the absence and presence of magnesium show that Ku oligomerization does not depend on the presence of Mg^{+2} (Supplemental Figure S4.3B). Thus, we conclude that magnesium plays a critical role in mediating the interaction between *Mtb* Ku and DNA duplexes. All subsequent binding studies were performed in the presence of magnesium.

4.4.5 Full activation of UvrD1 depends on multiple Ku dimers bound to the template.

Ku dimers are canonically thought to preferentially bind to blunt dsDNA ends. However, multiple Ku dimers have been shown to load at DNA ends and slide along the template. Furthermore, based on previous studies, UvrD1 is positioned on the ssDNA or at the ssDNA/dsDNA junction 32 bp away from the blunt end. Therefore, we wondered whether the activation of unwinding was due to a single or multiple Ku dimers interacting with the template. To determine how many Ku dimers were interacting with the template, we performed fluorescence quenching experiments titrating Ku over different constant 32bp-dT20 DNA template concentrations. At a DNA concentration of 300 nM, we observed nearly stoichiometric binding and a linear increase in quenching as a function of Ku dimer

concentration until saturation was reached (Supplemental Figure S4.4). The crossing point of lines fit to both the increase and saturation plateau occurs at a Ku dimer/DNA ratio of 2.9, suggesting that 3 Ku dimers are bound to the DNA at saturation. Given the length of the dsDNA region of our template, this suggests a footprint of 10-11 bp which is consistent with previous measurements of other Ku systems [49].

Using titrations performed under sub-stoichiometric conditions (*i.e.*, 20 nM and 50 nM DNA) in combination with the titration at 300 nM (Figure 4.5A), we performed a binding density analysis [41,42]. Using the fact that free ligand concentration determines the average fraction bound regardless of DNA concentration, the fraction bound at each quenching signal was obtained as the slope of total Ku dimer concentration plotted as a function of DNA concentration (Figure 4.5B). The quenching signal showed a linear dependence on the fraction bound up until saturation at a stoichiometry of 2.5 Ku dimers bound per template (Figure 4.5B, inset), consistent with the independent analysis of the stoichiometric titration (Supplemental Figure S4.4) and with previous studies of multiple Ku's binding a single template [50–52]. Models based on the hypothesis of three independent binding sites did not fit the titrations well. However, cooperative binding models incorporating a positive interaction between DNA-bound Ku dimers was able to account for the data (Figure 4.5A) and resulted in an estimate of a dissociation constant (K_d) with higher nanomolar affinity and an interaction factor (w) of 1000 (see Supplemental Methods for model equations). A comparison of the Ku concentration dependence of the unwinding activity of UvrD1 shows that activation increases throughout the entire binding isotherm of Ku on the unwinding template (Figure 4.5C,D). This suggests a

model where activation begins upon the binding of the first Ku and continues to increase up to saturations as a second and third Ku bind to the template.

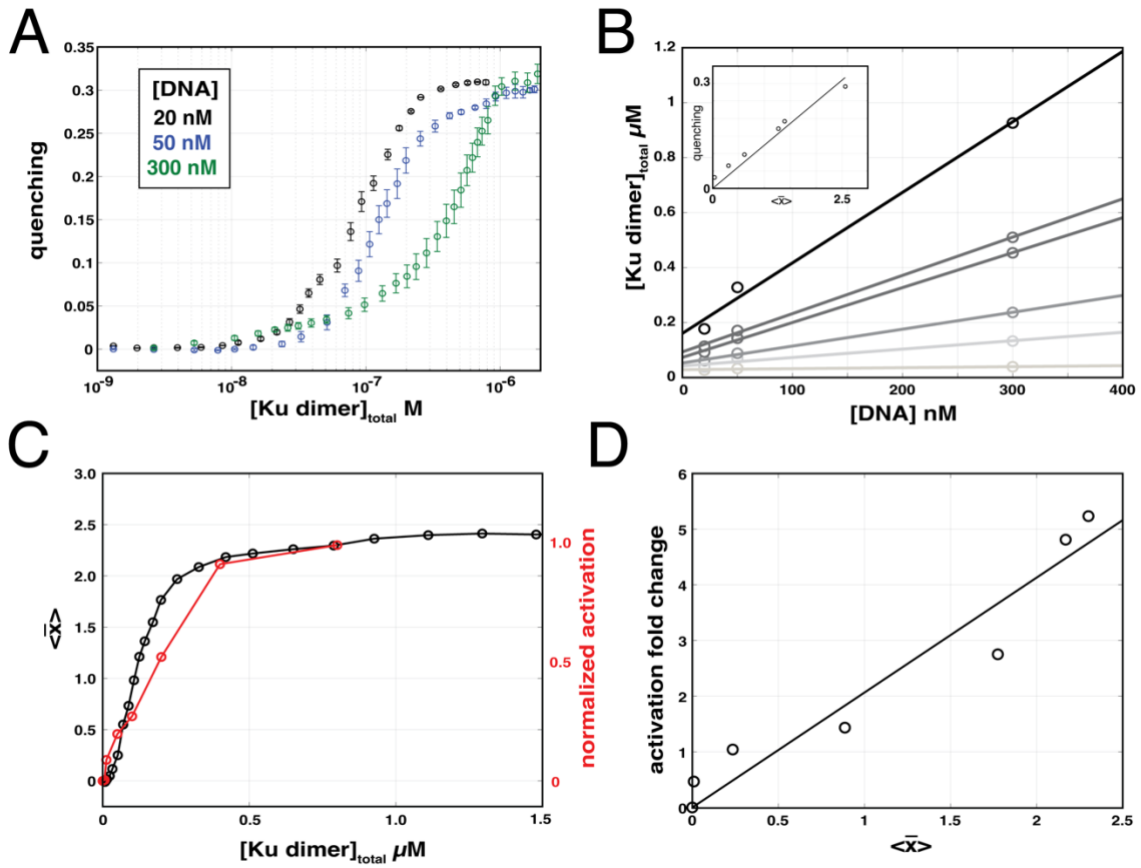


Figure 4.5: Multiple Ku dimers bind the template and contribute to UvrD1 activation. (A) Percent quenching of a FAM-labeled 32bp dT20 DNA template as a function of Ku dimer concentration using 20 nM (black), 50 nM (blue) and 300 nM (green) DNA concentrations. (B) The total Ku dimer concentration needed to reach a constant quenching signal for the three DNA concentrations show a linear dependence. The slope of each line represents an estimate of the fraction bound which allows for a plot of the quenching signal as a function of fraction bound (inset). (C) Binding density (black) and normalized unwinding activity (red, taken from the data in Figure 2) as a function of total Ku dimer concentration. (D) The activation fold-change (taken from Figure 2), as a function of binding density.

4.4.6 Mtb Ku can load onto and spread on DNA from a single-strand overhang.

Our *in vitro* unwinding template has two ends: a single-stranded end with which UvrD1 interacts and a double-stranded end with which Ku presumably interacts to load on the template. Considering an actual DSB in the cell will have only one free end it is not clear whether the

interactions we report here would be possible *in vivo*. However, previous studies have shown that *Mtb* Ku can also load onto templates with ssDNA overhangs [53].

To investigate the determinants of *Mtb* Ku DNA binding further, we performed sedimentation velocity experiments on Ku in the presence of labeled DNA with different types of ends. As expected, no Ku-ssDNA complexes were observed with 1 mM ssDNA (dT20) and 8 mM Ku (Figure 4.6A) and Ku-dsDNA complexes were observed with blunt-ended 50 bp duplex labeled internally to control for any potential dye effects on Ku binding/loading at the DNA end (Figure 6B, green). The observed complexes on the dsDNA show a broad distribution of sedimentation values consistent with the fact that multiple Ku dimers are loading onto the DNA (Figure 4.6B, green, inset). In the presence of a similarly internally labeled 50 bp dsDNA with a dT20 single-stranded tail on each end, a similar distribution of protein-DNA complexes was observed (Figure 4.6B, red, inset). These data show that *Mtb* Ku can bypass single-stranded tails of at least 20 bases and bind internal dsDNA duplex regions and suggests that, in the cell, *Mtb* Ku may load on free DNA ends regardless of the overhang status.

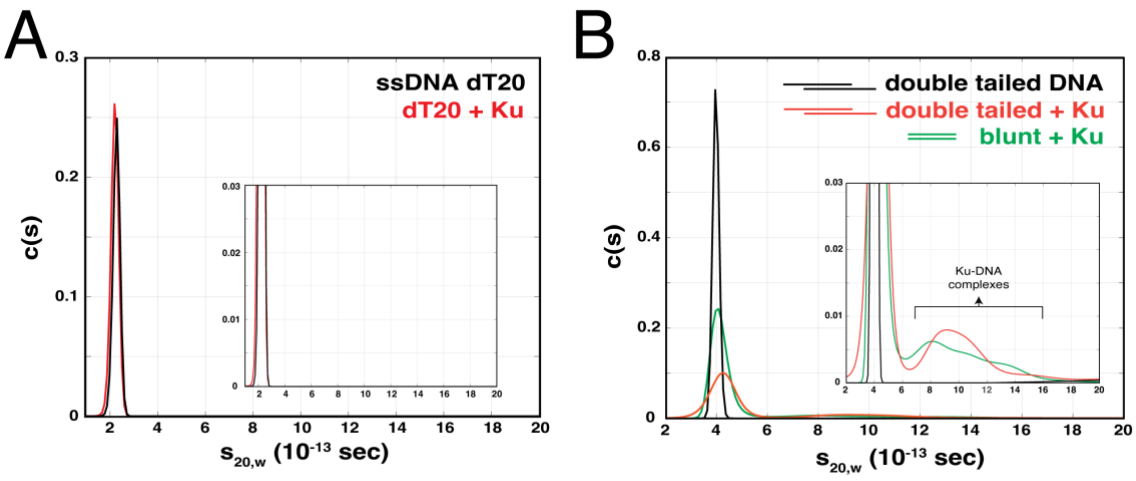


Figure 4.6: Mtb Ku binds DNA with double dT20 single-strand overhangs. (A) AUC sedimentation velocity of dT20 ssDNA labeled with Cy5 at 5' end alone (black) 1 mM and in the presence of 8 mM Ku (red). (B) AUC sedimentation velocity of 50 bp dsDNA internally labeled with Cy5 with two dT20 single-stranded tails on either end alone (black) 1 mM and in the presence of 8 µM Ku (red). The addition of Ku to blunt (50 bp internally labeled with Cy5 as control shown in green).

4.4.7 UvrD1 binding remodels a Ku-coated DNA structure.

Given the results of the previous sections, we hypothesised Ku-stimulated UvrD1-dependent unwinding was taking place in the context of Ku-coated DNA template. To test this idea, we used sedimentation velocity experiments to characterize the size and shape of Ku-DNA complexes in the presence and absence of UvrD1. The incubation of 1 mM of Cy5-labeled dT20-32bp DNA with 4 mM Ku results in a broad distribution of complexes with large $s_{20,w}$ values ranging between 7 and 20 (Figure 4.7A, red). Since the predicted $s_{20,w}$ value for a single Ku dimer bound to this DNA template is ~5.7, this is consistent with the binding of multiple Ku dimers as shown previously. Furthermore, $s_{20,w}$ values of the distribution further increase as the concentration of Ku is increased to 6 mM and 8 mM (Figure 4.7A, purple, and cyan).

To ask whether UvrD1 binds to these Ku-coated DNA templates, UvrD1 was added to the complexes generated by 4 mM, 6 mM, and 8 mM Ku. Upon addition of 2 mM UvrD1, we observed two main changes to the curves (Figure 4.7B-D, dashed curves). The first is the appearance of a peak around 6.5 which is consistent with the binding of UvrD1 to DNA templates uncoated with Ku(data not shown). The second is that the c(s) peaks at large values shift towards smaller $s_{20,w}$ values. This result, although difficult to interpret unequivocally, suggests that UvrD1 binds to the Ku-DNA complexes and remodels the Ku-DNA complex in a way that changes the size and/or shape of the complex. At a minimum, these data show that, at the concentrations needed for the stimulation of monomer unwinding, UvrD1 interacts with a Ku-coated DNA template.

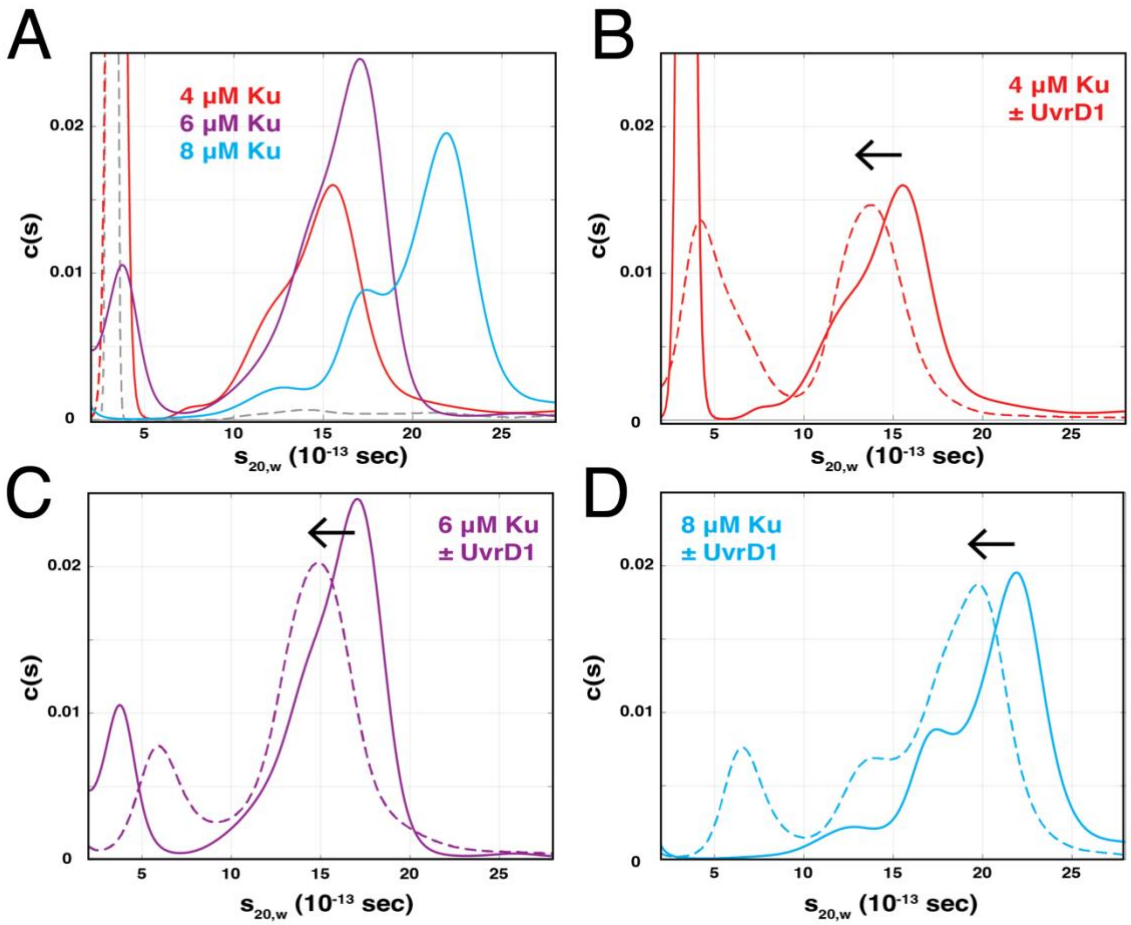


Figure 4.7: UvrD1 monomer binding remodels Ku coated DNA. (A) AUC sedimentation velocity plot of increasing concentration of Ku bound to 32bp-dT20 DNA substrate (1μM) labeled with Cy5. With increasing concentration of Ku the $s_{20,w}$ value increases suggesting multiple Ku bound to DNA. (B, C, and D) The addition of 2 mM UvrD1 shown in dashed curve remodels Ku bound nucleoprotein filament shown at 4, 6, and 8 uM concentration (compare dashed to solid curves in each figure).

4.4.8 The C-terminal Tudor domain of UvrD1 is required for the activation of unwinding.

Previous studies described an interaction between Ku and the C-terminal region of *Mtb* UvrD1 via yeast-two hybrid assays [38]. Further work using *M. smegmatis* UvrD1 showed that deleting the C-terminal 91 amino acids from UvrD1 prevented its interaction with Ku on native gels [40]. However, the same truncated construct maintained the ability to be activated by Ku

in gel-based unwinding assays. One caveat to the interpretation of these experiments is that they lacked the ability to separate dimer and monomer UvrD1 effects. Given our identification of the UvrD1 monomer as the activated oligomer by Ku under multiround conditions, we returned to these observations and asked whether the C-terminal region of UvrD1 was required for Ku-dependent helicase activation.

Based on alignments with other UvrD-family members including *M. smegmatis* UvrD1, we deleted the C-terminal 46 amino acids of *Mtb* UvrD1 which make up the conserved C-terminal Tudor domain (Supplementary Figure S4.5) [58]. We first characterized the effect of this UvrD1DTudor mutant on the interaction between UvrD1 and the DNA template itself. AUC with Cy5-labeled unwinding template revealed that both monomers and dimers of UvrD1DTudor bind DNA in about the same ratio as WT (Supplementary Figure S4.6). In fact, at the same concentrations, more monomers and dimers UvrD1DTudor were bound to DNA, suggestive of a higher affinity. Furthermore, UvrD1DTudor in the presence of DTT is monomeric similar to WT (Supplementary Figure S4.7) and shows increased DNA unwinding as compared to WTUvrD1 suggesting that the Tudor domain may be inhibitory under these conditions (Figure 4.8, cyan). However, unwinding assays performed with reduced UvrD1DTudor in the presence and absence of Ku revealed a lack of activation (Figure 4.8, purple and cyan) showing that Ku activation is Tudor domain dependent.

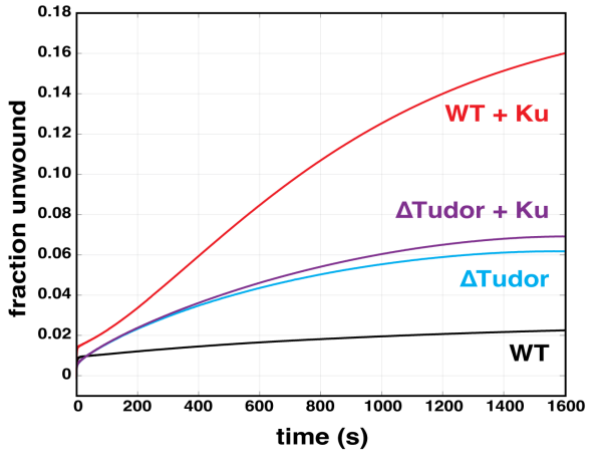


Figure 4.8: The C-terminal Tudor domain of UvrD1 is required for Ku-dependent activation of monomer unwinding. (A) The addition of Ku activates UvrD1 monomer unwinding in reducing conditions (red and black). UvrD1DTudor shows enhanced unwinding relative to WT (cyan and black) 200nM, but the addition of 400 nM Ku dimer does not have any effect (purple). Each trace shown was performed on 10 nM 32bp dT20 unwinding template.

4.5 Discussion

DNA helicases use energy from ATP hydrolysis to translocate on single-stranded DNA and unwind double-stranded DNA [31]. Extensive biochemical studies have shown that UvrD-like helicases unwind processively as dimers in absence of accessory factors, and we have previously shown this to be true for *Mtb* UvrD1 [37,59,60]. More specifically, we showed that dimerization of UvrD1 is linked to redox potential through the oxidative formation of a disulfide bond between 2B domains. In that work, we also showed that UvrD1 can be readily converted to monomers by the addition of a reducing agent or by single point mutation of cysteine to alanine at position 451 in the 2B domain, and proved that under single-turnover conditions, the monomeric form populated under reducing conditions did not possess helicase activity [37].

However, there are two fundamentally different ways to assay for helicase activity. Single-round experiments ask whether an enzyme (or enzyme complex) can productively unwind a template of a particular length in the context of a single binding event. To do so, the

enzyme must possess the ability to take processive steps along the DNA equal to the length of the duplex region without dissociating. If it does dissociate from the template, an excess of trap substrate prevents re-binding. In contrast, multiround experiments are performed in the absence of a trap and allow for dissociated enzymes to re-bind. Under these conditions, the enzyme does not need to possess a processivity on the order of the length of the template. Here, we show that although the monomeric form of UvrD1 does not display single-round unwinding even on the shortest templates, it does exhibit a small, but measurable, degree of unwinding under multiround conditions. This activity is approximately 100x slower than that exhibited by the dimer. So much slower in fact, that the two timescales of unwinding may be clearly observed in a single experiment when a mixture of dimers and monomers are present (Supplementary Figure S4.8).

Previous work by others shows that *Mtb* Ku, an accessory factor involved in NHEJ, stimulates the helicase activity of UvrD1 [38,40]. In other examples of UvrD-family enzyme activation, the effects of auxiliary factors could be seen on monomers, dimers, or both species [61–63]. Considering this, we asked which oligomeric species of UvrD1 was involved in the Ku-dependent mechanism. Since UvrD1 can be readily converted to monomers by the addition of a reducing agent or by single point mutation of cysteine to alanine at position 451 in the 2B domain, we could directly test whether Ku enhances the helicase activity of monomers, dimers, or both. We found that Ku did not affect unwinding by the processive dimer of UvrD1 but increased unwinding by UvrD1 monomers about five-fold. This stimulation of unwinding is only observed on a timescale of 10 minutes in multiround conditions or in conditions when ssDNA is used as a TRAP rather than blunt or partial duplex DNA substrates. Since Ku alone cannot bind to single-stranded DNA these results suggest the presence of a Ku-UvrD1 complex that unwinds via multiple binding events. This model is consistent with previous studies showing Ku-stimulated unwinding on longer time scales using single-stranded DNA traps [38].

Although published studies of Ku-stimulated unwinding by UvrD1 report 80-90% DNA unwound, we observed a maximum of only 15% DNA unwound. This discrepancy may be due to the absence of a reducing agent in the previous work that would lead to unwinding by both dimers and monomers. In addition, the solution conditions in which unwinding was performed were different from our current study. We performed experiments in TRIS pH 8.0, 20% glycerol, and 75 mM NaCl at 25 °C to be consistent with our first study. In contrast, the other study was performed at 37 °C and with no reported monovalent salt concentration although there is 150 mM NaCl in the protein storage buffers [38]. When we performed a comparison across temperatures, we saw an acceleration of the unwinding kinetics at 37 °C compared to 25 °C, but the fraction unwound at saturation on a dT20-18 bp template was comparable (Supplementary Figure S4.9). As changes in buffer conditions dramatically affect the properties of UvrD helicases, the difference in fraction unwound is surely due to differences in temperature, salt concentration, and redox potential between the different studies [59].

All studies we have found investigating the binding of Ku proteins to DNA were performed in presence of Mg^{+2} [50,52]. In the process of our studies presented here, we unexpectedly showed that the interaction of Ku and DNA is completely dependent on the presence of magnesium. We have not explored the concentration dependence or stoichiometry of this necessity, nor have we investigated the ability of other divalent cations to substitute for magnesium. These aspects of the Ku-DNA interaction would be interesting for future studies.

Our AUC sedimentation velocity data show that *Mtb* Ku exists as a dimer in solution and can bind to blunt ended dsDNA, consistent with previous studies [64,65]. We also showed that a single-strand overhang does not prevent the loading of *Mtb* Ku on dsDNA and that multiple Ku dimers cooperatively coat our DNA templates. The addition of UvrD1 monomer to Ku-bound DNA remodels the complex and appears to dissociate some fraction of Ku molecules. Taken together, our data suggest a model where activation occurs in the context of

a UvrD1 monomer bound to the single-stranded side of the junction interacting with a Ku dimer bound to the double-stranded side of the junction. Since activation increases as more Ku dimers are loaded onto the template, we hypothesize that the loading of additional Kus serve to increase the probability of a junction proximal Ku by restricting the available space on the duplex region of the template.

In addition, on templates with a 18bp-dT20 tail, this multiround activity of Ku-activated monomers leads to the unwinding of only about 5% of the available templates over the course of tens of minutes. The lengthening of the single-stranded tail of the template leads to more efficient unwinding over similar timescales (Supplementary Figure S4.10). This observation suggests that either the rate of binding of UvrD1 monomer to the templates limits the yield of the reaction and/or that the ability to bind multiple monomers simultaneously on a template increases the efficiency of the reaction. Future experiments would have to be performed to investigate this effect further.

The specific DNA repair pathways in which the interaction between Ku and UvrD1 plays a role is not clear. Double-stranded breaks are processed either by resection-independent or resection-dependent mechanisms in mycobacteria [10]. The canonical NHEJ pathway utilizes Ku-LigD complexes to seal breaks without single-stranded resection. In addition to ligase activity, the mycobacterial version of LigD has both nuclease and polymerase activities which aid in the trimming of overhanging flaps and the filling in of nucleotides prior to ligation [15]. The resection-dependent pathways require the production of single-stranded overhanging regions to be used either for homology searches (i.e., homologous recombination or gene conversion) or for single-stranded annealing (SSA) [9,66]. These two pathways depend on the combined helicase/nuclease activities of AdnAB and RecBCD respectively and are independent of Ku [67]. In these contexts, one possible hypothesis based on our data is that UvrD1 controls the timing of and partitioning between these different strategies. Its ability to

interact with Ku may inhibit NHEJ, while multiple rounds of binding and unwinding may prepare the substrate for resection by other enzymes. This would have the effect of funneling the repair away from non-homologous end joining and into homology-based mechanisms. The concentration and oligomeric status of UvrD1 would then provide a means for the regulation of fluxes in each pathway.

Dissecting the physiological role of UvrD1 should be aided by our results. Specifically, one could dissect the roles of the dimer and Ku-UvrD1 monomer complex *in vivo* by replacing WT UvrD1 with either the 2B cysteine mutant which is an obligate monomer or with the UvrD1DTudor construct which can still form dimers but can no longer interact with Ku. In particular, one could test the hypothesis that the UvrD1 dimer is active in NER while the Ku-UvrD1 monomer complex participates in alternative pathways of repair including DSBR.

SF1 family helicase monomers have previously been shown to be activated through protein-protein interactions with accessory factors. For instance, *E. coli* UvrD monomer and dimer helicase activities are both stimulated by MutL, and *E. coli* Rep monomer and dimer can be activated by PriC [61–63]. In contrast, we distinctly did not observe a Ku-dependent increase in unwinding by the UvrD1 dimer. We note that perhaps the presence of the covalent dimer, absent in the case of *E. coli* UvrD and Rep, contributes to this difference. Additionally, it may be that the activation of Ku is restricted to longer timescales in multiround conditions and that, due to the relative briskness of dimer activity, it is not affected.

The activation of *E. coli* Rep unwinding by PriC depends on the C-terminal domain and previous yeast-two-hybrid results suggest that the CTD of *Mtb* UvrD1 interacts with the N-terminus of Ku [38]. Previous studies have shown that removing the entire C-terminal domain of *M. smegmatis* UvrD1 including the disordered linker and the Tudor domain leads to abrogation of super shifted Ku-UvrD1-DNA complex in the native gel assay, but Ku can still

stimulate the unwinding of a similar construct of *Mtb* UvrD1 [39]. We removed the C-terminal Tudor domain (46 amino acids) of *Mtb* UvrD1 to test directly if the Tudor domain of *Mtb* UvrD1 played a role in the Ku-dependent activation. We found that deleting the Tudor domain increases the amplitude of the slow phase of multiround unwinding by UvrD1 both in the absence and presence of DTT (Supplementary Figure S4.11), suggesting that the Tudor domain itself is autoinhibitory for DNA unwinding. The observation is reminiscent of the finding that removal of the autoinhibitory 2B domain leads to activation of monomeric Rep helicase activity [68]. Furthermore, although basal levels of slow timescale unwinding were increased, the DTudor construct was no longer activated by Ku (Figure 4.6). This observation is consistent with the physical interaction of the Tudor domain with Ku and supports a hypothesis that this interaction releases the autoinhibitory effect of the Tudor domain on DNA unwinding.

This potential inhibitory role of the Tudor domain is particularly interesting given that the same domain has been shown to underlie the interaction of *E. coli* UvrD and *B. subtilis* PcrA with the RNA polymerase beta-subunit during transcription-coupled NER (TCR) [35,69,70]. In TCR, lesions on the template strand cause RNAP pausing during elongation. This paused state becomes a substrate for either UvrD or Mfd, which then process the complex to both remove the polymerase from the lesion by either backtracking or forward tracking respectively and recruit downstream NER factors [71–73]. In the case of UvrD, it has been suggested that a dimer of UvrD is needed to interact with RNAP to produce backtracking [71,74]. However, the results presented here suggest an alternative mechanism where a monomer of UvrD would be activated via the binding of the Tudor domain with RNAP. The Tudor domain results here are likely extendible to homologous systems as we observed that *Mtb* Ku activates the Tudor domain containing *E. coli* UvrD, but not the Tudor domain lacking *E. coli* Rep under these conditions (Supplementary Figure S4.12). This possibility would also suggest that the Tudor domain may serve as a protein-protein interaction hub for UvrD1 and

that the partitioning of UvrD1 into various repair pathways would ultimately be controlled via competition for this domain.

Author Contributions

A.C. and E.A.G. designed research; A.C. performed experiments and analyzed the data. A.K. helped with the analysis of quenching titrations and binding density analyses. A.C. and E.A.G. wrote the paper. T.M.L. and E.A.G. guided the project.

Acknowledgements

We thank Rahul Chadda from Edwin Antony's lab for their help in using their fluorimeter instrument for preliminary anisotropy assays. We thank Galburt lab members for their comments on the manuscript.

Funding

EAG is supported by grant R35GM144282 from NIGMS.

Conflict of Interest

The authors declare no conflict of interest related to this work.

4.6 References

1. S. Bagcchi, WHO's Global Tuberculosis Report 2022, Lancet Microbe. 4 (2022) e20. https://doi.org/10.1016/s2666-5247(22)00359-7.

2. H.T. Pacl, V.P. Reddy, V. Saini, K.C. Chinta, A.J.C. Steyn, Host-pathogen redox dynamics modulate Mycobacterium tuberculosis pathogenesis, Pathog Dis. 76 (2018) fty036. https://doi.org/10.1093/femspd/fty036.

3. M.I. Voskuil, I.L. Bartek, K. Visconti, G.K. Schoolnik, The Response of Mycobacterium Tuberculosis to Reactive Oxygen and Nitrogen Species, Front Microbiol. 2 (2011) 105. https://doi.org/10.3389/fmicb.2011.00105.

4. N. Chatterjee, G.C. Walker, Mechanisms of DNA damage, repair, and mutagenesis, Environ Mol Mutagen. 58 (2017) 235–263. https://doi.org/10.1002/em.22087.

5. T.D. Vultos, O. Mestre, T. Tonjum, B. Gicquel, DNA repair in Mycobacterium tuberculosis revisited, Fems Microbiol Rev. 33 (2009) 471–487. https://doi.org/10.1111/j.1574-6976.2009.00170.x.

6. M.R. Lieber, The Mechanism of Human Nonhomologous DNA End Joining*, J Biol Chem. 283 (2008) 1–5. https://doi.org/10.1074/jbc.r700039200.

7. A.C. Vítor, P. Huertas, G. Legube, S.F. de Almeida, Studying DNA Double-Strand Break Repair: An Ever-Growing Toolbox, Frontiers Mol Biosci. 7 (2020) 24. https://doi.org/10.3389/fmolb.2020.00024.

8. M.S. Glickman, Double-Strand DNA Break Repair in Mycobacteria, Microbiol Spectr. 2 (2014). https://doi.org/10.1128/microbiolspec.mgm2-0024-2013.

9. R. Gupta, D. Barkan, G. Redelman-Sidi, S. Shuman, M.S. Glickman, Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways, Mol Microbiol. 79 (2011) 316–330. https://doi.org/10.1111/j.1365-2958.2010.07463.x.

10. R. Gupta, M. Ryzhikov, O. Koroleva, M. Unciuleac, S. Shuman, S. Korolev, M.S. Glickman, A dual role for mycobacterial RecO in RecA-dependent homologous recombination and RecA-independent single-strand annealing, Nucleic Acids Res. 41 (2013) 2284–2295. https://doi.org/10.1093/nar/gks1298.

11. M. Vos, X. Didelot, A comparison of homologous recombination rates in bacteria and archaea, Isme J. 3 (2009) 199–208. https://doi.org/10.1038/ismej.2008.93.

12. R.S. Pitcher, L.M. Tonkin, J.M. Daley, P.L. Palmbos, A.J. Green, T.L. Velting, A. Brzostek, M. Korycka-Machala, S. Cresawn, J. Dziadek, G.F. Hatfull, T.E. Wilson, A.J. Doherty, Mycobacteriophage Exploit NHEJ to Facilitate Genome Circularization, Mol Cell. 23 (2006) 743–748. https://doi.org/10.1016/j.molcel.2006.07.009.

13. J. Aniukwu, M.S. Glickman, S. Shuman, The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends, Gene Dev. 22 (2008) 512–527. https://doi.org/10.1101/gad.1631908.

14. A. Brzostek, I. Szulc, M. Klink, M. Brzezinska, Z. Sulowska, J. Dziadek, Either Non-Homologous Ends Joining or Homologous Recombination Is Required to Repair Double-Strand Breaks in the Genome of Macrophage-Internalized Mycobacterium tuberculosis, Plos One. 9 (2014) e92799. https://doi.org/10.1371/journal.pone.0092799.

15. M. Della, P.L. Palmbos, H.-M. Tseng, L.M. Tonkin, J.M. Daley, L.M. Topper, R.S. Pitcher, A.E. Tomkinson, T.E. Wilson, A.J. Doherty, Mycobacterial Ku and Ligase Proteins Constitute a Two-Component NHEJ Repair Machine, Science. 306 (2004) 683–685. https://doi.org/10.1126/science.1099824.

16. C. Gong, P. Bongiorno, A. Martins, N.C. Stephanou, H. Zhu, S. Shuman, M.S. Glickman, Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system

driven by Ku, ligase D and ligase C, Nat Struct Mol Biol. 12 (2005) 304–312. https://doi.org/10.1038/nsmb915.

17. S. Zahid, M.S.E. Dahan, F. Iehl, P. Fernandez-Varela, M.-H.L. Du, V. Ropars, J.B. Charbonnier, The Multifaceted Roles of Ku70/80, Int J Mol Sci. 22 (2021) 4134. https://doi.org/10.3390/ijms22084134.

18. J.R. Walker, R.A. Corpina, J. Goldberg, Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair, Nature. 412 (2001) 607–614. https://doi.org/10.1038/35088000.

19. A.J. Doherty, S.P. Jackson, G.R. Weller, Identification of bacterial homologues of the Ku DNA repair proteins, Febs Lett. 500 (2001) 186–188. https://doi.org/10.1016/s0014-5793(01)02589-3.

20. S. McGovern, S. Baconnais, P. Roblin, P. Nicolas, P. Drevet, H. Simonson, O. Piétrement, J.-B. Charbonnier, E.L. Cam, P. Noirot, F. Lecointe, C-terminal region of bacterial Ku controls DNA bridging, DNA threading and recruitment of DNA ligase D for double strand breaks repair, Nucleic Acids Res. 44 (2016) 4785–4806. https://doi.org/10.1093/nar/gkw149.

21. S.W. Matson, D.W. Bean, J.W. George, DNA helicases: Enzymes with essential roles in all aspects of DNA metabolism, Bioessays. 16 (1994) 13–22. https://doi.org/10.1002/bies.950160103.

22. A.J. van Brabant, R. Stan, N.A. Ellis, DNA HELICASES, GENOMIC INSTABILITY, AND HUMAN GENETIC DISEASE, Genom Hum Genetics. 1 (2000) 409–459. https://doi.org/10.1146/annurev.genom.1.1.409.

23. N.K. Maluf, C.J. Fischer, T.M. Lohman, A Dimer of Escherichia coli UvrD is the Active Form of the Helicase In Vitro, J Mol Biol. 325 (2003) 913–935. https://doi.org/10.1016/s0022-2836(02)01277-9.

24. J. Houghton, C. Townsend, A.R. Williams, A. Rodgers, L. Rand, K.B. Walker, E.C. Böttger, B. Springer, E.O. Davis, Important Role for Mycobacterium tuberculosis UvrD1 in Pathogenesis and Persistence apart from Its Function in Nucleotide Excision Repair, J Bacteriol. 194 (2012) 2916–2923. https://doi.org/10.1128/jb.06654-11.

25. K.M. Sinha, M.S. Glickman, S. Shuman, Mutational Analysis of Mycobacterium UvrD1 Identifies Functional Groups Required for ATP Hydrolysis, DNA Unwinding, and Chemomechanical Coupling, Biochemistry-Us. 48 (2009) 4019–4030. https://doi.org/10.1021/bi900103d.

26. R.M. Brosh, DNA helicases involved in DNA repair and their roles in cancer, Nat Rev Cancer. 13 (2013) 542–558. https://doi.org/10.1038/nrc3560.

27. J.T. Reardon, A. Sancar, Nucleotide Excision Repair, Prog Nucleic Acid Re. 79 (2005) 183–235. https://doi.org/10.1016/s0079-6603(04)79004-2.

28. V. Epshtein, V. Kamarthapu, K. McGary, V. Svetlov, B. Ueberheide, S. Proshkin, A. Mironov, E. Nudler, UvrD facilitates DNA repair by pulling RNA polymerase backwards, Nature. 505 (2014) 372. https://doi.org/10.1038/nature12928.

29. G. Kokic, A. Chernev, D. Tegunov, C. Dienemann, H. Urlaub, P. Cramer, Structural basis of TFIIH activation for nucleotide excision repair, Nat Commun. 10 (2019) 2885. https://doi.org/10.1038/s41467-019-10745-5.

30. M.E. Fairman-Williams, U.-P. Guenther, E. Jankowsky, SF1 and SF2 helicases: family matters, Curr Opin Struc Biol. 20 (2010) 313–324. https://doi.org/10.1016/j.sbi.2010.03.011.

31. K.D. Raney, A.K. Byrd, S. Aarattuthodiyil, DNA Helicases and DNA Motor Proteins, Adv Exp Med Biol. 767 (2012) 17–46. https://doi.org/10.1007/978-1-4614-5037-5_2.

32. T.M. Lohman, E.J. Tomko, C.G. Wu, Non-hexameric DNA helicases and translocases: mechanisms and regulation, Nat Rev Mol Cell Bio. 9 (2008) 391–401. https://doi.org/10.1038/nrm2394.

33. M.S. Dillingham, D.B. Wigley, M.R. Webb, Direct Measurement of Single-Stranded DNA Translocation by PcrA Helicase Using the Fluorescent Base Analogue 2-Aminopurine †, Biochemistry-Us. 41 (2002) 643–651. https://doi.org/10.1021/bi011137k.

34. M.S. Dillingham, D.B. Wigley, M.R. Webb, Demonstration of Unidirectional Single-Stranded DNA Translocation by PcrA Helicase: Measurement of Step Size and Translocation Speed †, Biochemistry-Us. 39 (2000) 205–212. https://doi.org/10.1021/bi992105o.

35. K. Sanders, C.-L. Lin, A.J. Smith, N. Cronin, G. Fisher, V. Eftychidis, P. McGlynn, N.J. Savery, D.B. Wigley, M.S. Dillingham, The structure and function of an RNA polymerase interaction domain in the PcrA/UvrD helicase, Nucleic Acids Res. 45 (2017) gkx074-. https://doi.org/10.1093/nar/gkx074.

36. E.J. Gwynn, A.J. Smith, C.P. Guy, N.J. Savery, P. McGlynn, M.S. Dillingham, The Conserved C-Terminus of the PcrA/UvrD Helicase Interacts Directly with RNA Polymerase, Plos One. 8 (2013) e78141. https://doi.org/10.1371/journal.pone.0078141.

37. A. Chadda, D. Jensen, E.J. Tomko, A.R. Manzano, B. Nguyen, T.M. Lohman, E.A. Galburt, Mycobacterium tuberculosis DNA repair helicase UvrD1 is activated by redox-dependent dimerization via a 2B domain cysteine, Proc National Acad Sci. 119 (2022) e2114501119. https://doi.org/10.1073/pnas.2114501119.

38. K.M. Sinha, N.C. Stephanou, F. Gao, M.S. Glickman, S. Shuman, Mycobacterial UvrD1 Is a Ku-dependent DNA Helicase That Plays a Role in Multiple DNA Repair Events, Including Double-strand Break Repair*, J Biol Chem. 282 (2007) 15114–15125. https://doi.org/10.1074/jbc.m701167200.

39. K.M. Sinha, M.S. Glickman, S. Shuman, Mutational Analysis of Mycobacterium UvrD1 Identifies Functional Groups Required for ATP Hydrolysis, DNA Unwinding, and Chemomechanical Coupling, Biochemistry-Us. 48 (2009) 4019–4030. https://doi.org/10.1021/bi900103d.

40. K.M. Sinha, N.C. Stephanou, M.-C. Unciuleac, M.S. Glickman, S. Shuman, Domain Requirements for DNA Unwinding by Mycobacterial UvrD2, an Essential DNA Helicase, Biochemistry-Us. 47 (2008) 9355–9364. https://doi.org/10.1021/bi800725q.

41. W. Bujalowski, Thermodynamic and Kinetic Methods of Analyses of Protein−Nucleic Acid Interactions. From Simpler to More Complex Systems, Chem Rev. 106 (2006) 556–606. https://doi.org/10.1021/cr040462l.

42. A.G. Kozlov, R. Galletto, T.M. Lohman, Single-Stranded DNA Binding Proteins, Methods and Protocols, Methods Mol Biology. 922 (2012) 55–83. https://doi.org/10.1007/978-1-62703-032-8_4.

43. J. Dam, P. Schuck, Calculating Sedimentation Coefficient Distributions by Direct Modeling of Sedimentation Velocity Concentration Profiles, Methods Enzymol. 384 (2004) 185–212. https://doi.org/10.1016/s0076-6879(04)84012-6.

44. P. Schuck, Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling, Biophys J. 78 (2000) 1606–1619. https://doi.org/10.1016/s0006-3495(00)76713-0.

45. T.M. Laue, B.D. Shah, T.M. Ridgeway, S.L. Pelletier, Computer-aided interpretation of analytical sedimentation data for proteins, in: S. Harding, A. Rowe, J. Horton (Eds.), Analytical Ultracentrifugation in Biochemistry and Polymer Science, Royal Society of Chemistry, 1992: pp. 90–125.

46. R. Öz, J.L. Wang, R. Guerois, G. Goyal, S. KK, V. Ropars, R. Sharma, F. Koca, J.-B. Charbonnier, M. Modesti, T.R. Strick, F. Westerlund, Dynamics of Ku and bacterial non-homologous end-joining characterized using single DNA molecule analysis, Nucleic Acids Res. 49 (2021) gkab083-. https://doi.org/10.1093/nar/gkab083.

47. D. Arosio, S. Costantini, Y. Kong, A. Vindigni, Fluorescence Anisotropy Studies on the Ku-DNA Interaction ANION AND CATION EFFECTS*, J Biol Chem. 279 (2004) 42826–42835. https://doi.org/10.1074/jbc.m406529200.

48. B. Kysela, A.J. Doherty, M. Chovanec, T. Stiff, S.M. Ameer-Beg, B. Vojnovic, P.-M. Girard, P.A. Jeggo, Ku Stimulation of DNA Ligase IV-dependent Ligation Requires Inward Movement along the DNA Molecule*, J Biol Chem. 278 (2003) 22466–22474. https://doi.org/10.1074/jbc.m303273200.

49. R. Öz, J.L. Wang, R. Guerois, G. Goyal, S. KK, V. Ropars, R. Sharma, F. Koca, J.-B. Charbonnier, M. Modesti, T.R. Strick, F. Westerlund, Dynamics of Ku and bacterial non-homologous end-joining characterized using single DNA molecule analysis, Nucleic Acids Res. 49 (2021) gkab083-. https://doi.org/10.1093/nar/gkab083.

50. Y. Ma, M.R. Lieber, DNA Length-Dependent Cooperative Interactions in the Binding of Ku to DNA †, Biochemistry-Us. 40 (2001) 9638–9646. https://doi.org/10.1021/bi010932v.

51. A.K. Kushwaha, A. Grove, C-terminal low-complexity sequence repeats of Mycobacterium smegmatis Ku modulate DNA binding, Bioscience Rep. 33 (2013) e00016. https://doi.org/10.1042/bsr20120105.

52. S. McGovern, S. Baconnais, P. Roblin, P. Nicolas, P. Drevet, H. Simonson, O. Piétrement, J.-B. Charbonnier, E.L. Cam, P. Noirot, F. Lecointe, C-terminal region of bacterial Ku controls DNA bridging, DNA threading and recruitment of DNA ligase D for double strand breaks repair, Nucleic Acids Res. 44 (2016) 4785–4806. https://doi.org/10.1093/nar/gkw149.

53. G.R. Weller, B. Kysela, R. Roy, L.M. Tonkin, E. Scanlan, M. Della, S.K. Devine, J.P. Day, A. Wilkinson, F. d'Adda di Fagagna, K.M. Devine, R.P. Bowater, P.A. Jeggo, S.P. Jackson, A.J. Doherty, Identification of a DNA Nonhomologous End-Joining Complex in Bacteria, Science. 297 (2002) 1686–1689. https://doi.org/10.1126/science.1074584.

54. B. Kysela, A.J. Doherty, M. Chovanec, T. Stiff, S.M. Ameer-Beg, B. Vojnovic, P.-M. Girard, P.A. Jeggo, Ku Stimulation of DNA Ligase IV-dependent Ligation Requires Inward Movement along the DNA Molecule*, J Biol Chem. 278 (2003) 22466–22474. https://doi.org/10.1074/jbc.m303273200.

55. D.J. Sowa, M.M. Warner, A. Tetenych, L. Koechlin, P. Balari, J.P. Rascon Perez, C. Caba, S.N. Andres, The Mycobacterium tuberculosis Ku C-terminus is a multi-purpose arm for binding DNA and LigD and stimulating ligation, Nucleic Acids Res. (2022). https://doi.org/10.1093/nar/gkac906.

56. Y. Ma, M.R. Lieber, DNA Length-Dependent Cooperative Interactions in the Binding of Ku to DNA †, Biochemistry-Us. 40 (2001) 9638–9646. https://doi.org/10.1021/bi010932v.

57. R. Öz, J.L. Wang, R. Guerois, G. Goyal, S. KK, V. Ropars, R. Sharma, F. Koca, J.-B. Charbonnier, M. Modesti, T.R. Strick, F. Westerlund, Dynamics of Ku and bacterial non-homologous end-joining characterized using single DNA molecule analysis, Nucleic Acids Res. 49 (2021) gkab083-. https://doi.org/10.1093/nar/gkab083.

58. K. Sanders, C.-L. Lin, A.J. Smith, N. Cronin, G. Fisher, V. Eftychidis, P. McGlynn, N.J. Savery, D.B. Wigley, M.S. Dillingham, The structure and function of an RNA polymerase interaction domain in the PcrA/UvrD helicase, Nucleic Acids Res. 45 (2017) gkx074-. https://doi.org/10.1093/nar/gkx074.

59. N.K. Maluf, T.M. Lohman, Self-association Equilibria of Escherichia coli UvrD Helicase Studied by Analytical Ultracentrifugation, J Mol Biol. 325 (2003) 889–912. https://doi.org/10.1016/s0022-2836(02)01276-7.

60. N.K. Maluf, T.M. Lohman, Self-association Equilibria of Escherichia coli UvrD Helicase Studied by Analytical Ultracentrifugation, J Mol Biol. 325 (2003) 889–912. https://doi.org/10.1016/s0022-2836(02)01276-7.

61. B. Nguyen, M.K. Shinn, E. Weiland, T.M. Lohman, Regulation of E. coli Rep helicase activity by PriC, J Mol Biol. 433 (2021) 167072. https://doi.org/10.1016/j.jmb.2021.167072.

62. Y.A. Ordabayev, B. Nguyen, A. Niedziela-Majka, T.M. Lohman, Regulation of UvrD Helicase Activity by MutL, J Mol Biol. 430 (2018) 4260–4274. https://doi.org/10.1016/j.jmb.2018.08.022.

63. Y.A. Ordabayev, B. Nguyen, A.G. Kozlov, H. Jia, T.M. Lohman, UvrD helicase activation by MutL involves rotation of its 2B subdomain, Proc National Acad Sci. 116 (2019) 16320–16325. https://doi.org/10.1073/pnas.1905513116.

64. M. Koike, T. Shiomi, A. Koike, Dimerization and Nuclear Localization of Ku Proteins*, J Biol Chem. 276 (2001) 11167–11173. https://doi.org/10.1074/jbc.m010902200.

65. R. Bowater, A.J. Doherty, Making Ends Meet: Repairing Breaks in Bacterial DNA by Non-Homologous End-Joining, Plos Genet. 2 (2006) e8. https://doi.org/10.1371/journal.pgen.0020008.

66. R. Gupta, M. Ryzhikov, O. Koroleva, M. Unciuleac, S. Shuman, S. Korolev, M.S. Glickman, A dual role for mycobacterial RecO in RecA-dependent homologous recombination and RecA-independent single-strand annealing, Nucleic Acids Res. 41 (2013) 2284–2295. https://doi.org/10.1093/nar/gks1298.

67. K.M. Sinha, M.-C. Unciuleac, M.S. Glickman, S. Shuman, AdnAB: a new DSB-resecting motor–nuclease from mycobacteria, Gene Dev. 23 (2009) 1423–1437. https://doi.org/10.1101/gad.1805709.

68. K.M. Brendza, W. Cheng, C.J. Fischer, M.A. Chesnik, A. Niedziela-Majka, T.M. Lohman, Autoinhibition of Escherichia coli Rep monomer helicase activity by its 2B subdomain, Proc National Acad Sci. 102 (2005) 10076–10081. https://doi.org/10.1073/pnas.0502886102.

69. A.A. Kawale, B.M. Burmann, UvrD helicase–RNA polymerase interactions are governed by UvrD's carboxy-terminal Tudor domain, Commun Biology. 3 (2020) 607. https://doi.org/10.1038/s42003-020-01332-2.

70. I. Urrutia-Irazabal, J.R. Ault, F. Sobott, N.J. Savery, M.S. Dillingham, Analysis of the PcrA-RNA polymerase complex reveals a helicase interaction motif and a role for PcrA/UvrD helicase in the suppression of R-loops, Elife. 10 (2021) e68829. https://doi.org/10.7554/elife.68829.

71. V. Epshtein, V. Kamarthapu, K. McGary, V. Svetlov, B. Ueberheide, S. Proshkin, A. Mironov, E. Nudler, UvrD facilitates DNA repair by pulling RNA polymerase backwards, Nature. 505 (2014) 372. https://doi.org/10.1038/nature12928.

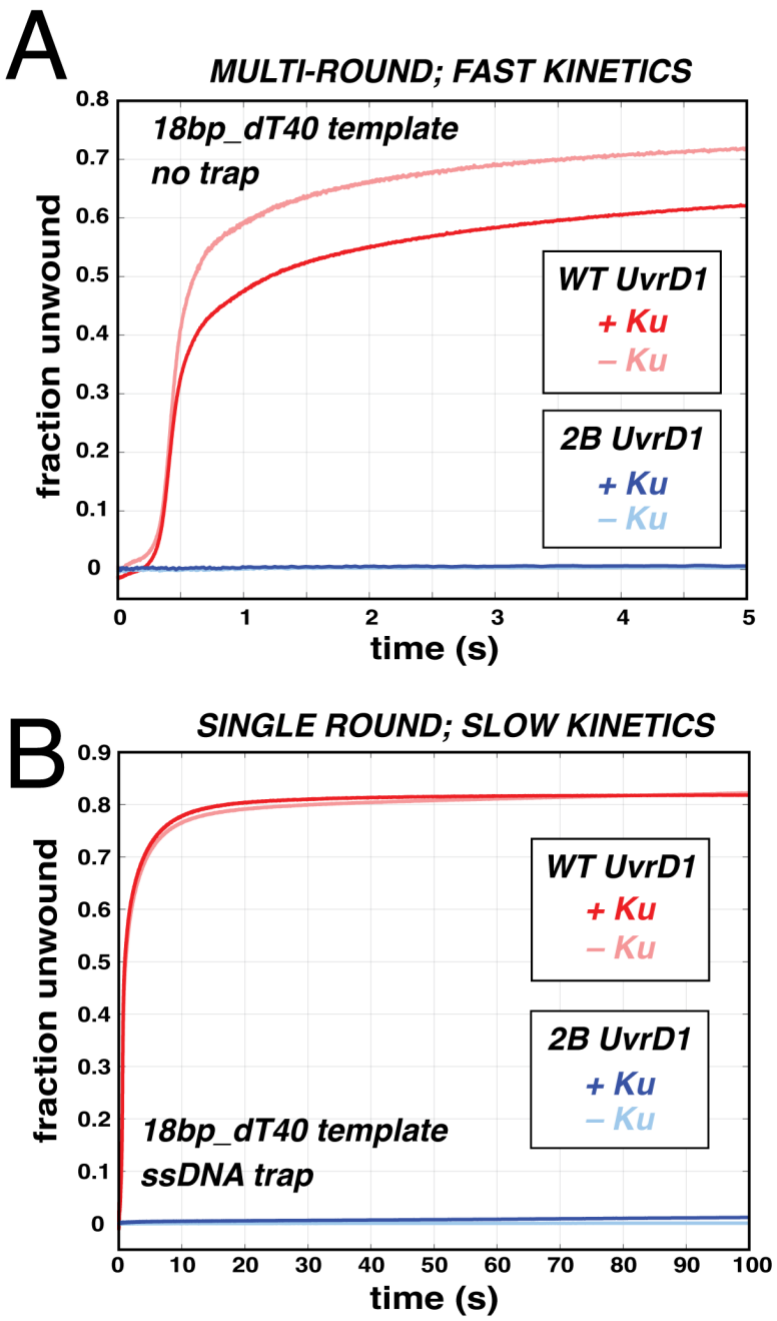
72. J.-S. Park, M.T. Marr, J.W. Roberts, E. coli Transcription Repair Coupling Factor (Mfd Protein) Rescues Arrested Complexes by Promoting Forward Translocation, Cell. 109 (2002) 757–767. https://doi.org/10.1016/s0092-8674(02)00769-9.

73. J.Y. Kang, E. Llewellyn, J. Chen, P.D.B. Olinares, J. Brewer, B.T. Chait, E.A. Campbell, S.A. Darst, Structural basis for transcription complex disruption by the Mfd translocase, Elife. 10 (2021) e62117. https://doi.org/10.7554/elife.62117.

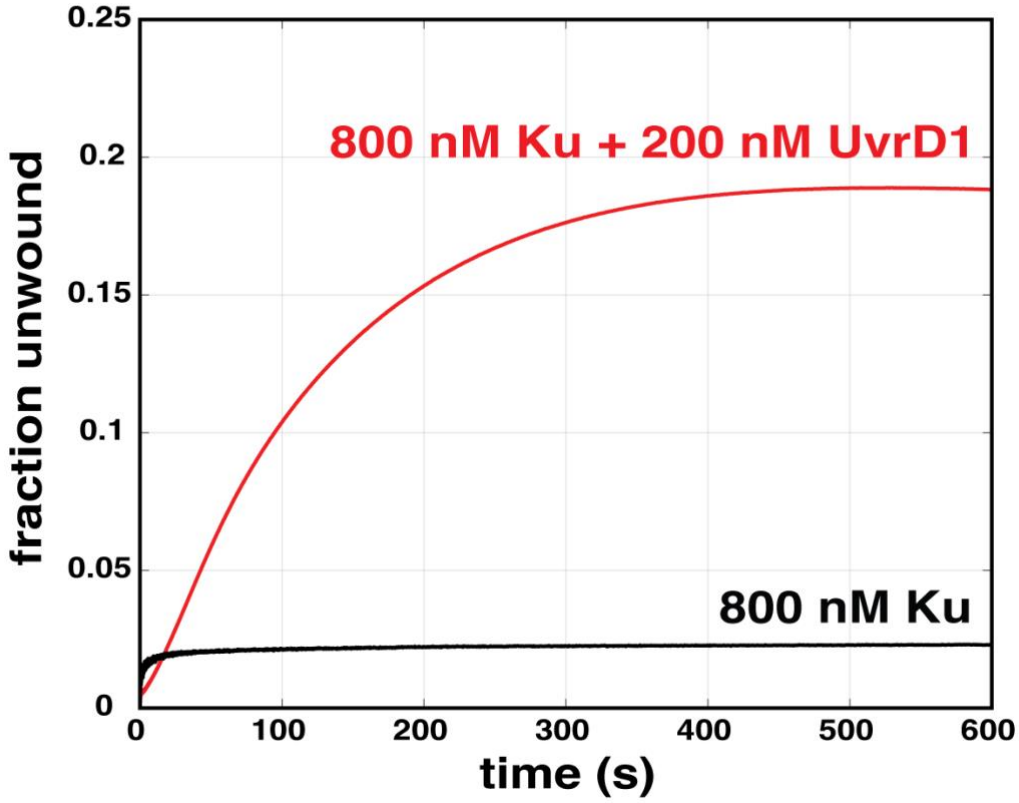
74. B.K. Bharati, M. Gowder, F. Zheng, K. Alzoubi, V. Svetlov, V. Kamarthapu, J.W. Weaver, V. Epshtein, N. Vasilyev, L. Shen, Y. Zhang, E. Nudler, Crucial role and mechanism of transcription-coupled DNA repair in bacteria, Nature. (2022) 1–8. https://doi.org/10.1038/s41586-022-04530-6.

Supplemental information

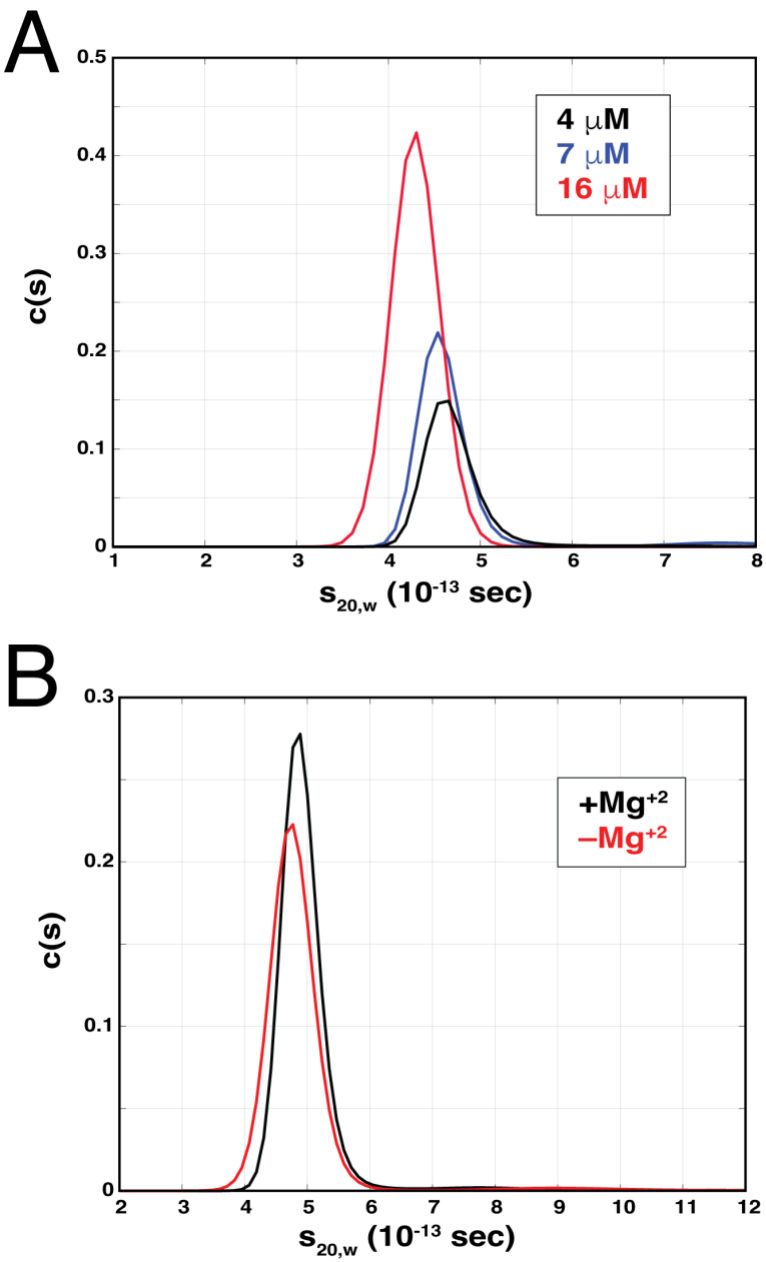
Chapter 4: *Mycobacterium tuberculosis* Ku Stimulates Multiround DNA Unwinding by UvrD1 Monomers.



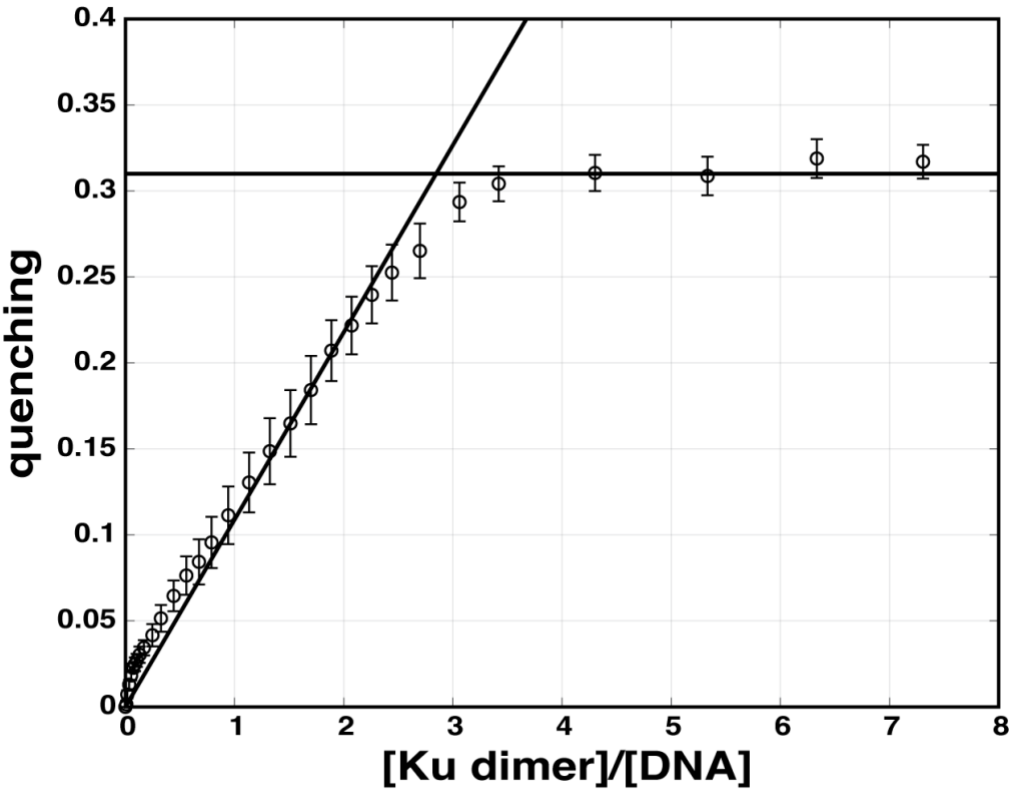
Supplementary Figure 4.1: DNA unwinding by UvrD1 dimers is not stimulated by Ku on short time scales in multi-round conditions nor in single-round conditions at longer timescales. (A) Multi-round DNA unwinding by WT UvrD1 in the absence of DTT where it's a mixture of monomers and dimers in the presence and absence of Ku (red traces). Multi-round DNA unwinding by 2B UvrD1 where removal of cysteine 451 leads to abrogation of dimerization in the presence and absence of Ku (blue traces). (B) Ku did not affect single-round unwinding by WT UvrD1 in the absence of DTT at longer timescales (red traces). Nor did Ku affect single-round unwinding by 2B UvrD1 under these conditions (blue traces).



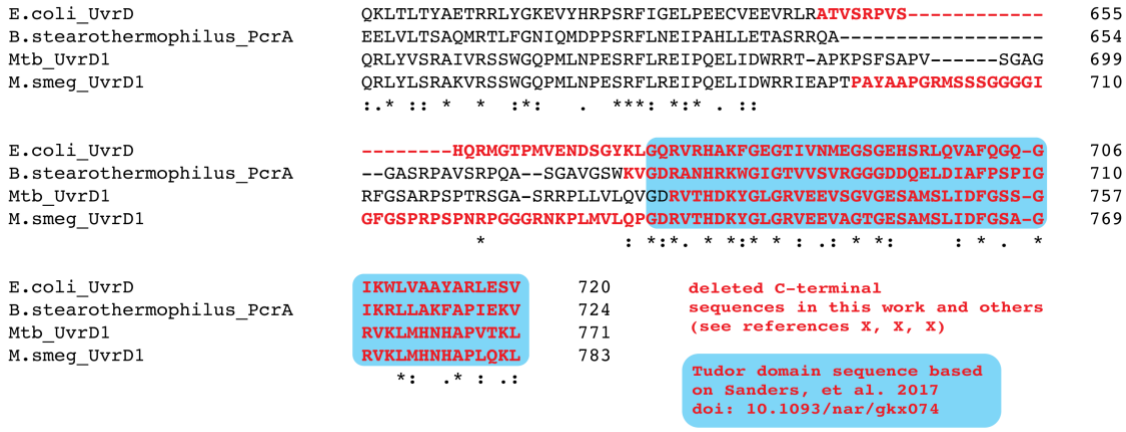
Supplementary Figure 4.2: Ku alone does not account for the DNA unwinding observed in the presence of UvrD1. DNA unwinding traces are shown with 400 nM Ku in the presence (red) and absence (black) of 200 nM UvrD1.



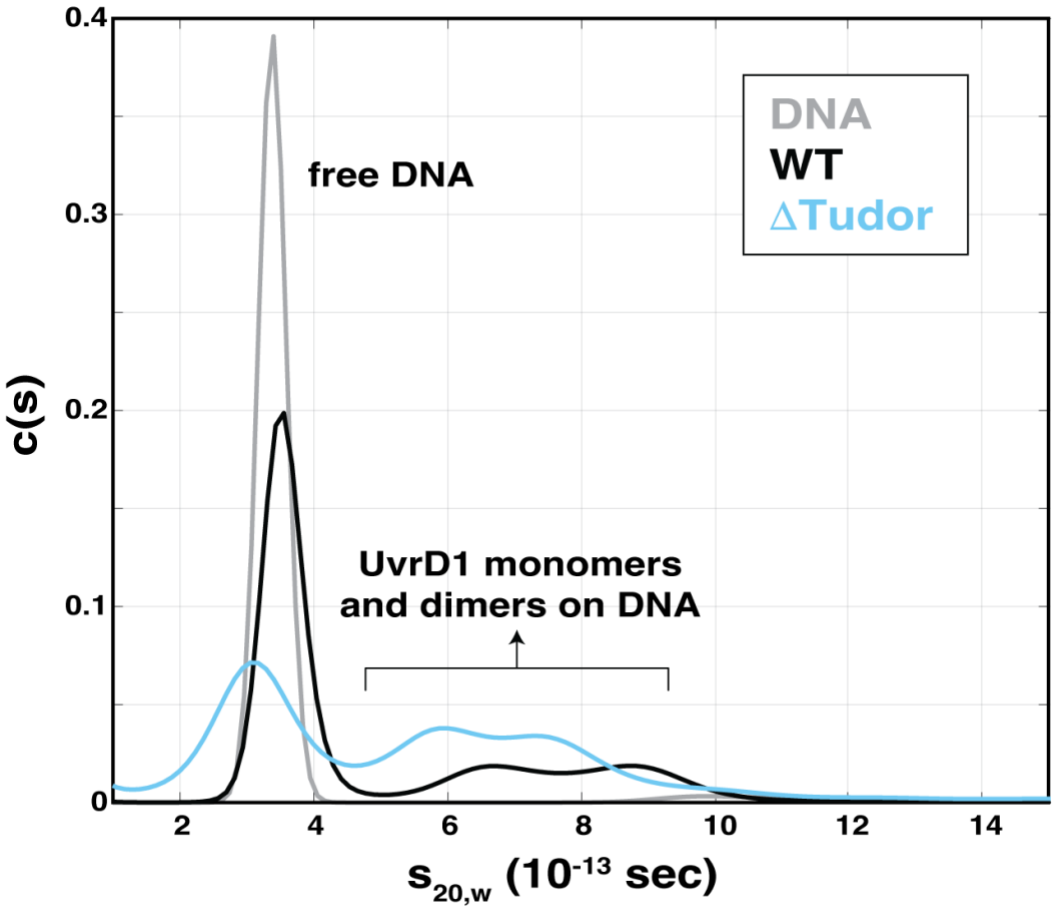
Supplementary Figure 4.3: Mtb Ku is a dimer in the presence and absence of magnesium. (A) Ku was run in AUC sedimentation velocity at the three concentrations indicated and, in each case, the data suggested a molecular species consistent with that of a dimer. (B) 3 mM Ku was run in AUC sedimentation velocity in the presence (black) and absence (red) of 5 mM $MgCl_2$. The data are consistent with a model where magnesium has no effect on the oligomeric state of Ku.



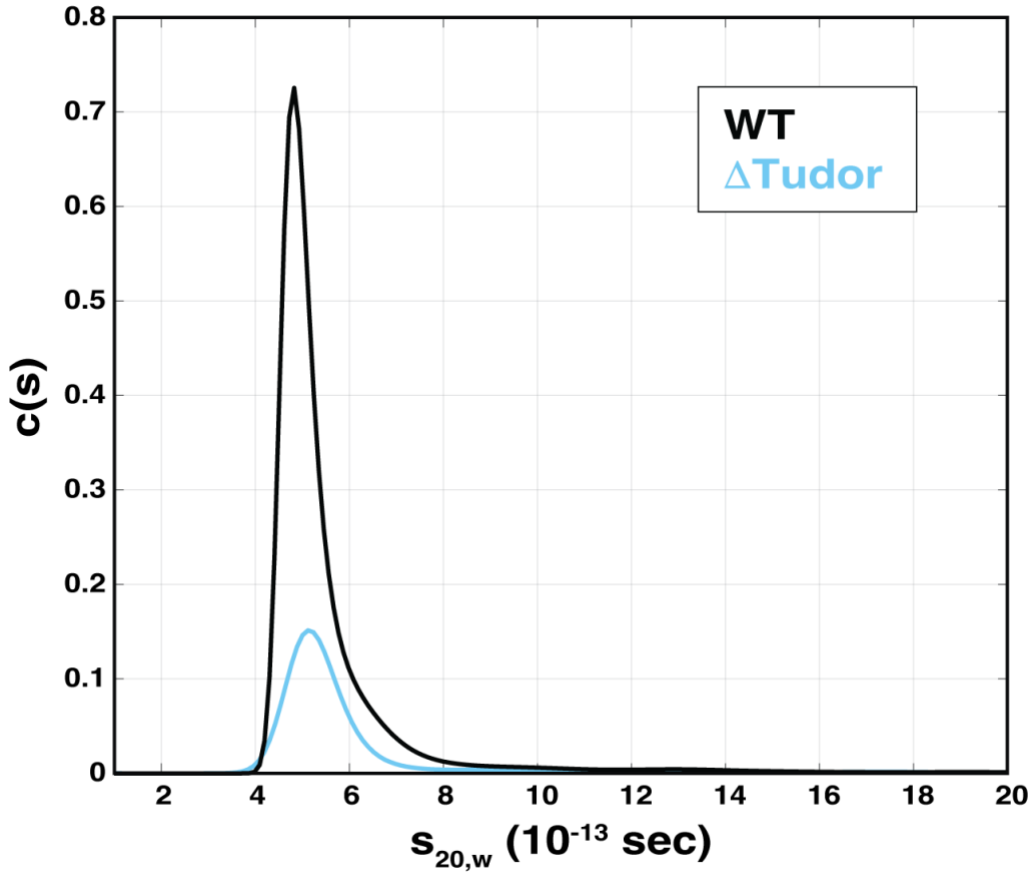
Supplementary Figure 4.4: Stoichiometric titration of Ku over DNA. The quenching signal from a 32 bp dT20 blunt end labeled template (as in Figure 5A in the main text) plotted as a function of the ratio of Ku dimer and template ratios. The linear increase in signal meets the saturation point at a ratio of 2.85, slightly higher than, but consistent with the estimate of stoichiometry obtained from the binding density analysis presented in the main text of 2.5.



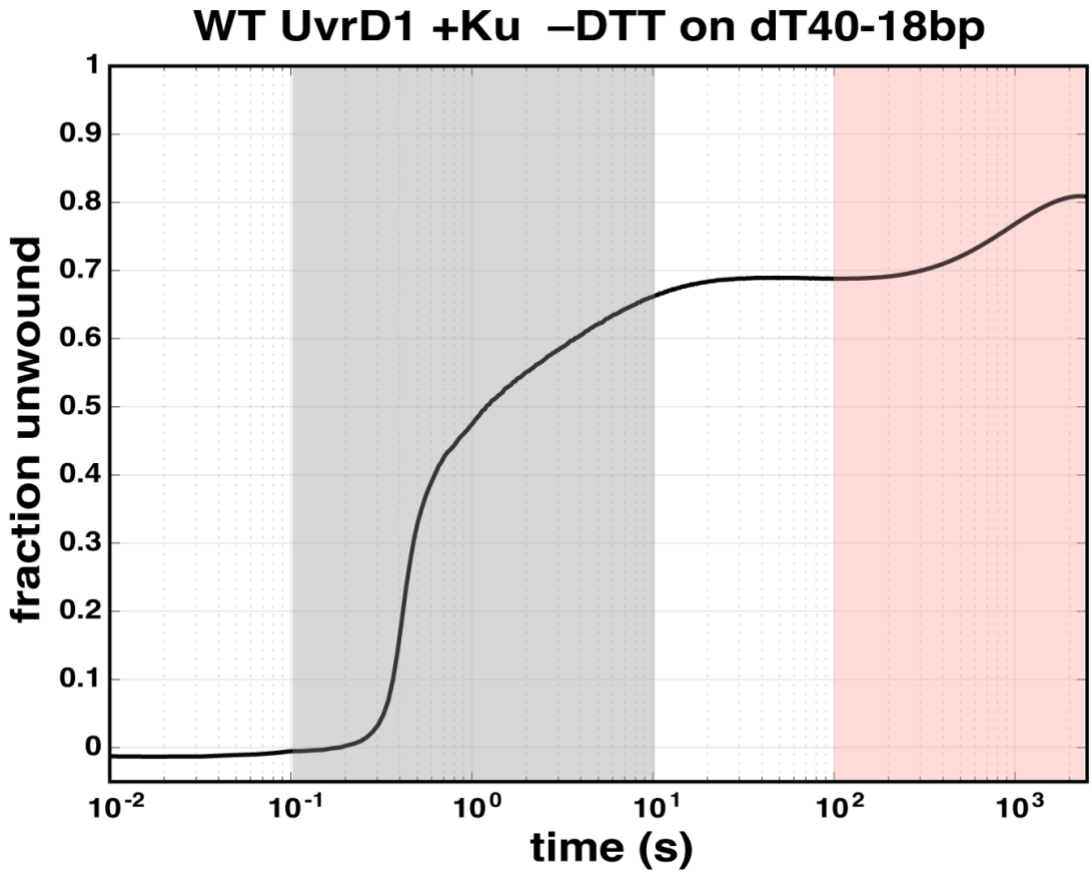
Supplementary Figure 4.5: Sequence alignment of C-terminal domain of SF1 helicases from E. coli, B. stearothermophilus, Mtb and M. smeg. Amino acids colored in red are those deleted in the referenced studies. The conserved Tudor domain identified by Sanders, et al. is highlighted in blue.



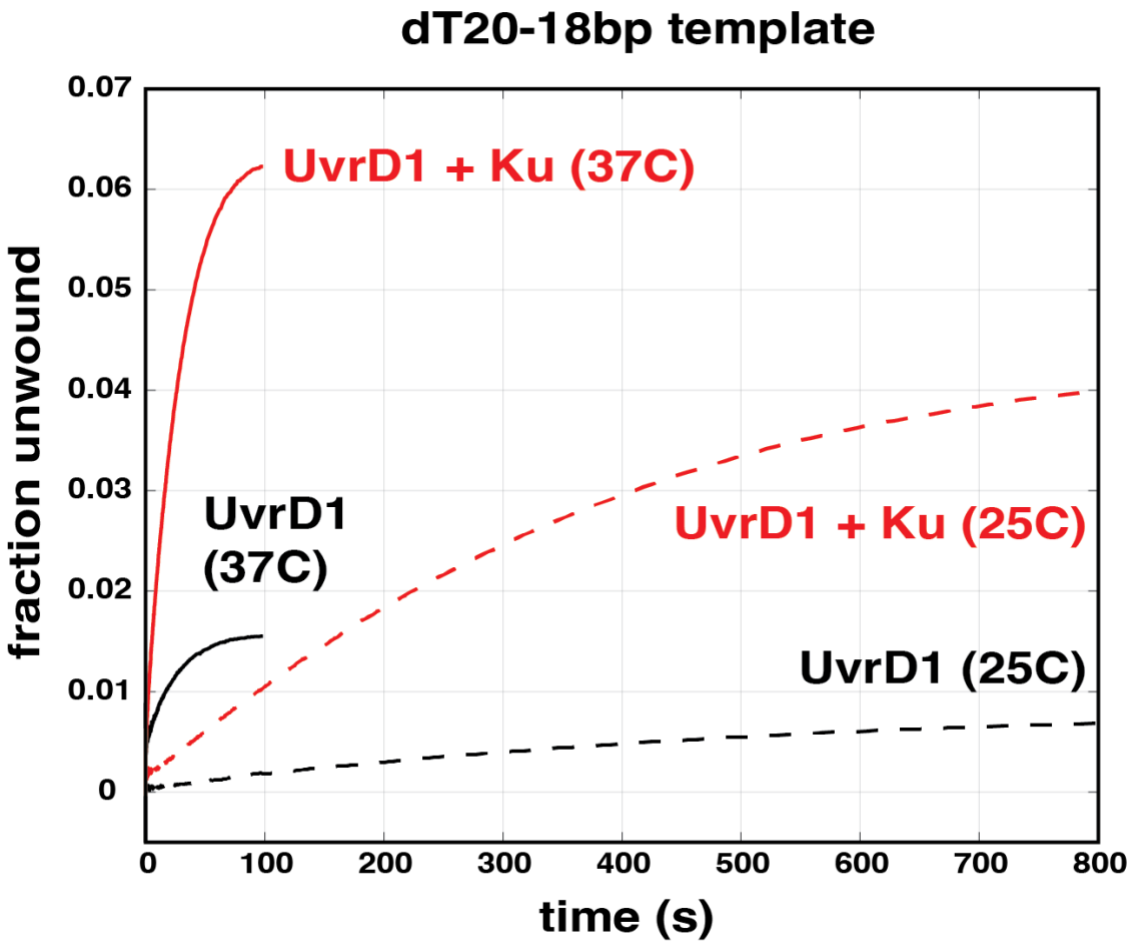
Supplementary Figure 4.6: Delta Tudor domain UvrD1 still binds DNA as a mixture of monomers and dimers in the absence of DTT. AUC velocity sedimentation data for labeled DNA alone (gray), WT UvrD1 (black), and DTudor UvrD1 (cyan). Only complexes containing the Cy5-labeled DNA were observed in this experiment.



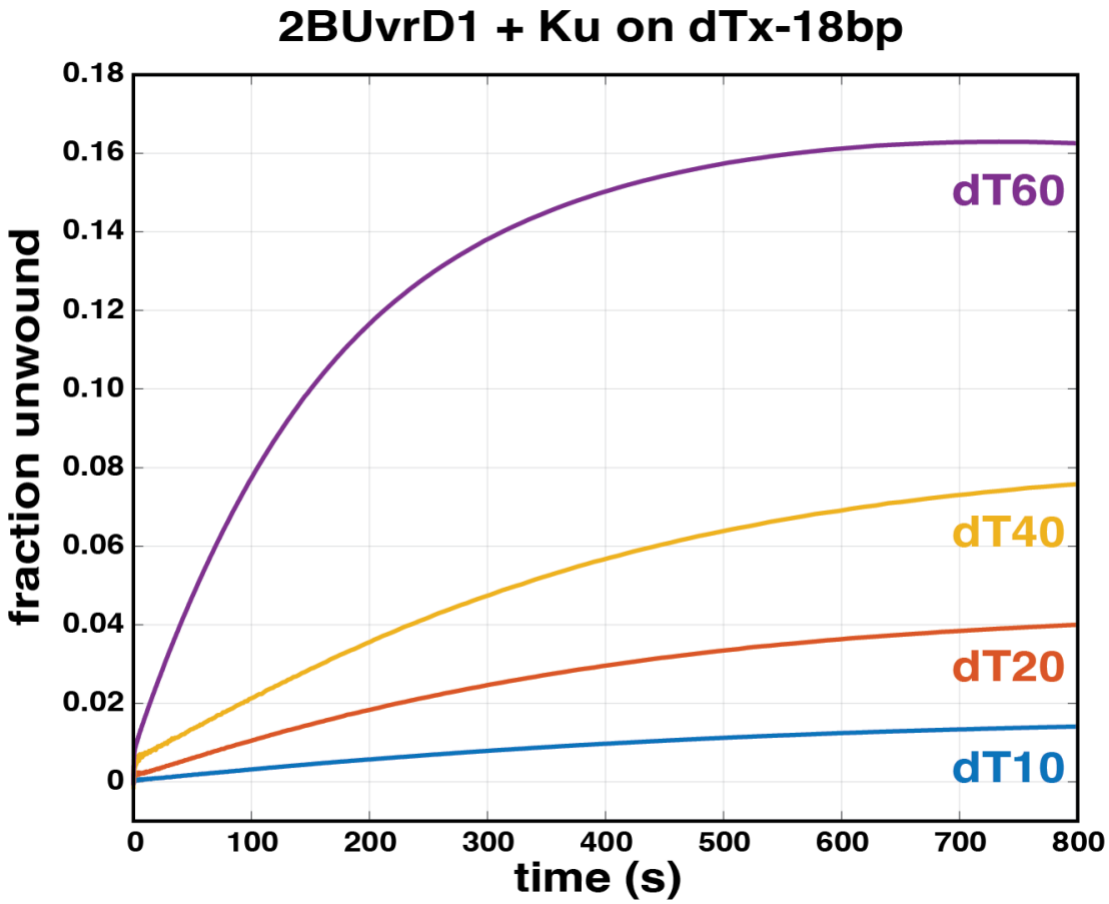
Supplementary Figure 4.7: D Tudor UvrD1 is monomeric in the presence of DTT. AUC velocity data for WT (black) and DTudor UvrD1 in presence of DTT (cyan).



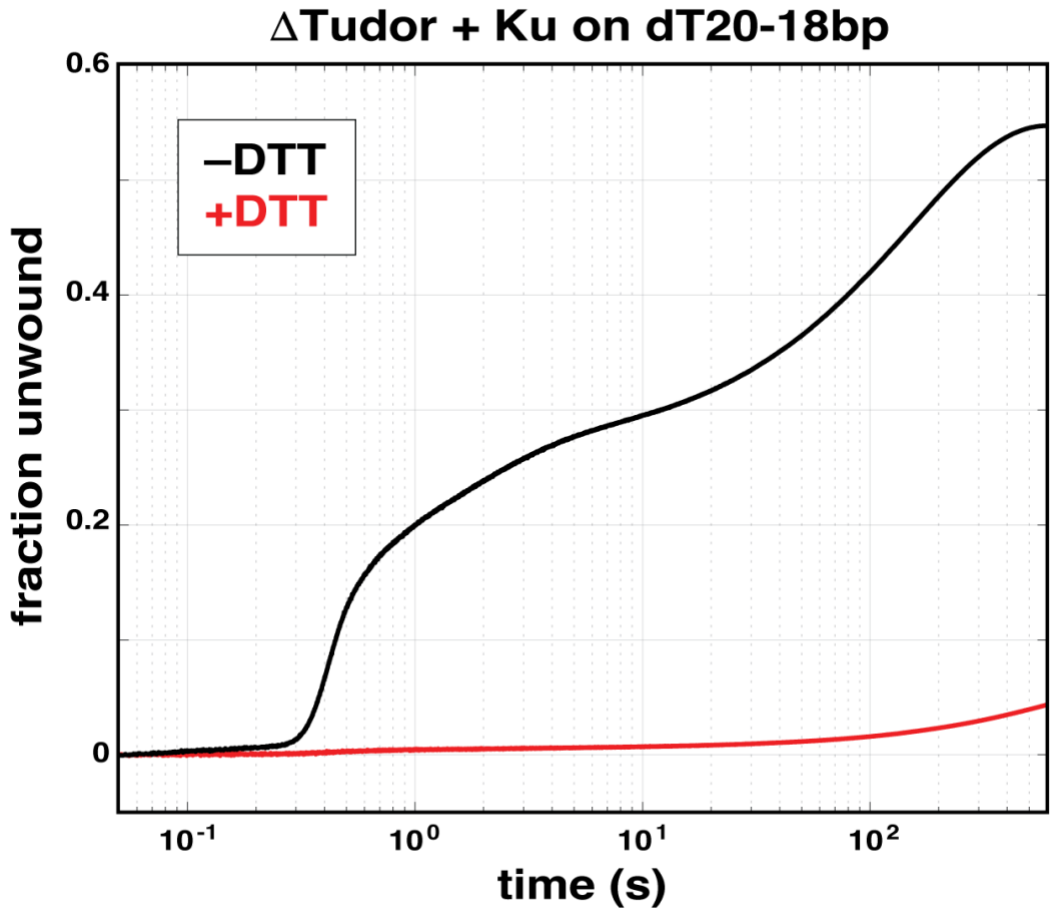
Supplementary Figure 4.8: Unwinding of WT UvrD1 with 18bp-dT40 in absence of DTT and presence of Ku on longer timescales: Multi-round DNA unwinding traces under oxidative conditions in the presence of Ku. The trace shows the sub-second unwinding of the UvrD1 dimer (gray shading) followed by a minutes-long second phase produced by monomeric UvrD1 in the presence of Ku (pink area).



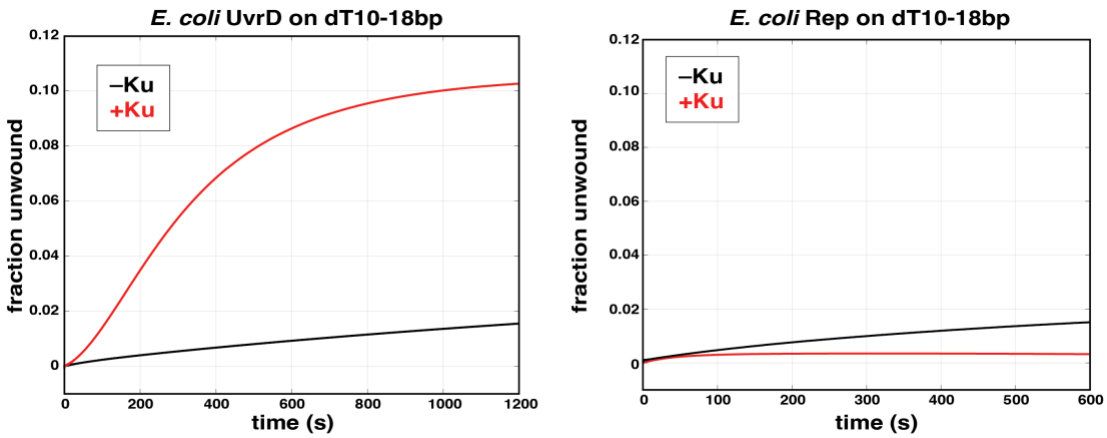
Supplementary Figure 4.9: The rate of unwinding of 2B UvrD1 in affected by temperature. Multi-round DNA unwinding traces of UvrD1 in the presence (red) and absence (black) of Ku at 37°C (solid) and 25°C (dashed). The rate of unwinding of an 18bp-dT20 substrate increases with temperature, but the fold-stimulation by Ku is comparable.



Supplementary Figure 4.10: Ku-dependent activation of monomer unwinding is single-strand tail dependent. Multi-round DNA unwinding traces in the presence of 2B UvrD1 on templates with increasing ssDNA tail lengths. The increase in fraction unwound with increasing tail length suggests that the ability to bind multiple monomers on a template increases the efficiency of the unwinding reaction.



Supplementary Figure 4.11: Unwinding of DTudor UvrD1 on a 32bp-dT20 DNA template. Multi-round DNA unwinding traces of DTudor UvrD1 in the presence (red) and absence (black) of DTT. The oxidation-dependent unwinding by UvrD1 dimers at short times can been observed in the black trace only while the slow phase due to the presence of monomers and stimulated by the removal of the Tudor domain can be observed in both traces.



Supplementary Figure 4.12: Mtb Ku stimulates DNA unwinding by E. coli UvrD monomers but not Rep monomers. (A) Multi-round DNA unwinding traces of E. coli UvrD in the presence (red) and absence (black) of Ku. (B) Multi-round DNA unwinding traces of E. coli Rep in the presence (red) and absence (black) of Ku. E. coli UvrD contains a C-terminal Tudor domain whereas Rep does not further suggest that activation by Ku is through this Tudor domain.

Supplemental Table 4.1: Sequences of single-stranded DNA used for unwinding and AUC

Leng	Sequence	sequence type/assay
18	5'GTT GGT CGG CAG CAG GGC-3'	ssDNA TRAP
10bp-	5'GCCTCGCTGCTTTTTGCAGCGAGGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3'	hairpin TRAP
52	5'-FAM GGA ATG TAA CCA TCG TTG GTC GGC AGC AG dT(20)-3'	Anisotropy assays
32	5'- GCC CTG CTG CCG ACC AAC GAT GGT TAC ATT CC	
32	5'-FAM GGA ATG TAA CCA TCG TTG GTC GGC AGC AG	
20	5'-FAM dT(20)3'	
38	5'-FAM GTT GGT CGG CAG CAG GGC dT(20)-3’	
18	5'-GCC CTG CTG CCG ACC AAC -3’	
39	5'-cy5 GTT GGT CGG CAG CAG GGC dT(40)-3’	Helicase assays
52	5'-cy5 GTT GGT CGG CAG CAG GGC dT(20)-3’	
32	5'-cy5 GGA ATG TAA CCA TCG TTG GTC GGC AGC AGG dT(20)-3'	
18	5'-cy5 3' GCC CTG CTG CCG ACC AAC GAT GGT TAC AT (BHQ_2)-3’	
78	5'-cy5 GTT GGT CGG CAG CAG GGC dT(60)-3’	
18	5'-GCC CTG CTG CCG ACC AAC (BHQ_2)-3’	
50	5'-CCT GCA CTA GCA/iCy5/GCA GCG GGA ATG TAA CC TCG TTG GTC GGC AGC AGG GC-3’	Analytical ultracentrifug
50	5'-GCC CTG CTG CCG ACC AAC GAT GGT TAC ATT CCC GCT GCT AGT GCA GG-3’	
90	5'-dT(20) CCT GCA CTA GCA/iCy5/GCA GCG GGA ATG T CCA TCG TTG GTC GGC AGC AGG GC dT(20)-3’	
38	5'-cy5 GTT GGT CGG CAG CAG GGC dT(20)-3’	

Chapter 5: Cryo-EM of *Mtb* UvrD1 Dimer

5.1 Introduction

Living organisms require motor proteins to move, transport cargo, divide, and perform work. All motor proteins depend on chemical or electrical energy to perform work[1]. Helicases are one such motor proteins that use energy in the form of ATP hydrolysis to translocate on single-stranded DNA (ssDNA) or double stranded DNA and unwind double-stranded DNA (dsDNA) to (ssDNA)[2,3]. As a result, they play a role in many important processes in the cell that involve interaction with nucleic acids such as DNA replication, repair, and transcription[4,5]. Helicases have been divided into two superfamilies (SF1 and SF2) with SF1 helicases unwinding only DNA and SF2 being able to unwind both DNA and RNA and both families sharing the helicase motifs. SF1 family includes monomeric or dimeric enzymes such as PcrA, UvrD, RecB and RecD that mostly interact with ssDNA while many SF2 interact with dsDNA[6,7]. SF1 helicases are ATPase motors and contain a Walker A or P-loop motif and two RecA like domains. Two RecA-like motor domains may belong to one polypeptide chain as in monomeric proteins or to adjacent protomers in a dimeric protein[8]

UvrD helicase belongs to the SF1 family of helicases and contains 2 RecA-like domains (1A and 2A) and two DNA binding domains (1B and 2B). Topologically, the B domains are inserted in domains 1A and 2A. UvrD and its homologs Rep and PcrA have been extensively characterized by kinetic, single-molecule, and crystallographic studies that revealed key elements of mechano-chemical coupling, which may be generally applicable to other helicases and motor proteins[9–14].

Crystal structures of a PcrA monomer bound to DNA with ATP analog complexes were the first to resolve the structure of an SF1 helicase and led to the proposal of a monomer unwinding DNA[12]. Following these studies, the crystal structures of UvrD bound to DNA, AMPPNP, and an ATP state were resolved that also showed a monomer of UvrD bound to DNA[14]. Since biochemical and kinetic studies have reported that UvrD-like helicases function as dimers and the structural studies have only revealed a monomer bound to DNA, we still don't have clear structural understanding of the mechanism of unwinding by UvrD like helicases[15].

During characterization of the UvrD1 an SF1 helicase from *Mycobacterium tuberculosis* (*Mtb*), I discovered that it forms a redox-dependent dimer which is the active form of the helicase[18]. Although a monomer of UvrD1 can bind and translocate on DNA it is unable to unwind DNA. To understand the conformational dynamics of *Mtb* UvrD1 dimer and how the two monomer subunits interact and unwind DNA the goal of this study is to obtain a high-resolution structure of UvrD1 dimer alone and bound to DNA using Cryo-electron microscopy (Cryo-EM) which has been used to resolve the structures of many helicases[16,17].

5.2 Structure of *Mtb* UvrD1 dimer in presence of crosslinker and detergent

5.2.1 Negative staining and screening for *Mtb* UvrD1 dimer

For acquiring the structure of *Mtb* UvrD1 dimer, I first performed negative staining of purified *Mtb* UvrD1 to check homogeneity, monodispersity, and distribution of protein on the grid. Analyzing the images suggested that the protein was not monodisperse (data not shown), so we decided to optimize conditions and freeze the protein in the presence of crosslinking agents such as BS3 along with detergents such as Digitonin and CHAPSO (Figure 1A and 1B).

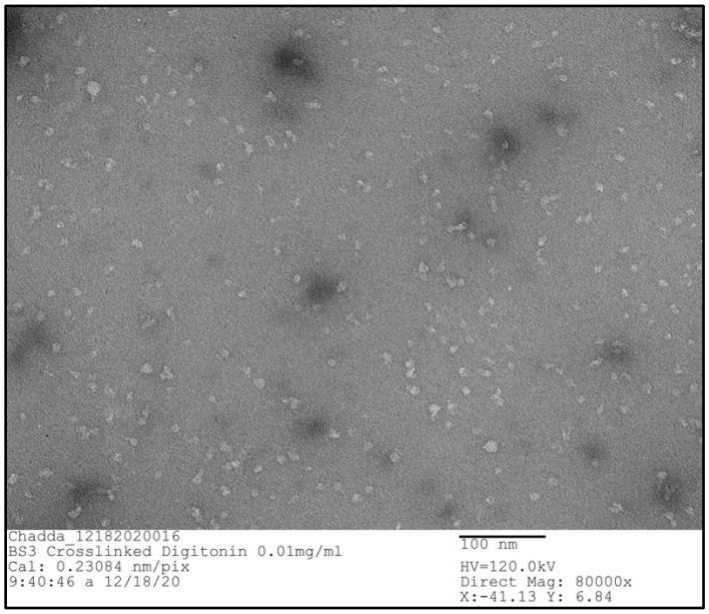


Figure 5.1.1: Negative staining Cryo-EM image of *Mtb* UvrD1 dimer shows monodisperse population in presence of BS3 crosslinker and Digitonin detergent.

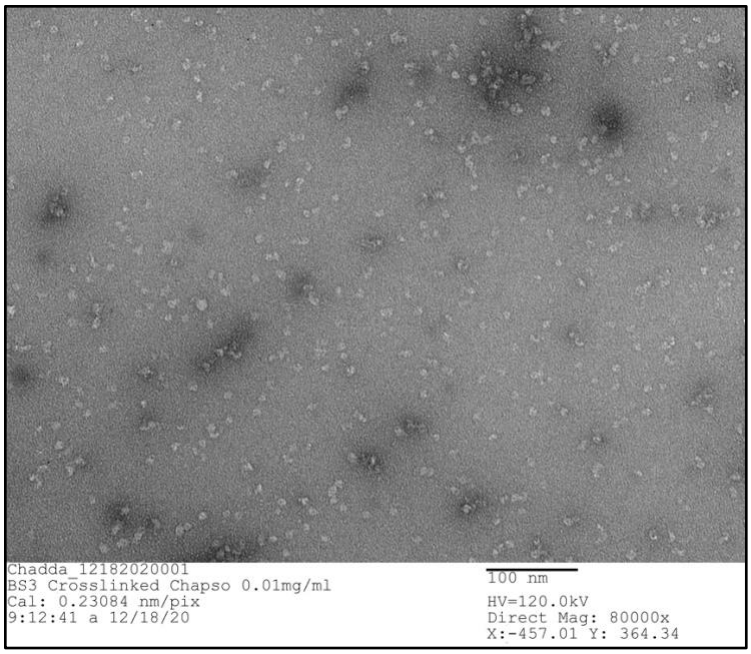


Figure 5.1.2: Negative staining Cryo-EM image of *Mtb* UvrD1 dimer shows the monodisperse population in presence of BS3 crosslinker and CHAPSO detergent.

Since we could see monodisperse population of UvrD1 proteins by using crosslinker at a concentration of 0.5 mM BS3 in presence of Digitonin and CHAPSO detergents we decided to freeze and acquire images of *Mtb* UvrD1 under these conditions. *Mtb* UvrD1 dimer was purified by Ni-NTA chromatography and ion exchange as mentioned [18]. However, during size exclusion, the TRIS buffer was replaced with 50 mM sodium phosphate pH 7.4, 3% glycerol, 100 mM NaCl. The protein yield was reduced because of lower glycerol but about 1 ml UvrD1 (20 µM) UvrD1 was obtained from 3 Liter of LB culture. The protein was crosslinked using 0.5 mM BS3 for 45 minutes at room temperature and after quenching the reaction with TRIS pH 8.0 for 15 minutes the protein was dialyzed for 3 hours at room temperature to remove unreacted crosslinker in dialysis buffer (50 mM sodium phosphate pH 7.4, 3% glycerol, 100 mM NaCl) and after concentrating it to 8 µM or 0.3 mg/ml and 0.005% Digitonin detergent it was given to cryo-EM facility for freezing. Sample images were tested during grid screening for monodisperse population and particle distribution as shown in (Figure 1C). Following grid screening a set of 2500 images were acquired using Glacios 200 keV microscope. The images were gain corrected and collected at Raw pixel size (A): 0.94 A, Accelerating voltage (kV): 200 kV, Spherical aberration (mm): 2.7 mm, and Total exposure dose (e/A^2): 50 e/A^2.

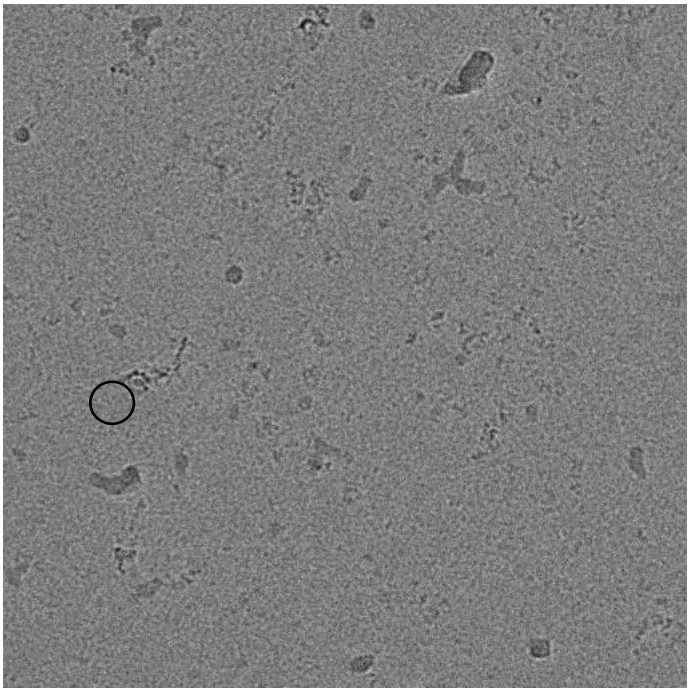


Figure 5.1.3: Cryo-EM image of *Mtb* UvrD1 dimer during screening of grids. The circled regions depict monodisperse UvrD1 particles shown at a low magnification that were evaluated for ice quality at low magnification and particle distribution at high magnification.

5.2.2 Workflow of image processing in Cryo-SPARC

The raw images were collected by WashU center for Cryo-electron microscopy (WUCCI) and provided to us on 10 TB hard drive for transfer and analysis on remote cryo-EM computer belonging to BMB department. I transferred and analyzed the images using the cryo-SPARC software. The images were first imported into cryoSPARC software which can be accessed via bmb.cryosparc.wustl.edu website. After logging into cryoSPARC, the webpage shows information about the software version and job history as shown in (Figure 2A).

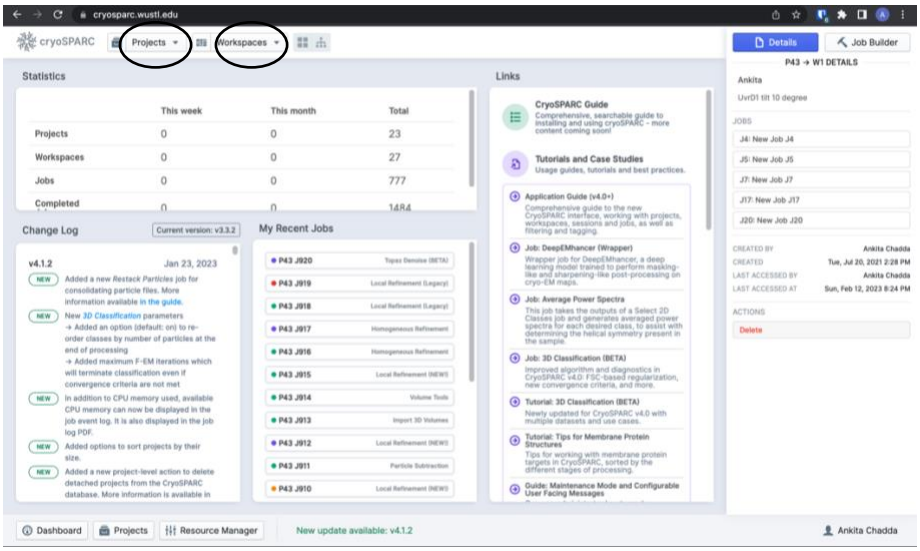


Figure 5.2.1: Dashboard for cryoSPARC showing projects and workspaces.

5.2.3 Image preprocessing

By clicking "Projects," you can add the title of the project and start preprocessing movies. By clicking "Import movies," 4400 movies were imported by defining the path and the parameters mentioned above. Moving from one job to the next in cryo-SPARC involves the Drag and drop approach where the output from one job can be dragged and used as input for another job. After the movies have been imported, Patch motion correction is performed in which the raw movie frames are aligned to account for sample and stage movement during the exposure before producing an aligned average micrograph (Figure 2B). After correction for motion of the particles, the Patch CTF estimation uses the output from Patch motion correction as input and attempt to measure additional parameters that vary from one micrograph to another like astigmatism, defocus, and estimated resolution (Figure 2C).

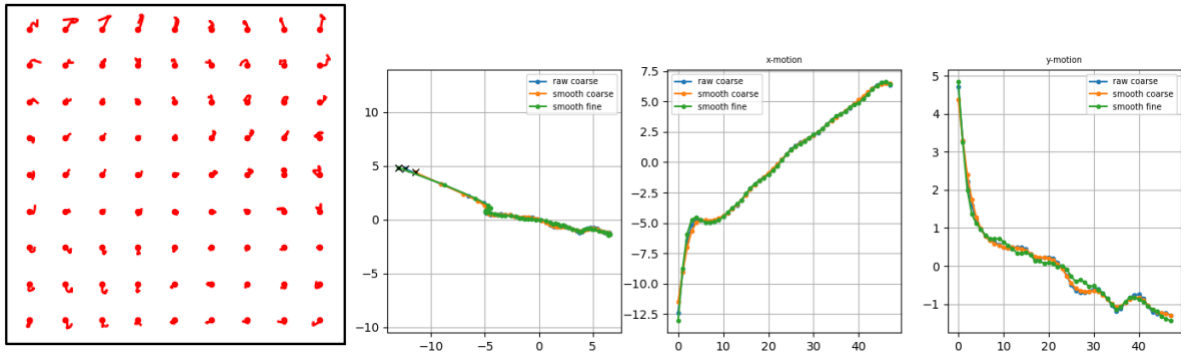


Figure 5.2.2: Patch motion micrograph showing correction for sample and stage movement showing motion of the particle in X and Y axis.

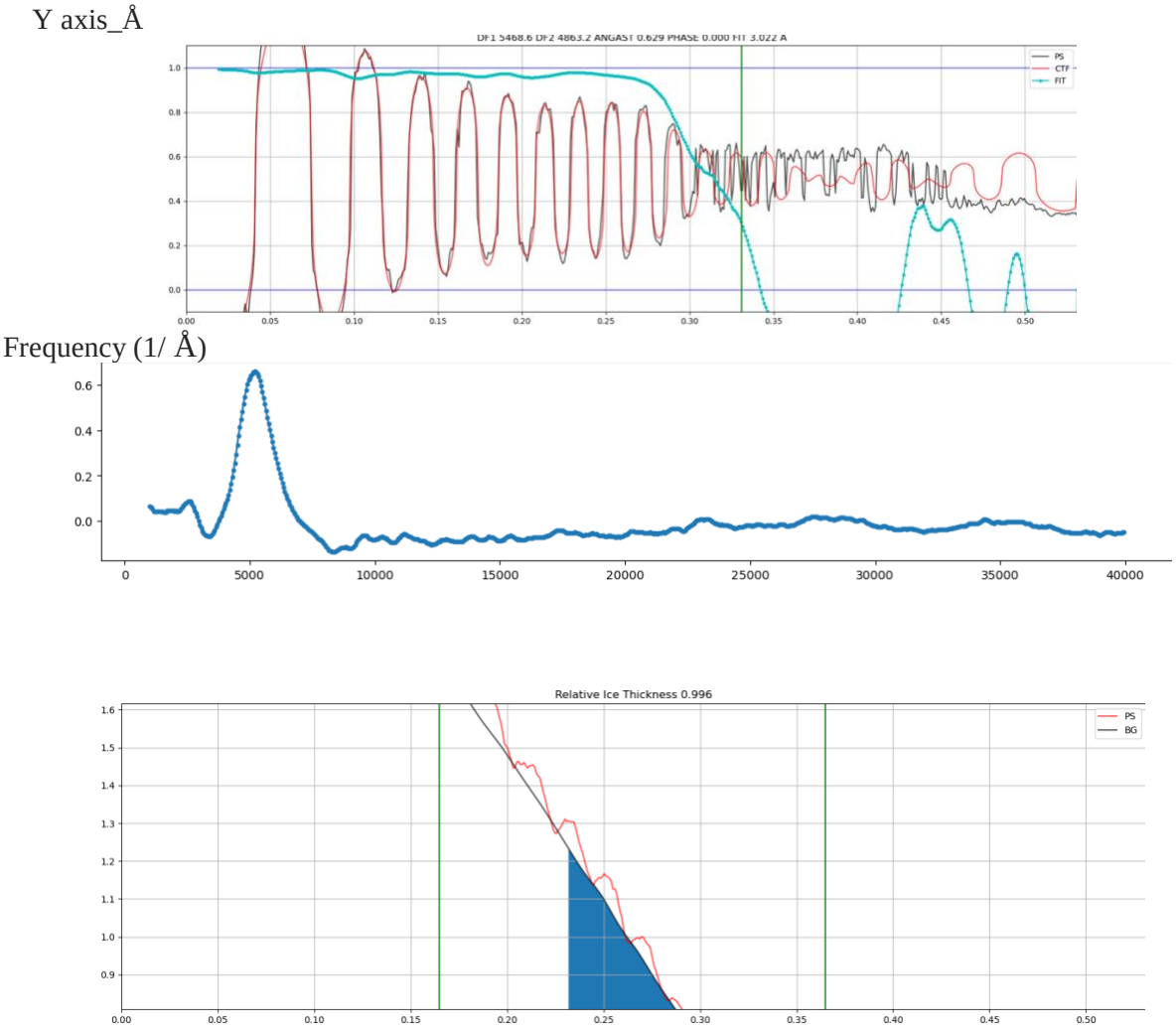


Figure 5.2.3: CTF correction of one of the input micrographs showing frequency. The X, Y and Z axes units are in Armstrong.

5.2.4 Particle picking and 2D classification

After preprocessing of images, the next step is to pick particles using a job called, “blob picker” that picks particles using Gaussian signals. The output from the Patch CTF was transferred as input to the blob picker and the minimum particle diameter was chosen to be 130 Å as the diameter of a monomer of UvrD1 in crystal structure is approximately 65 Å[19]. The image of the micrograph that assigned particles based on the blob picker is shown in Figure 2D.

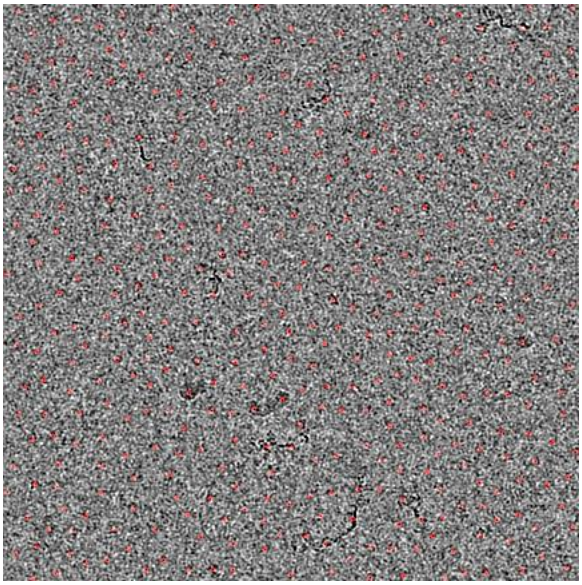


Figure 5.2.4: Micrograph showing particles picked by blob picker.

After blob-picking the particles must be extracted from the blobs before running them for 2D classification. There are two parameters that need to be assigned before running the extraction job. One the extraction box size and the other is the Fourier crop size. The extraction box size is picked using the equation:

$$\text{Box(Å)} = \text{Maximum defocus(Å)}/25 * \text{Best resolution(Å)}$$

Keeping the extraction box size at 350 Å (approximately two times the size of my particle) and the Fourier crop box size at 200 Å, which is generally set at half the value of extraction box size initially to increase the processing speed and then after obtaining the initial structure can be increased so that high resolution information is not lost. I extracted 1,135,265 particles from a dataset of 4400 movies. After the extraction step, the particles are input to a 2D classification job where there are defined set of parameters by the software to be used for 2D classification. I modified the number of assigned classes to 150 that results in 2D classes as shown in (Figure 2E). The increase in the number of classes during the process of 2D classification helps in finding different orientations of protein.

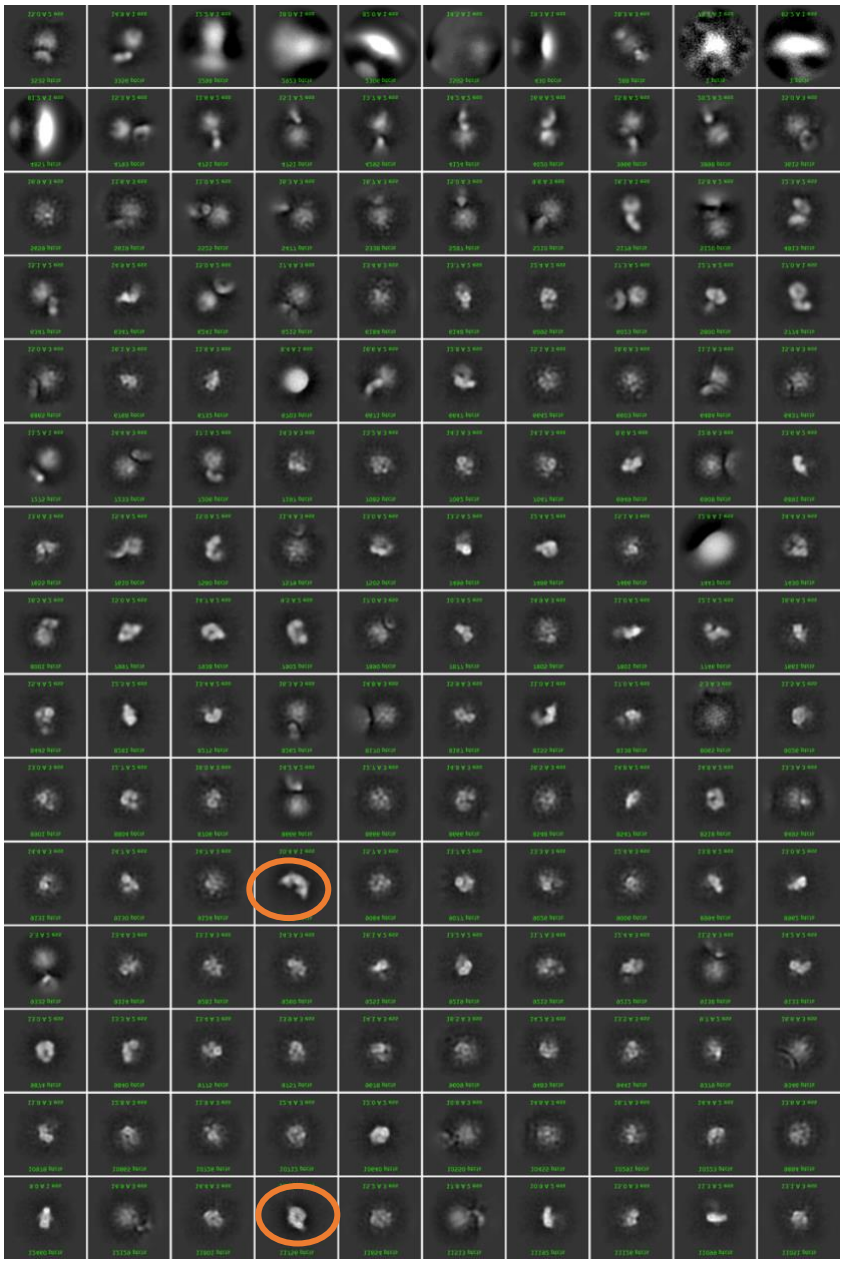


Figure 5.2.5: First 2D classification after extraction of particles using blob picker. The two classes marked by orange circles may represent two different conformations of the *Mtb* UvrD1.

After 2D classification, create a Select 2D Classes job and use the outputs from 2D Classification as inputs by drag and drop approach (the “class_averages” output will go into the “templates” input). Run the job. The results of the select 2D classes are shown below. The classes that showed particles in closed and open conformations were picked during this process.

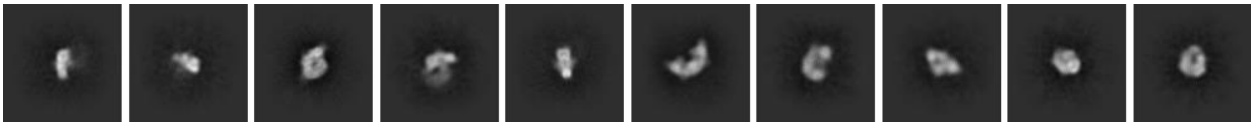


Figure 5.2F: Select 2D class job used to pick templates of UvrD1 for the Template picker.

The outputs from the 2D select job that showed defined features were used to pick templates and create another particle-picking job called Template picker providing select 2D classes as a template. 267,142 particles were extracted from 2500 micrographs and passed through 2 rounds of 2D classification.

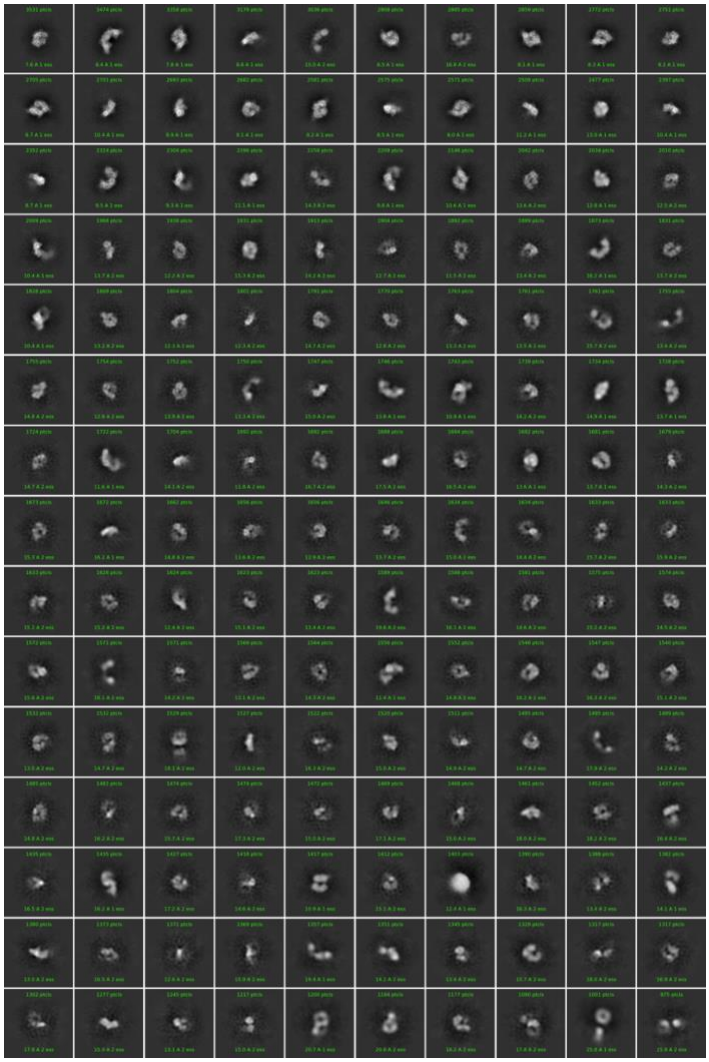


Figure 5.2.6: Third round of 2D classification job with good number of classes.

After the 3rd round of 2D classification, 55 classes were selected as shown below with 118,913 particles to be further refined via 3D classification. To improve the quality of my particle dataset, I avoided selecting classes that contain only a partial particle, two or more particles, or a non-particle junk image (e.g., ice crystals) and therefore selected only 55 classes.

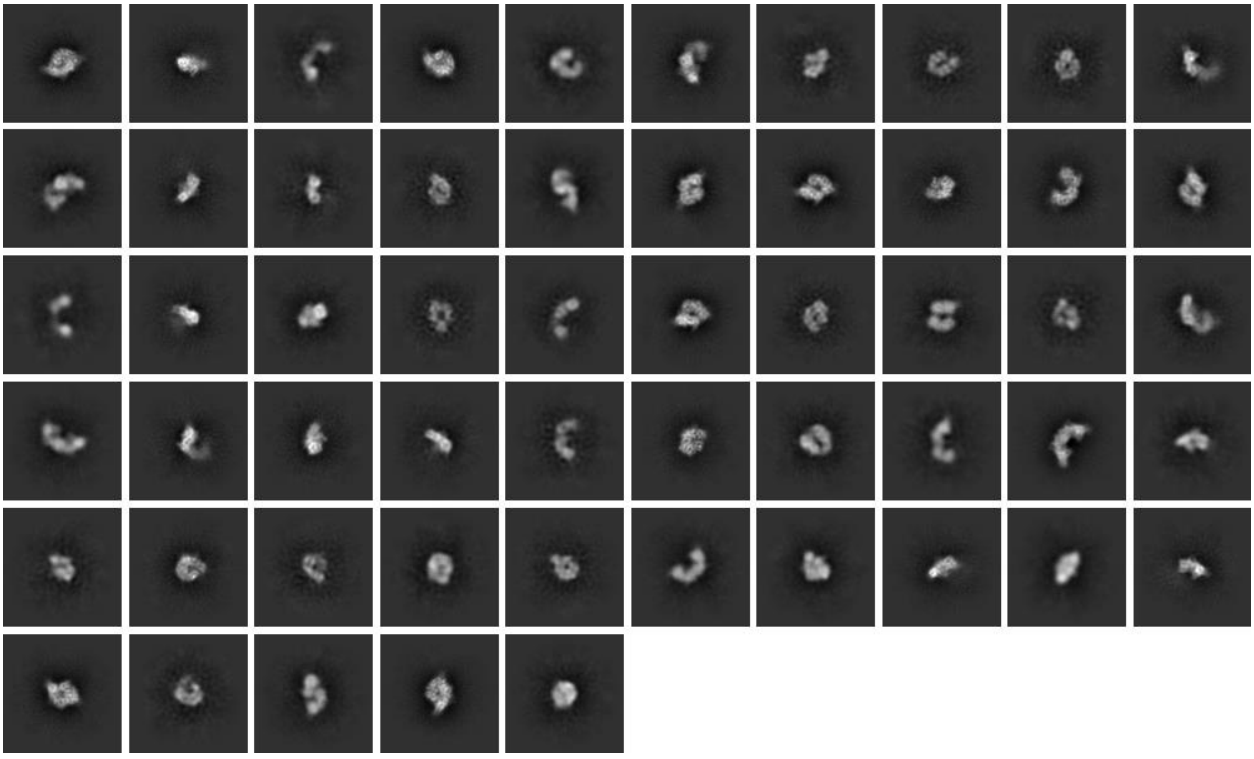


Figure 5.2.7: 2D select job used to pick the classes for 3D classification.

5.2.5 3D reconstruction and refinement

As shown above, good classes were obtained after the 3rd round of 2D classification, so the next step was to obtain a 3D reconstruction of the particles selected in 2D classes. The 3D reconstruction involves various steps with the first step being ab initio modeling to obtain an initial low-resolution 3D map of the structure. In this process, single or multiple 3D maps are

reconstructed from a set of particles, without any initial models or starting structures. The default number of 3D classes for this program is 1. However, by increasing this value between 1 to 4 one can expect to find multiple conformations/states in the data and identify junk particles. I increased the value to 4 as that led me to a cleaner dataset and removal of junk particles.

The results of Ab-initio reconstruction used to reconstruct one or more 3D maps from a set of particles, with a model independent analysis are shown below in (Figure 3A). Using 4 classes during Ab-initio gave the best separation of classes as shown below:

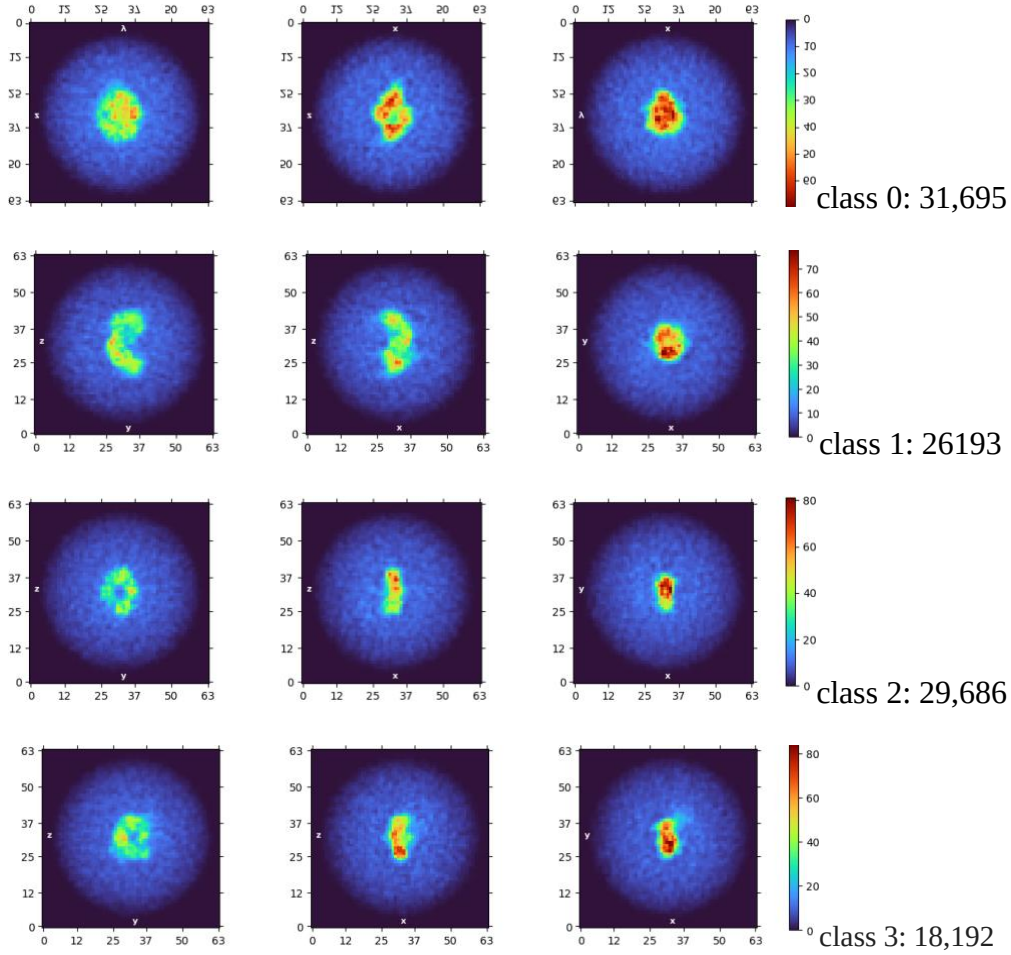


Figure 5.3.1: Model-independent 3D ab-initio reconstruction job that assigns particles for 3D classification where two different views of UvrD1 can be seen in class 1 and 2.

After Ab-initio reconstruction the output of the particles from class 1 and 2 is added as input in another 3D classification job called heterogeneous refinement for further removal of junk classes.

Class 1 and 2 obtained using Ab-initio reconstruction were refined using Heterogenous refinement to facilitate the ability to look for small differences between structures that may not be obvious at low resolutions, and to re-classify particles and remove junk particles to aid in sorting.

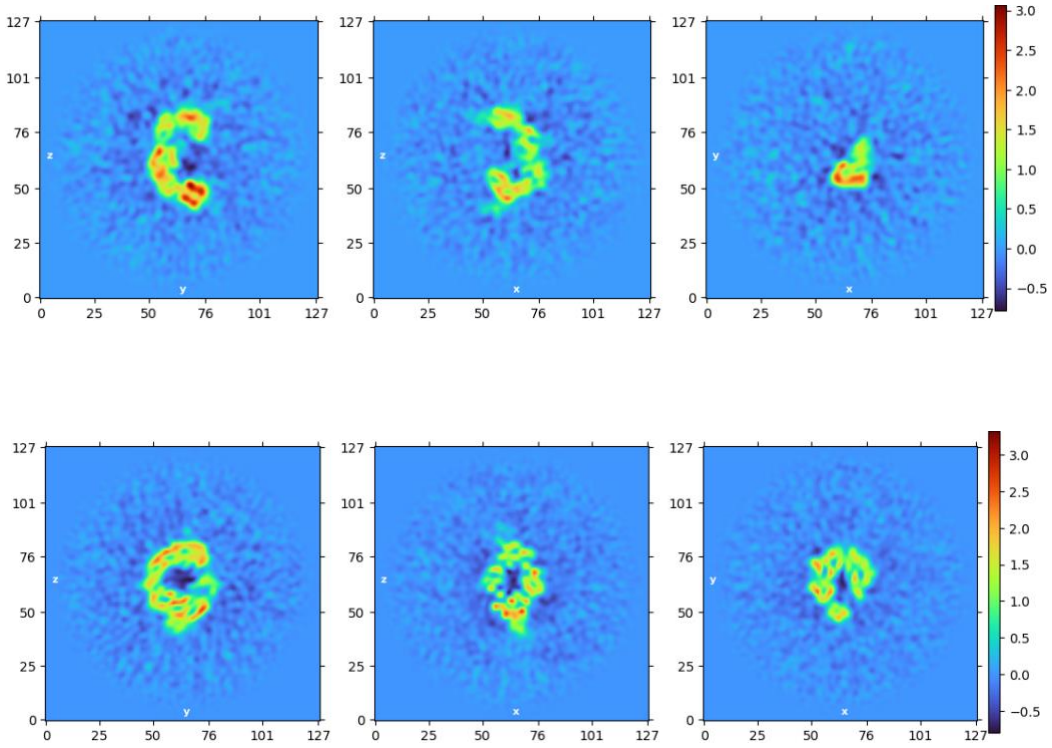


Figure 5.3.2: 3D Heterogenous refinement job that assigns particles for 3D classification with different conformations of UvrD1.

After Heterogenous refinement, the particle stack composed of 30K particles in each group is passed through another set of refinement jobs to get a high-resolution structure. Out of 4 classes during heterogenous refinement two classes were junk as they were noisy and did not show any structural features and two classes that have refined features were passed first through homogenous refinement to get a high-resolution structure. Viewing those structures in cryoSPARC reveals two different conformations that I will call closed and open based on the movement of RecA domains (Figure 5.3.3 and 5.3.4).

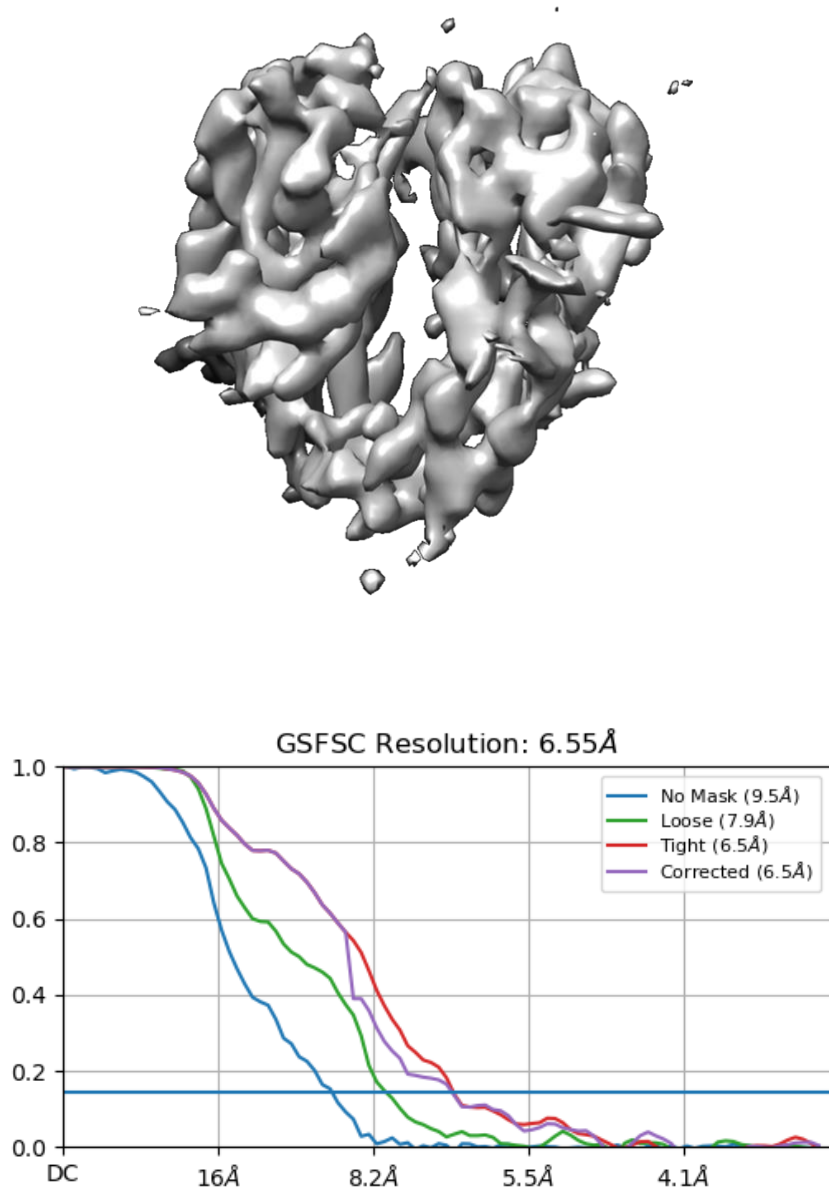


Figure 5.3.3: Homogenous refinement yield electron density map of closed conformation where the RecA domains of both monomeric subunits are closer to each other as shown above with the Fourier shell correlation (FSC) of 6.55Å.

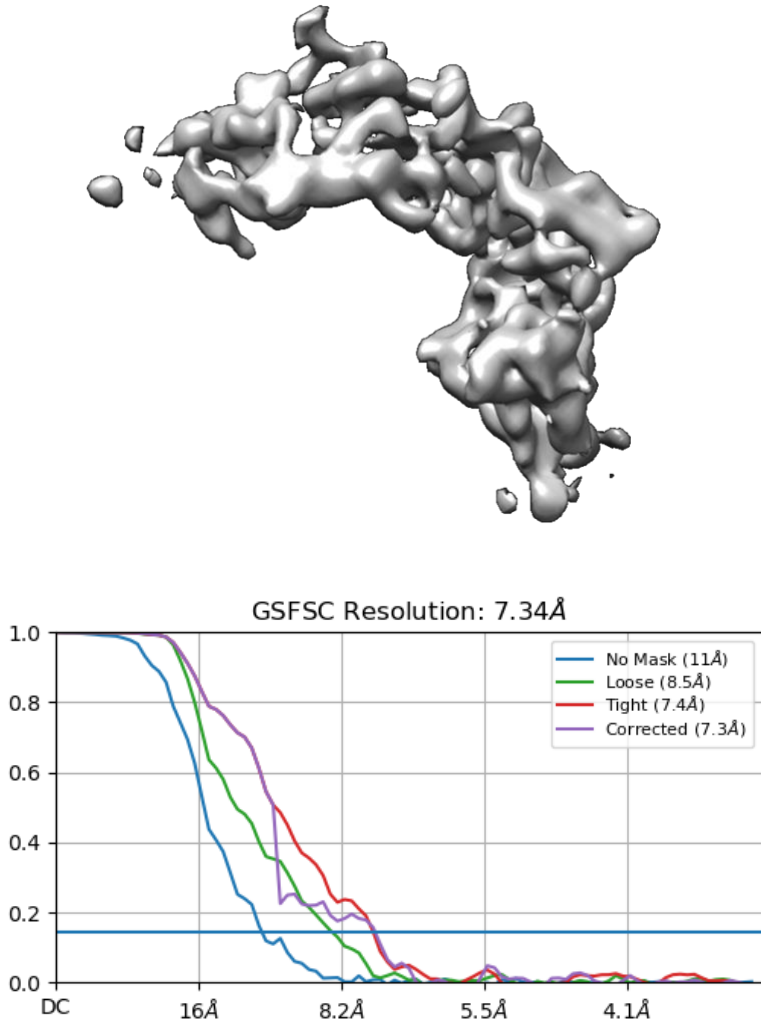


Figure 5.3.4: Homogenous refinement yield electron density map of open form of *Mtb* UvrD1 dimer where the RecA domains of both monomeric subunits are further apart as shown above with the Fourier shell correlation (FSC) of 7.34Å.

After homogenous refinement, non-uniform refinement was used that uses non-uniform adaptive regularization and performs highly efficient pose and shift marginalization over particle alignments during reconstruction steps, with a sampling strategy that automatically tunes to each particle's signal quality and noise level. Adaptive Marginalization can improve 3D structure quality on smaller proteins, as well as sometimes limit the influence of outlier images. The non-uniform refinement for closed conformation of *Mtb* UvrD1 yielded a relatively higher resolution

density map as shown in (Figure 3E). Another refinement procedure in Cryo-SPARC called local refinement was used for the open form of *Mtb* UvrD1 where the number of particles were increased by changing the symmetry of particles from C1 to C2 as both monomeric subunits are similar, and that increased the resolution by addressing the issue of sample flexibility as shown in (Figure 3F).

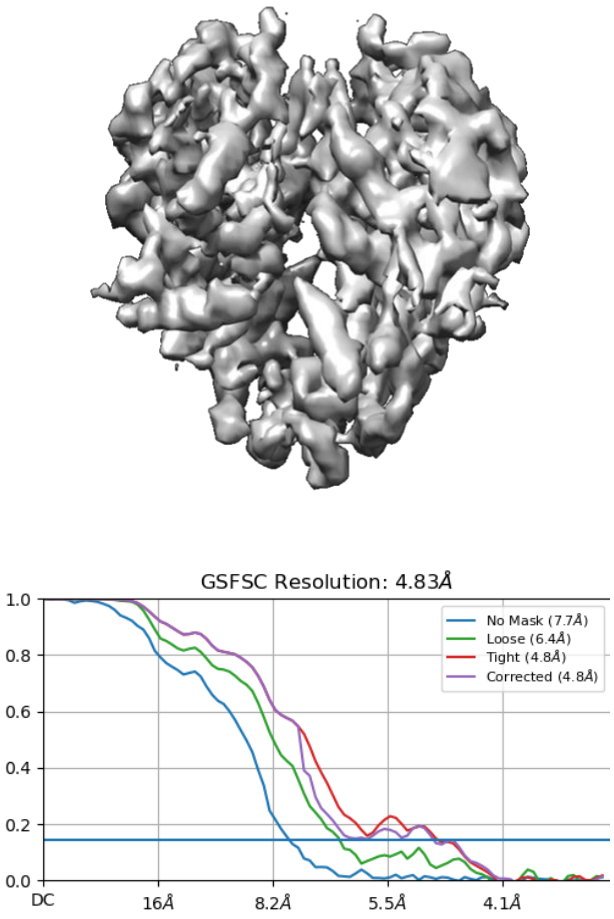


Figure 5.3.5: Non uniform refinement yield electron density map of closed form of *Mtb* UvrD1 dimer shown above with the Fourier shell correlation (FSC) of 4.83Å.

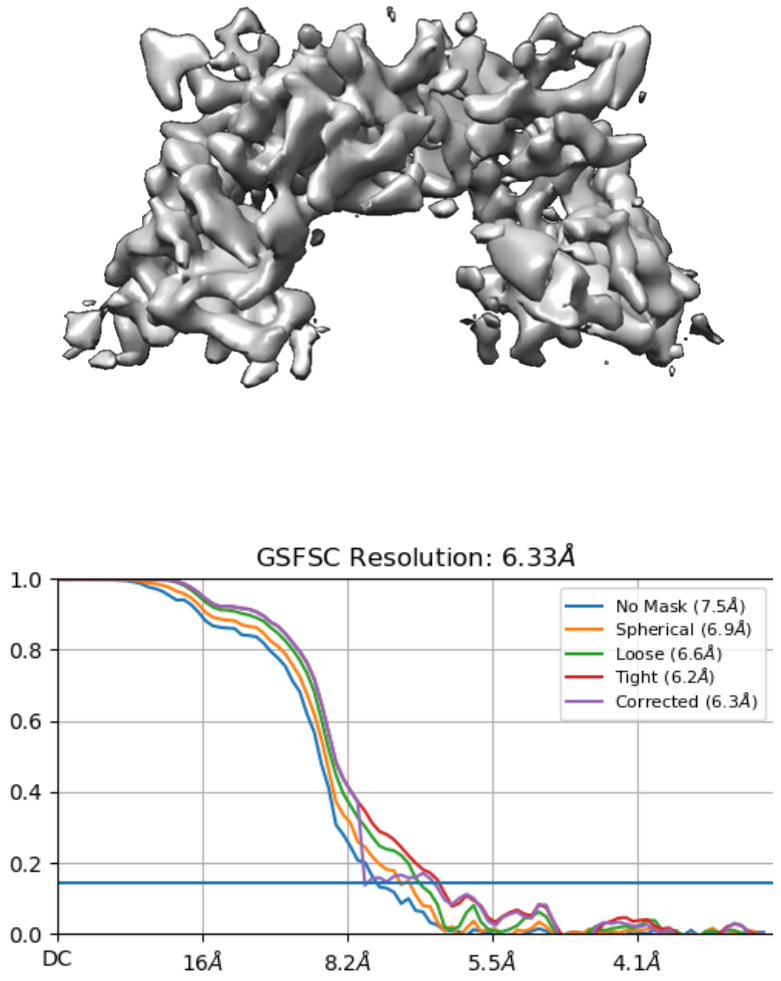


Figure 5.3.6: Local refinement yield electron density map of open conformation of *Mtb* UvrD1 dimer shown above with the Fourier shell correlation (FSC) of 6.33Å.

After local refinement the open conformation of UvrD1 dimer appeared noisier, so I decided to perform contrast CTF Refinement that includes two major components: local (per-particle) CTF refinement and global (per-group) CTF refinement. Local CTF Refinement adjusts each particle's defocus value to estimate the z-position of the particle in the sample/ice. Global CTF Refinement however adjusts the higher-order CTF terms (beam-tilt, trefoil, spherical aberration, tetrafoil)

across an entire group of images to find the optimum values, accounting for misalignment or aberrations in the microscope itself. Performing second round of local refinement after passing particle through CTF refinement removed some noise but did not improve the overall resolution for the open form of UvrD1 dimer as shown in (Figure 3G).

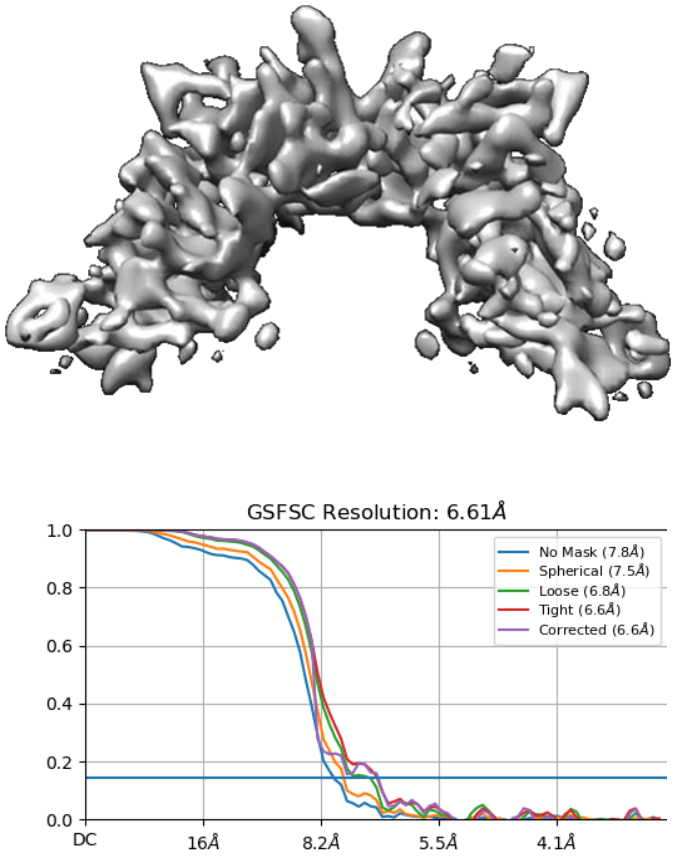


Figure 5.3.7: Local refinement of open conformation of *Mtb* UvrD1 yield electron density map of open conformation of *Mtb* UvrD1 dimer shown above with the Fourier shell correlation (FSC) of 6.33Å.

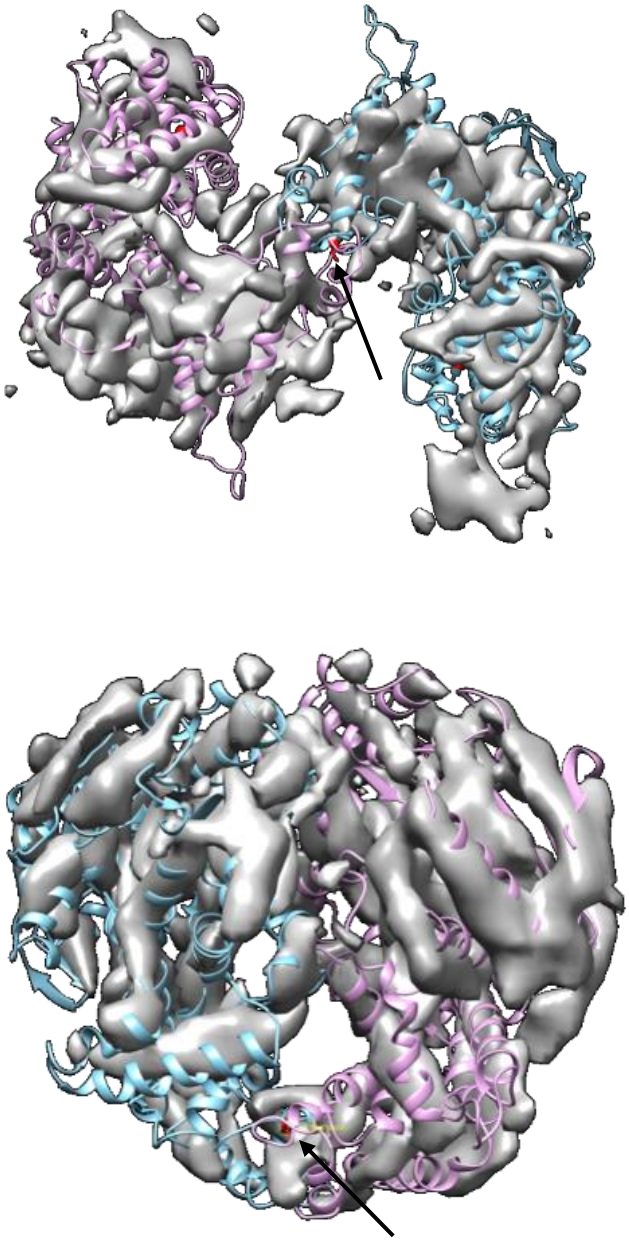


Figure 5.3.8: The electron density maps of both open and closed form of *Mtb* UvrD1 are fit with PDB model of predicted structure of *Mtb* UvrD1 dimer using Chimera. The closed and open forms fit with the correlation value of 0.84 and 0.503 calculated using Chimera software.

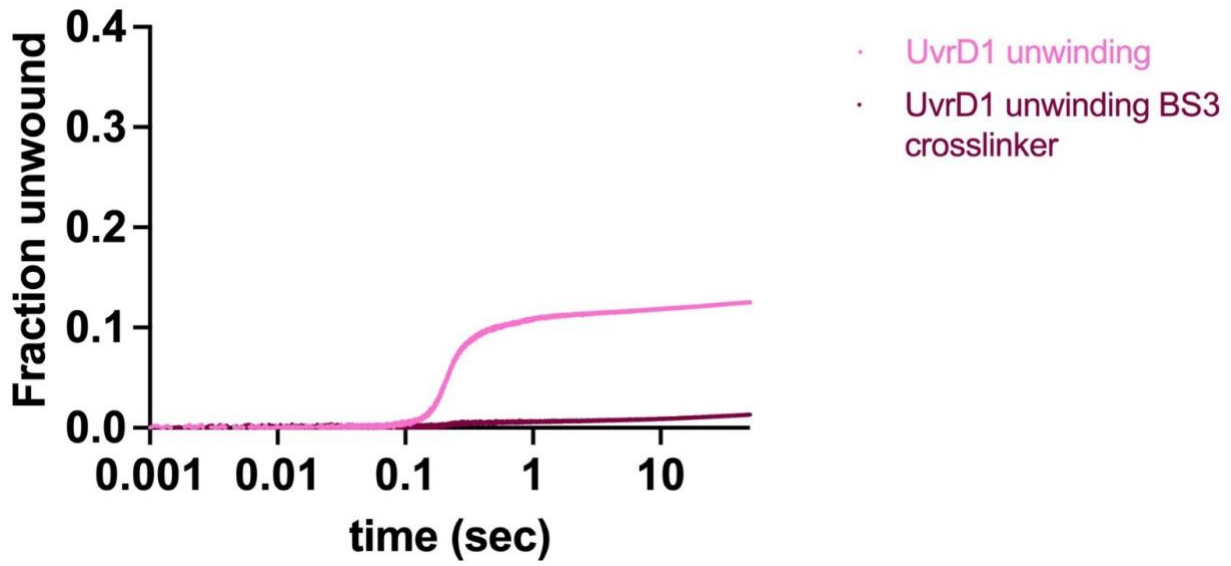


Figure 5.3.9: Unwinding assay of UvrD1 showing fraction DNA unwound with 18bp with dT20 in presence and absence of a crosslinker.

As shown in figure 3I, addition of a crosslinker abrogated UvrD1 dimer helicase unwinding. While many cryoEM structures are solved and interpreted in the presence of crosslinkers, we decided to try get a structure of UvrD1 dimer in absence of crosslinker but in the presence of detergent.

5.3 Structure of *Mtb* UvrD1 dimer in absence of crosslinker

Mtb UvrD1 dimer sample was prepared for negative staining in presence of two detergents either Digitonin (concentration 0.005% or Amphipol concentration 0.01%) in final buffer conditions (TRIS pH 8.0, 75 mM NaCl and 5% glycerol). The pellet collected from 3 Liter LB was sonicated and ion exchange in buffers containing TRIS in 10% glycerol as mentioned in the methods[18]. However, sizing was performed in a TRIS pH 8.0, 100 mM NaCl containing 5% glycerol and after collecting the sample from sizing it was immediately concentrated to a desired concentration of 0.3 mg/ml or 8 mM with detergents Digitonin (0.005%) or Amphipol 4 mM (between 0.01-0.025%) and the samples were provided for negative staining and freezing.

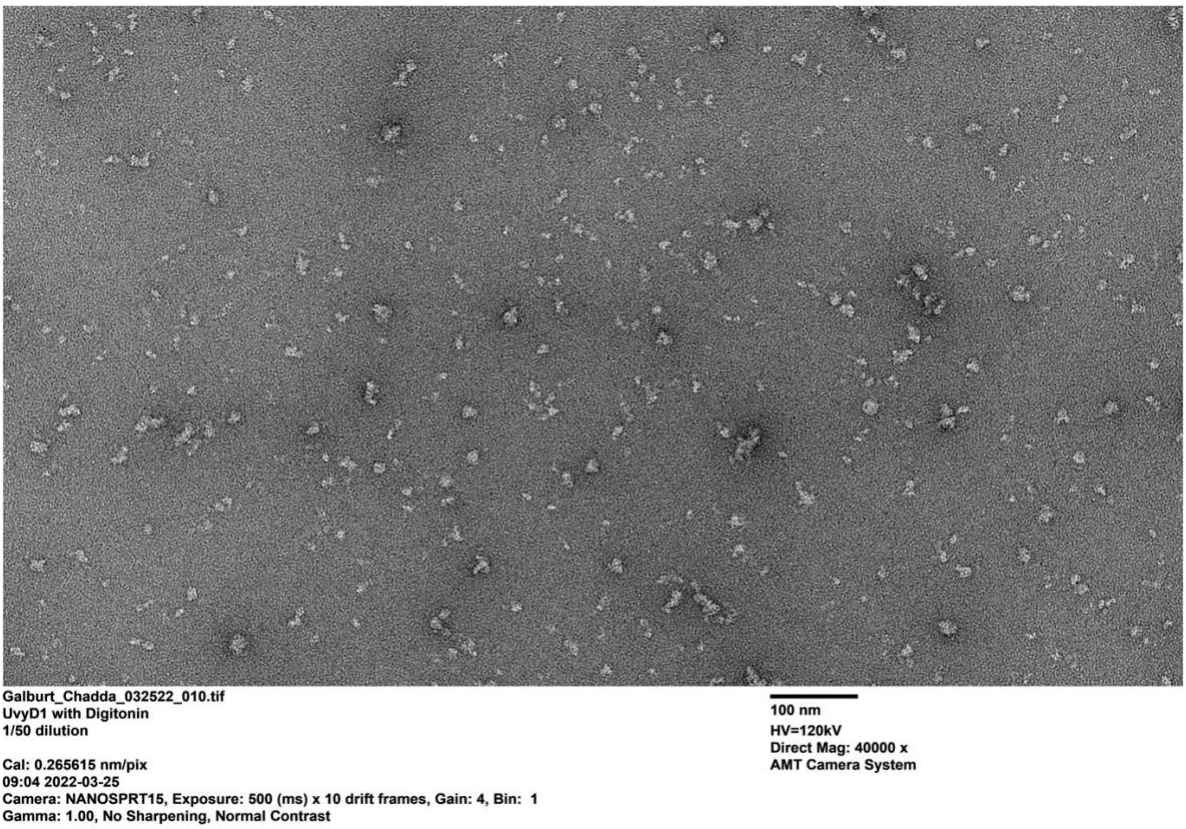


Figure 5.4.1: The negative staining of the *Mtb* UvrD1 dimer collected at 1:50 dilution of UvrD1 in presence of Digitonin detergent.

The images were screened on Glacios 200 keV microscope for presence of any clumps. Since the sample appeared monodisperse, they were imaged on (Krios 300kV) were gain corrected with raw pixel size (Å): 0.7 Å, accelerating voltage (kV): 300 kV, spherical aberration (mm): 0.01 mm, and total exposure dose (e/Å^2): 51.88 e/Å^2.

5.3.1 Image preprocessing

A new project was created by importing 3963 movies (defining the path and the parameters as mentioned above). Patch motion correction was performed to account for sample and stage movement during the exposure before producing an aligned average micrograph shown in (Figure 4B).

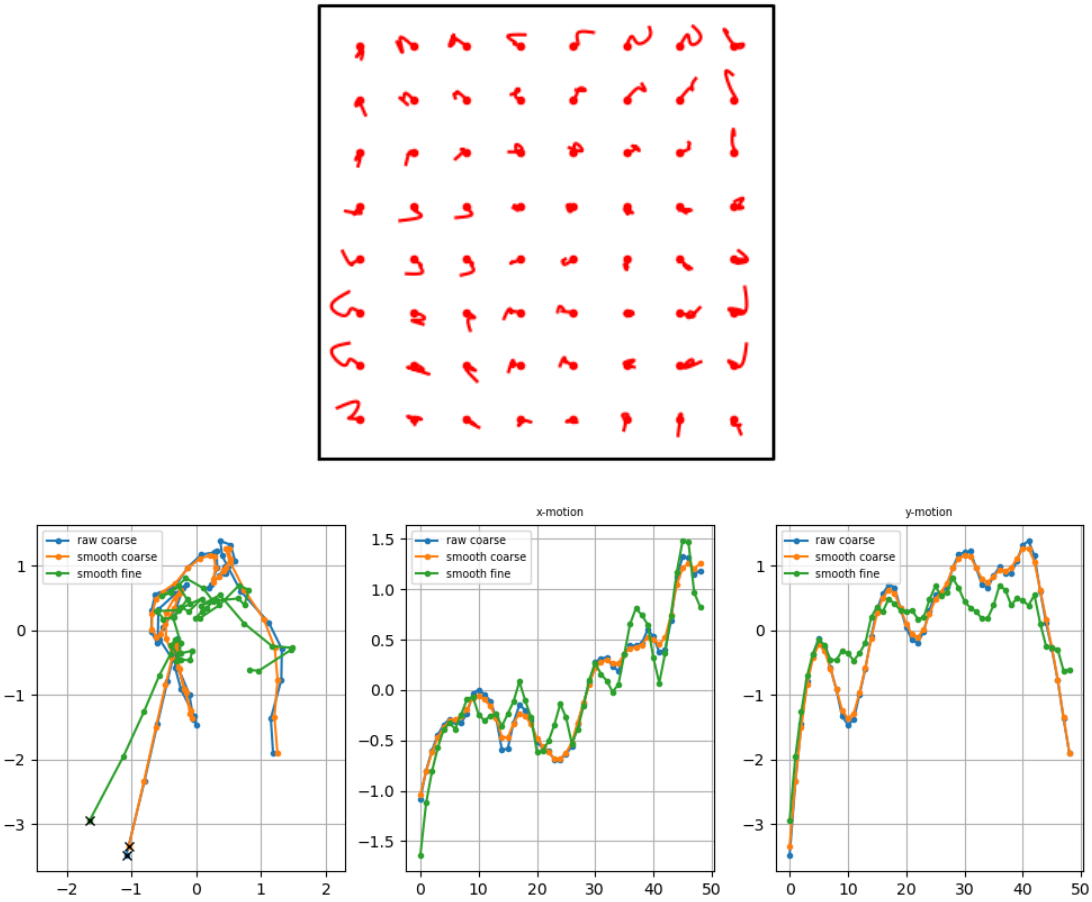


Figure 5.4.2: Patch motion correction showing sample and stage movement of the particles in X and Y axis.

Patch CTF estimation was used from the output of Patch motion correction as shown above to correct for and attempt to measure additional parameters that vary from one micrograph to another like astigmatism, defocus, and estimated resolution, shown in (Figure 4C).

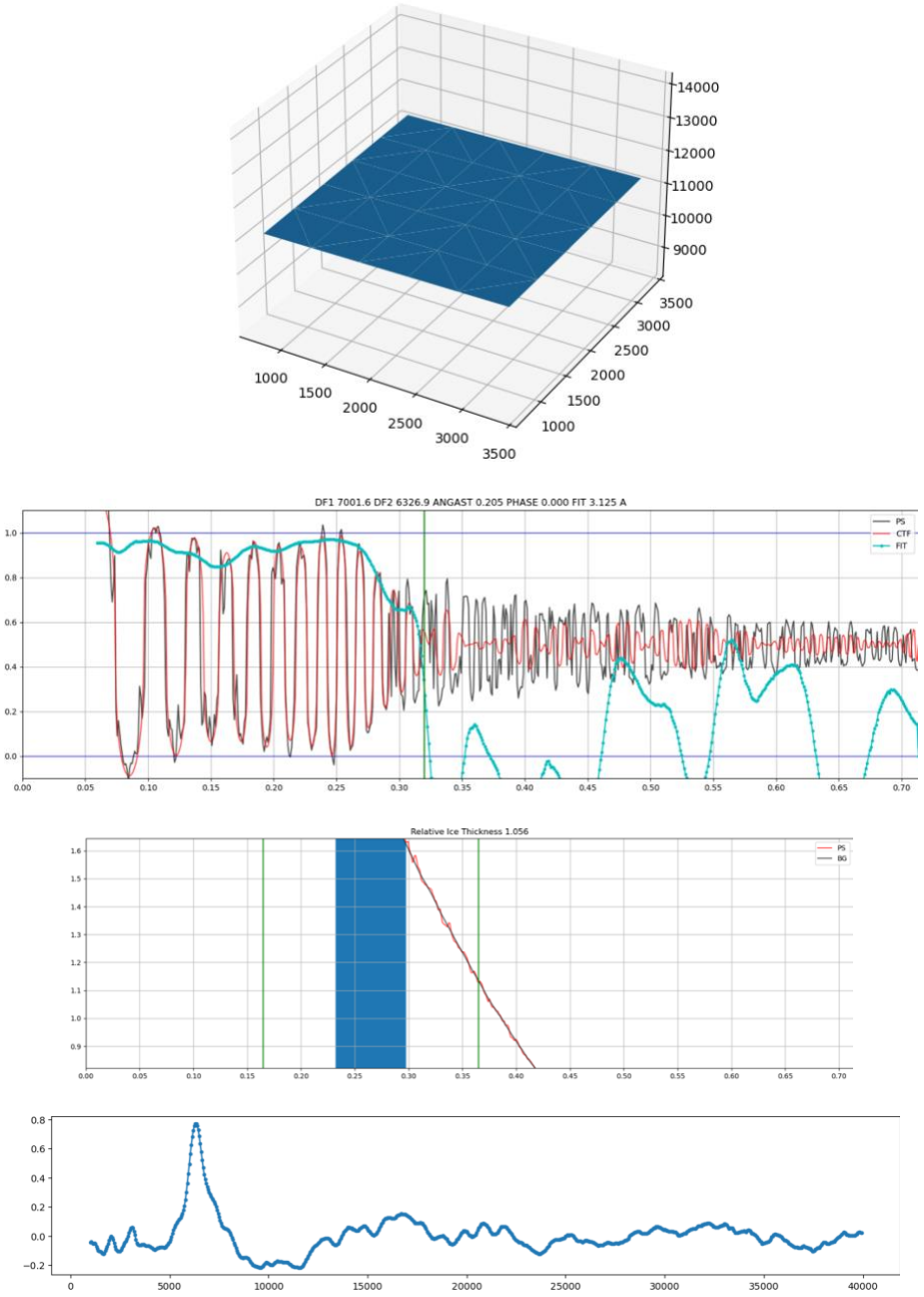


Figure 5.4.3: CTF correction of one of the input micrographs showing relative ice thickness. The X, Y and Z axes units are in Armstrong.

Curate exposures job was used after Patch CTF estimation to remove images with broken or thick ice, carbon edges and too much motion. Out of 3963 movies 769 were discarded because they lacked particles and only 3169 were used for particle picking.

5.3.2 Particle picking and 2D classification

Using blob picker and providing a particle size of 140 Å led to extraction of 658,417 particles from 3169 micrographs and performing 2D classification into 100 classes shows the first round of 2D classification in which classes were fuzzy and showed conformational heterogeneity.

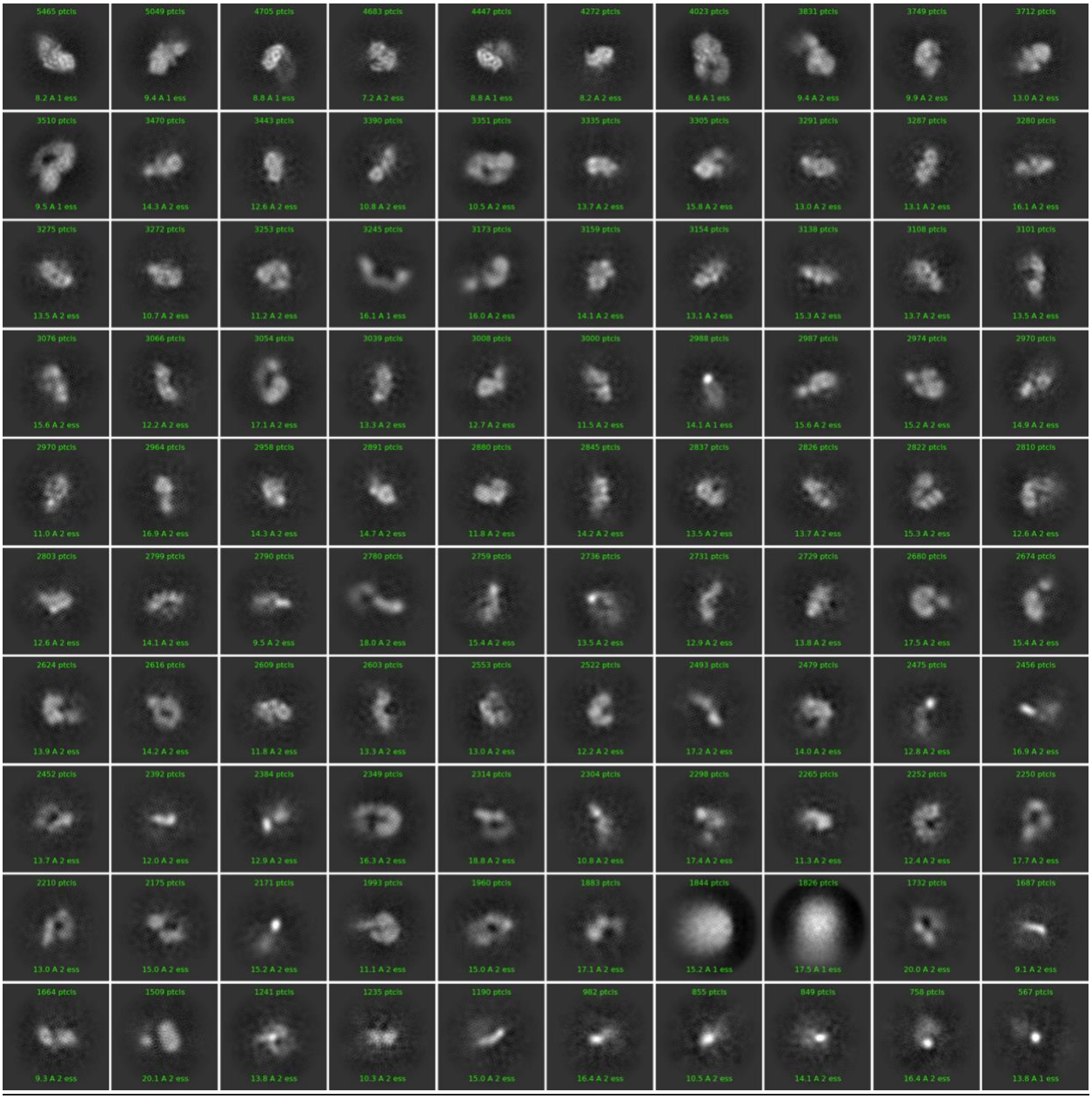


Figure 5.4.4: First 2D classification after extraction of particles using blob picker. The classes appear fuzzy and shows conformational heterogeneity.

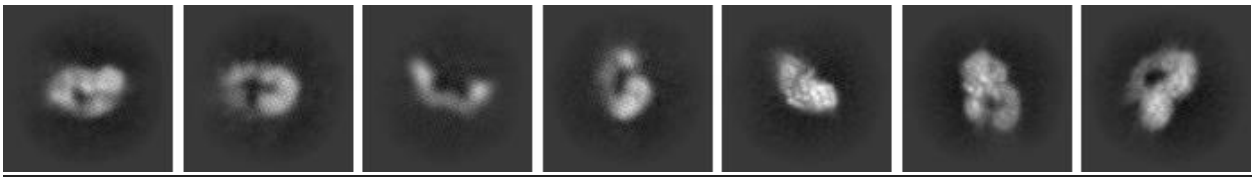


Figure 5.4.5: Templates picked after 2D classification using select 2D job.

After first round of 2D classification, 7 classes were used as templates for template picking job via select 2D job. After picking particles using template picker 819,340 particles were obtained which were next passed through extraction job with parameters of particle extraction box size of 350 Å and Fourier crop size 200 Å. The number of particles obtained from 3179 micrographs and 405,726 particles are passed through second and third round of 2D classification.

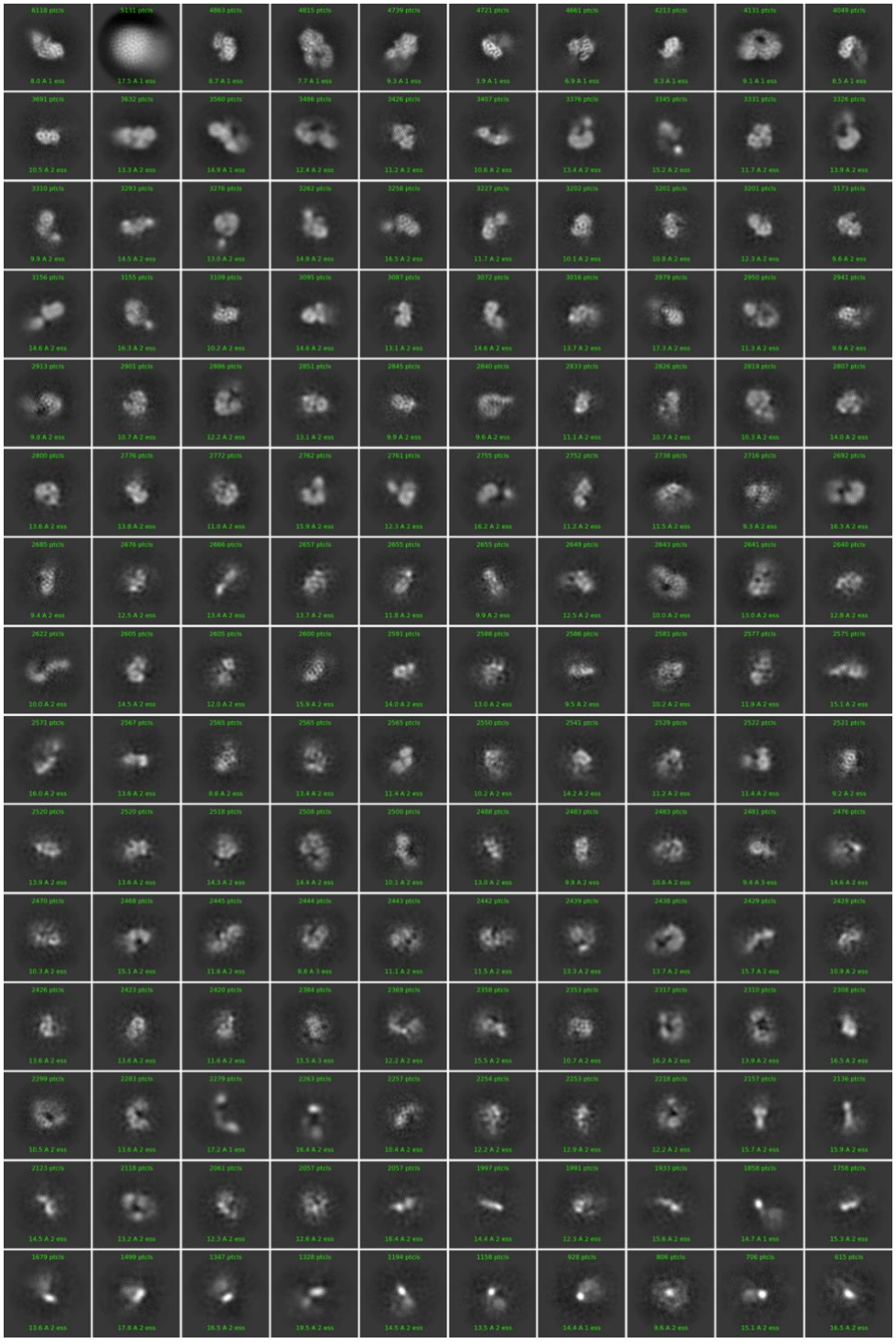


Figure 5.4.6: Third round of 2D classification still shows heterogeneity and most of the classes appear fuzzy.

Since the third round of 2D classification did not improve the resolution, I went ahead with the 2D select job to use 91 classes with 272,960 particles for 3D reconstruction.

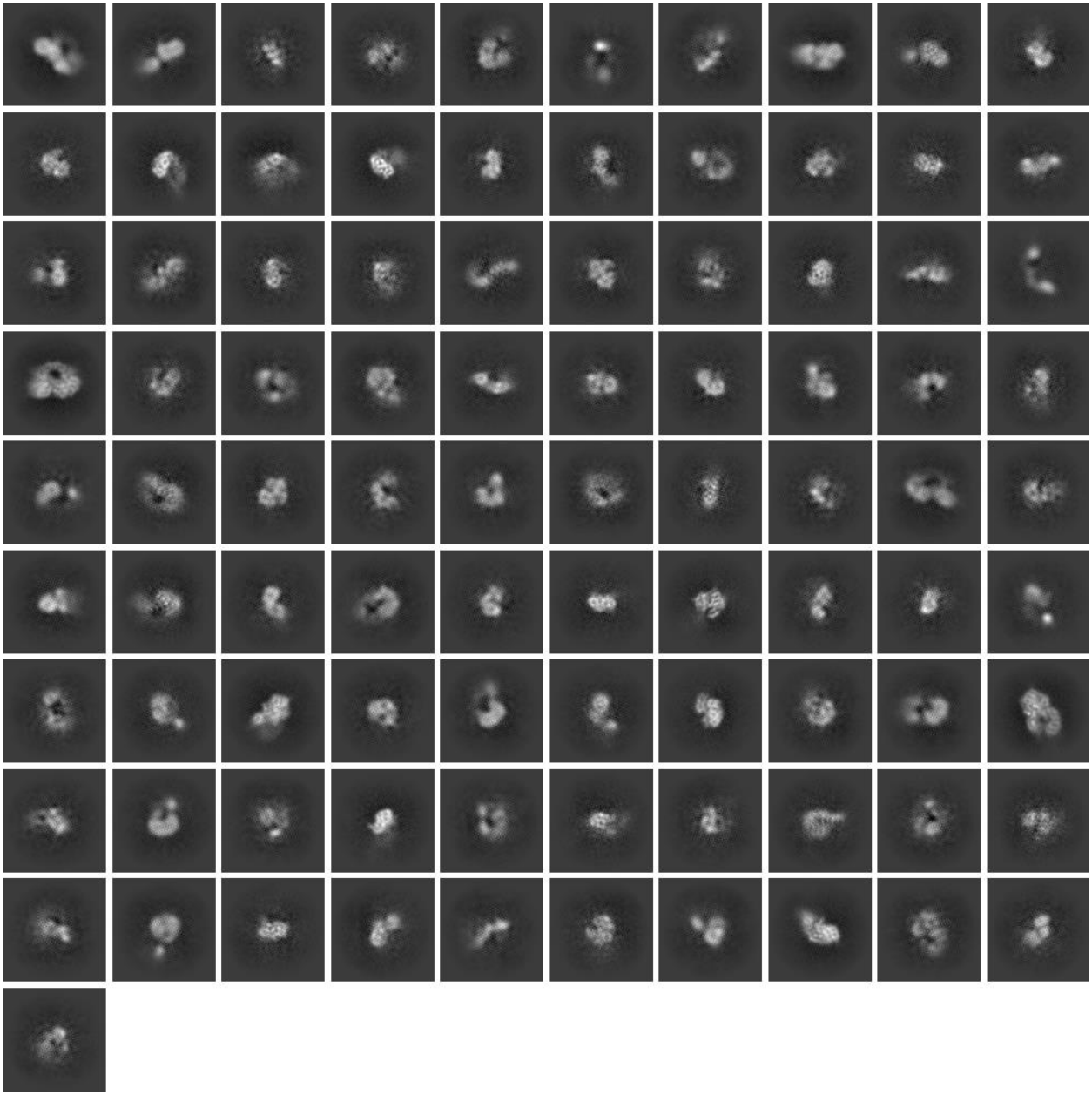


Figure 5.4.7: 2D select job used to select 91 classes with refined features for 3D reconstruction.

5.3.3 3D reconstruction and refinement

As shown above, good classes from select 2D job were used for obtaining a 3D reconstruction from initial particle set of 272,960. The first step of 3D reconstruction being the ab initio modeling where an initial low-resolution 3D map of structure is obtained. I increased the number of classes to 4 to find multiple conformations/states in your data and identify junk particles.

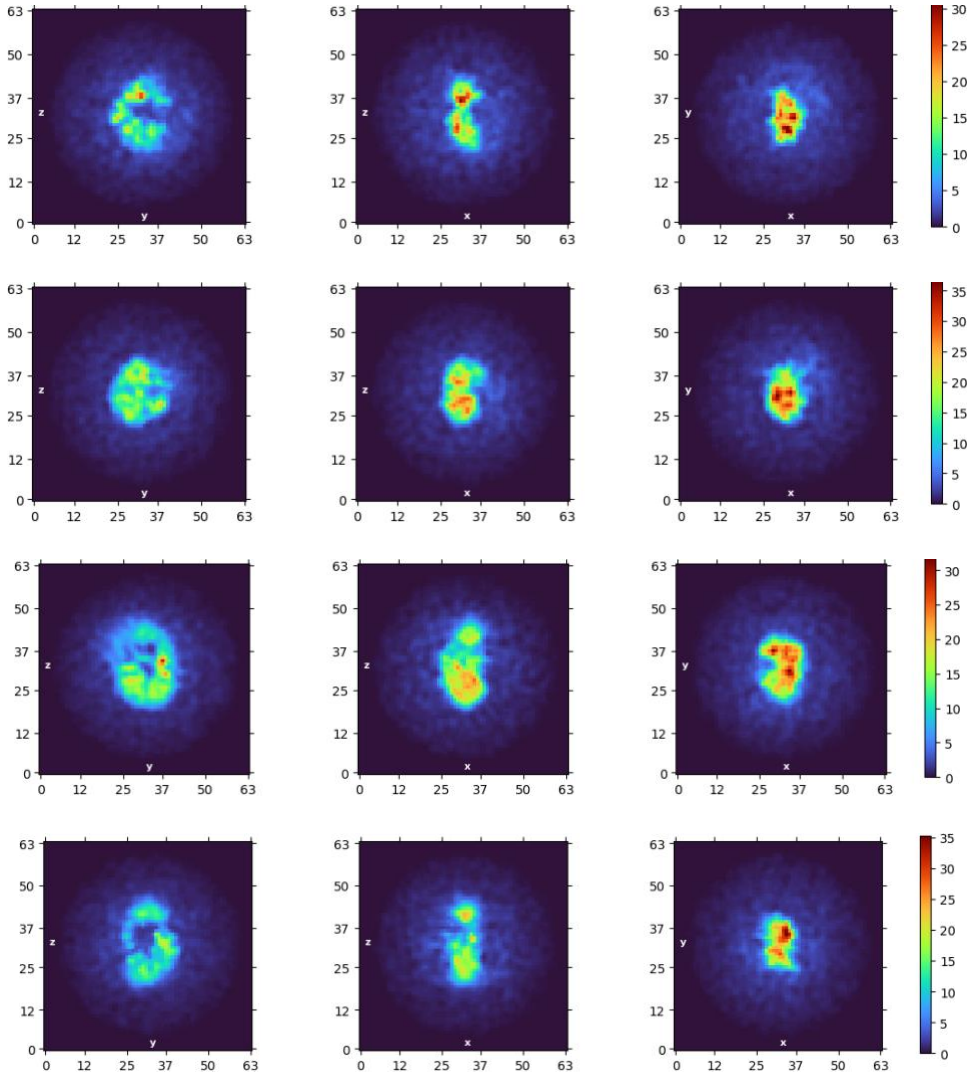
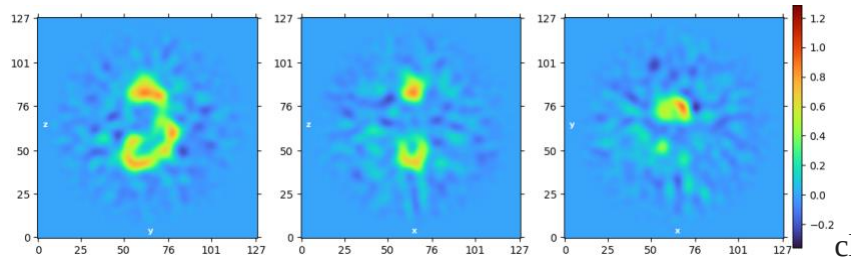
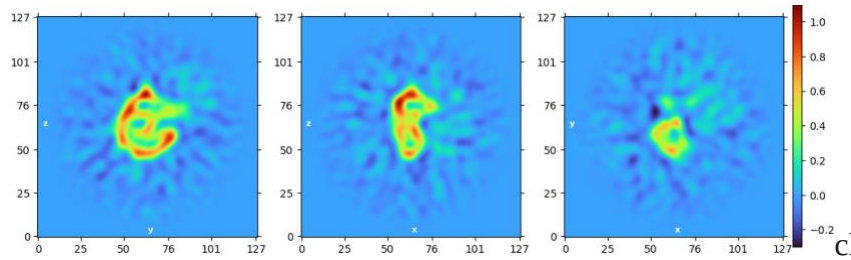
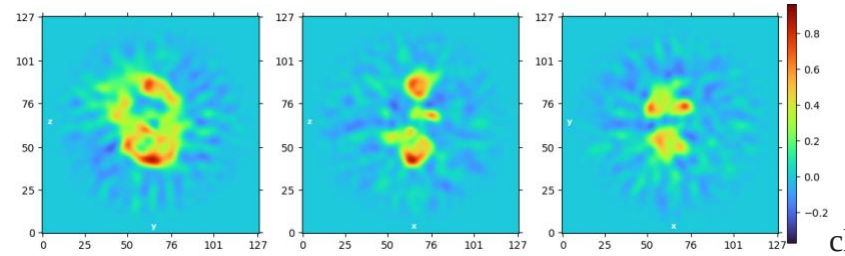


Figure 5.5.1: Model-independent 3D ab-initio reconstruction job that assigns particles for 3D classification where class2 shows dimer structure with one subunit of monomer being flexible.

After passing through initial ab-initio modeling the particles are passed through Heterogenous refinement to sort out junk particles and assign the particles in various groups. Class 2 after heterogeneous refinement showed 91,112 particles shown in (Figure 5B).







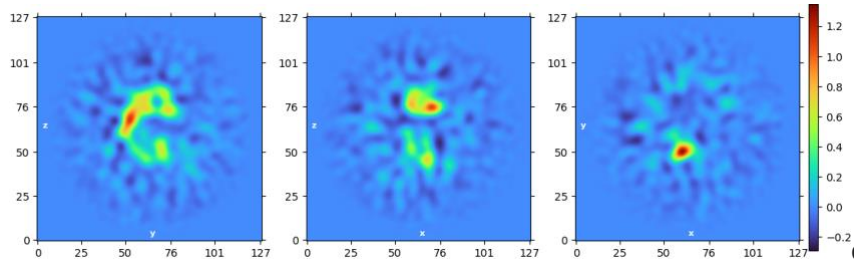
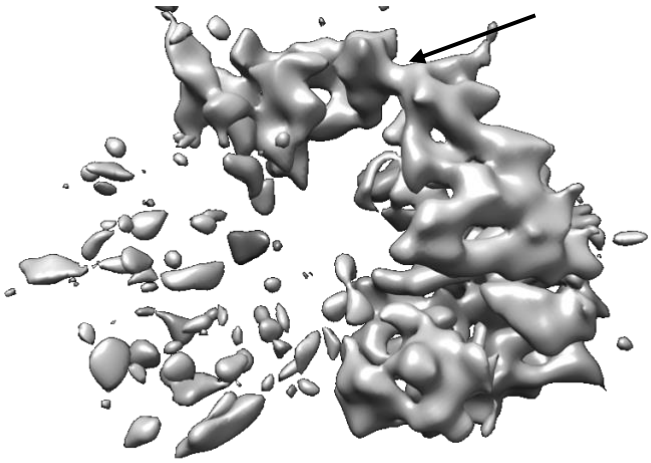
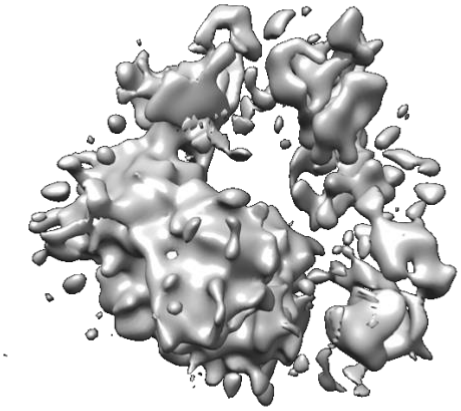


Figure 5.5.2: 3D Heterogenous refinement into four classes that sorted particles and assigned then in respective groups.

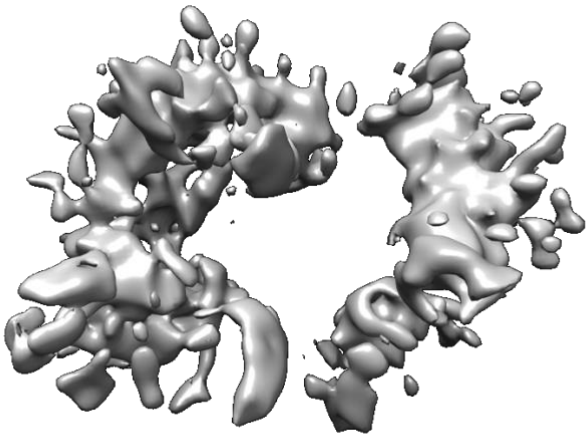
Each of the classes obtained via heterogenous refinement were passed through homogenous refinement to improve resolution. However, likely because of flexibility and conformation heterogeneity, the structures cannot be improved to more than 7-8 Å resolution.



Class I



Class II



Class III

Figure 5.5.3: 3D Homogenous refinement of classes 0, 1 and 2 obtained via heterogenous refinement showing different conformations of UvrD1 dimer where one subunit is not resolved due to flexibility. The arrow in the first image represents the cysteine bond between two monomer subunits.

The presence of conformational heterogeneity prevented a high-resolution structure of the UvrD1 dimer, and the resolution of class I was improved by ~1 Å by passing it through another 3D refinement job called local refinement.

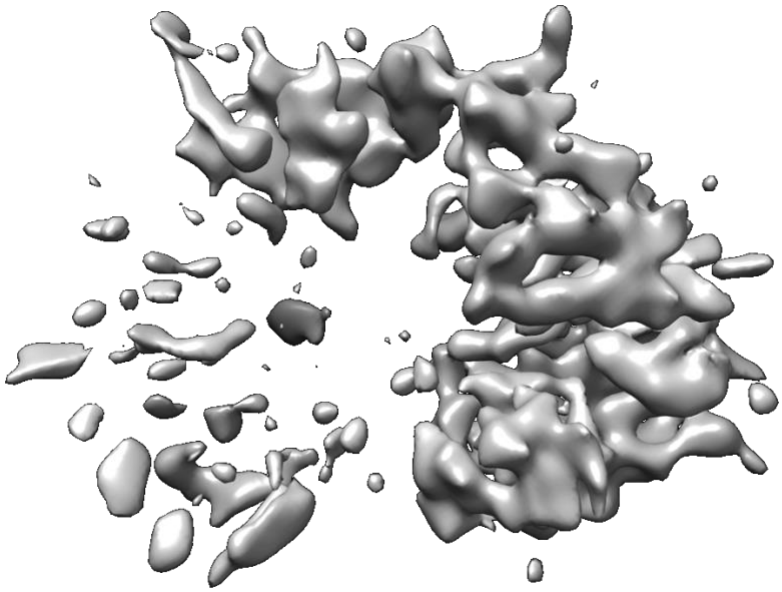


Figure 5.5.4: Local refinement of particles of class 1 that improved overall resolution to 6.6 Å.

5.4 Discussion and future directions

UvrD like helicases belong to SF1 family of helicases that translocate on ssDNA and unwind dsDNA to ssDNA[10]. Previous biochemical and kinetic studies have reported that UvrD helicases translocate on DNA as monomers but unwind DNA as dimers in absence of accessory factors[2,9,20]. However, the structural studies have only revealed UvrD helicase monomers bound to DNA[12,14]. Recently, we reported that the SF1 helicase *Mtb* UvrD1 forms a redox dependent dimer required for DNA unwinding [21]. The goal of this study was to understand the structural basis of DNA unwinding by UvrD1 dimer by obtaining a high-resolution cryo-EM structure of UvrD1 dimer unwinding the DNA. During the first part of this study, I optimized conditions for freezing UvrD1 in presence of crosslinker and observed two different conformations of UvrD1 (open and closed) at a resolution of 5-6 Å. Since UvrD1 becomes inactive after addition of crosslinker, and to try to resolve the full conformational heterogeneity of the dimer, I tried to

optimize conditions to freeze UvrD1 in presence of detergents but in the absence of a crosslinker. Removal of crosslinker results in significant heterogeneity and flexibility across the dimer which results in an inability to resolve any one orientation of one monomer relative to the other. This result is consistent with the ability of the monomers to explore many conformations while not bound to DNA.

Future studies should aim at getting high-resolution structure of UvrD1 dimer in presence of DNA and ATP analog AMPPNP and Mg^{+2}. Protein freezing can be performed in presence of crosslinker to stabilize the dimer conformation on DNA. The conditions for freezing the protein should include excess DNA so that ~100% of the proteins picked on the grid should be bound to DNA. The presence of crosslinker can show non-specific interactions that can be confirmed by making mutations in protein and testing if it affects the binding with DNA. Another strategy would be to use the C-terminal mutant (Delta Tudor) of UvrD1 and freeze with DNA in presence of BS3 crosslinker. One would first purify UvrD1 and perform sizing in the presence of phosphate buffer. One strategy would be to use crosslinker in presence of 18bp-dT20 DNA substrate where only a dimer can bind to DNA. Another way is to use a variety of DNA substrates with mismatch towards the partial duplex region that would represent the intermediates involved in unwinding process. The presence of ATP analog AMPPNP and ADP along with DNA could also be added in the reaction to increase affinity of UvrD1-DNA complexes. As the Delta Tudor UvrD1 has higher affinity for binding to the unwinding DNA substrate, using it for freezing may further promote the formation of UvrD1-DNA complexes (unpublished). To confirm that the closed and open confirmations of UvrD1 seen in case of crosslinked UvrD1 dimer represent true conformations one could also use FRET experiments where the RecA domain of protein may be labeled with either Cy3/Cy5 and single molecule FRET may be used to measure movement of domains.

5.5 References

1. C. Bustamante, D. Keller, G. Oster, The Physics of Molecular Motors †, Accounts Chem Res. 34 (2001) 412–420. https://doi.org/10.1021/ar0001719.

2. T.M. Lohman, N.T. Fazio, How Does a Helicase Unwind DNA? Insights from RecBCD Helicase, Bioessays. 40 (2018) 1800009. https://doi.org/10.1002/bies.201800009.

3. A.A. Mahdi, G.S. Briggs, G.J. Sharples, Q. Wen, R.G. Lloyd, A model for dsDNA translocation revealed by a structural motif common to RecG and Mfd proteins, Embo J. 22 (2003) 724–734. https://doi.org/10.1093/emboj/cdg043.

4. B. Sun, A. Singh, S. Sultana, J.T. Inman, S.S. Patel, M.D. Wang, Helicase promotes replication re-initiation from an RNA transcript, Nat Commun. 9 (2018) 2306. https://doi.org/10.1038/s41467-018-04702-x.

5. V. Epshtein, V. Kamarthapu, K. McGary, V. Svetlov, B. Ueberheide, S. Proshkin, A. Mironov, E. Nudler, UvrD facilitates DNA repair by pulling RNA polymerase backwards, Nature. 505 (2014) 372. https://doi.org/10.1038/nature12928.

6. M.R. Singleton, M.S. Dillingham, D.B. Wigley, Structure and Mechanism of Helicases and Nucleic Acid Translocases, Annu Rev Biochem. 76 (2007) 23–50. https://doi.org/10.1146/annurev.biochem.76.052305.115300.

7.M.R. Webb, C.P. Toseland, Encyclopedia of Biophysics, (2013) 950–958. https://doi.org/10.1007/978-3-642-16712-6_463.

8. H.S. Subramanya, L.E. Bird, J.A. Brannigan, D.B. Wigley, Crystal structure of a DExx box DNA helicase, Nature. 384 (1996) 379–383. https://doi.org/10.1038/384379a0.

9. J.A. Ali, T.M. Lohman, Kinetic Measurement of the Step Size of DNA Unwinding by Escherichia coli UvrD Helicase, Science. 275 (1997) 377–380. https://doi.org/10.1126/science.275.5298.377.

10. T.M. Lohman, E.J. Tomko, C.G. Wu, Non-hexameric DNA helicases and translocases: mechanisms and regulation, Nat Rev Mol Cell Bio. 9 (2008) 391–401. https://doi.org/10.1038/nrm2394.

11. M.S. Dillingham, D.B. Wigley, M.R. Webb, Direct Measurement of Single-Stranded DNA Translocation by PcrA Helicase Using the Fluorescent Base Analogue 2-Aminopurine †, Biochemistry-Us. 41 (2002) 643–651. https://doi.org/10.1021/bi011137k.

12. S.S. Velankar, P. Soultanas, M.S. Dillingham, H.S. Subramanya, D.B. Wigley, Crystal Structures of Complexes of PcrA DNA Helicase with a DNA Substrate Indicate an Inchworm Mechanism, Cell. 97 (1999) 75–84. https://doi.org/10.1016/s0092-8674(00)80716-3.

13. S. Korolev, J. Hsieh, G.H. Gauss, T.M. Lohman, G. Waksman, Major Domain Swiveling Revealed by the Crystal Structures of Complexes of E. coli Rep Helicase Bound to Single-Stranded DNA and ADP, Cell. 90 (1997) 635–647. https://doi.org/10.1016/s0092-8674(00)80525-5.

14. J.Y. Lee, W. Yang, UvrD Helicase Unwinds DNA One Base Pair at a Time by a Two-Part Power Stroke, Cell. 127 (2006) 1349–1360. https://doi.org/10.1016/j.cell.2006.10.049.

15. N.K. Maluf, C.J. Fischer, T.M. Lohman, A Dimer of Escherichia coli UvrD is the Active Form of the Helicase In Vitro, J Mol Biol. 325 (2003) 913–935. https://doi.org/10.1016/s0022-2836(02)01277-9.

16. J. Chase, A. Catalano, A.J. Noble, E.T. Eng, P.D. Olinares, K. Molloy, D. Pakotiprapha, M. Samuels, B. Chait, A. des Georges, D. Jeruzalmi, Mechanisms of opening and closing of the bacterial replicative helicase, Elife. 7 (2018) e41140. https://doi.org/10.7554/elife.41140.

17. F.A. Ali, L. Renault, J. Gannon, H.L. Gahlon, A. Kotecha, J.C. Zhou, D. Rueda, A. Costa, Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate, Nat Commun. 7 (2016) 10708. https://doi.org/10.1038/ncomms10708.

18. A. Chadda, D. Jensen, E.J. Tomko, A.R. Manzano, B. Nguyen, T.M. Lohman, E.A. Galburt, Mycobacterium tuberculosis DNA repair helicase UvrD1 is activated by redox-dependent dimerization via a 2B domain cysteine, Proc National Acad Sci. 119 (2022) e2114501119. https://doi.org/10.1073/pnas.2114501119.

19. H. Jia, S. Korolev, A. Niedziela-Majka, N.K. Maluf, G.H. Gauss, S. Myong, T. Ha, G. Waksman, T.M. Lohman, Rotations of the 2B Sub-domain of E. coli UvrD Helicase/Translocase Coupled to Nucleotide and DNA Binding, J Mol Biol. 411 (2011) 633–648. https://doi.org/10.1016/j.jmb.2011.06.019.

20. N.K. Maluf, C.J. Fischer, T.M. Lohman, A Dimer of Escherichia coli UvrD is the Active Form of the Helicase In Vitro, J Mol Biol. 325 (2003) 913–935. https://doi.org/10.1016/s0022-2836(02)01277-9.

21. A. Chadda, D. Jensen, E.J. Tomko, A.R. Manzano, B. Nguyen, T.M. Lohman, E.A. Galburt, Mycobacterium tuberculosis DNA repair helicase UvrD1 is activated by redox-dependent dimerization via a 2B domain cysteine, Proc National Acad Sci. 119 (2022) e2114501119.

Chapter 6: Interaction studies of *Mtb* UvrD1 with RNA polymerase

6.1 Introduction

In all cellular organisms, RNA is synthesized by a multisubunit DNA-dependent RNA polymerase (RNAP) that consists of two alpha, one beta, one beta prime, and one omega subunit depicted as α2ββ′ω and is evolutionarily conserved in sequence, structure, and function from bacteria to humans [1,2]. The bacterial RNAP has a crab claw shape, the claws comprising the large β and β′ subunits. The two claws define the enzyme's main channel, which accommodates the DNA template and the RNA–DNA hybrid that forms during transcription[1]. The active center, which is marked by a stably bound Mg^{2+} ion, lies at the back of the main channel. A second, smaller channel links the external milieu with the enzyme's active site. Because the DNA largely blocks access to the active center via the main channel during transcription, this secondary channel is the presumed entryway for substrate nucleotides (NTPs) to reach the site of catalysis. The two identical α subunits are located distal to the cleft formed by the claws of Beta and beta prime. One α subunit interacts with β and the other interacts with β'. The ω subunit is also located distal to the cleft, near the base, and interacts exclusively with β prime. This multisubunit bacterial RNAP is called core enzyme and is catalytically proficient but is unable to initiate transcription from promoters. Binding of the sigma (σ) factor to the RNAP core results in the formation of the holoenzyme which can recognize and initiate transcription from promoters[3]. The process of transcription consists of three distinct phases called initiation, elongation, and termination. During initiation the RNAP holoenzyme first binds to a duplex DNA promoter sequence to form the

RNAP promoter closed complex (RPc)[4,5]. This binding drives multiple conformational changes in both the RNAP and DNA that lead to the isomerization of RPc to RPo (RNAP open complex), where the duplex DNA is unwound around the transcription start site, resulting in a ∼13 base-pair (bp) single-stranded DNA bubble[5,6]. After binding of the initiating nucleotide in the active site of RNAP, RNA polymerization begins as nucleotides complementary to the template DNA strand are incorporated into a growing RNA transcript resulting in an initially transcribing complex (RP itc). Once 9–15 nucleotides are incorporated into the RNA transcript, upstream contacts are broken, the σ factor may dissociate, the transcription bubble shrinks, and RNAP fully escapes the promoter[7,8]. This marks the end of initiation and the beginning of processive elongation (RPe) [9,10]. The dissociation of the σ factor allows the core RNA polymerase enzyme to proceed along with the DNA template, synthesizing mRNA in the 5′ to 3′ direction[11]. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it. Once a gene is transcribed, the RNA polymerase dissociates from the DNA template and liberates the newly synthesized mRNA the process called termination [12]. Transcription occurs at the average rate of 50 nucleotides/sec but various factors can bind RNAP and regulate the process of transcription. For instance, a transcription factor called CarD in Mycobacterium tuberculosis (*Mtb*) has been shown to bind RNAP holoenzyme during initiation and helicases such as UvrD/Mfd during elongation when RNAP is stalled at the site of lesion[13–17] .

CarD is an essential transcription factor in *Mtb* and has been shown to stabilize the RNA polymerase open complex during the initiation step of transcription[16,18]. Whereas Mutation frequency decline (Mfd) is a superfamily 2 helicase protein that interacts with RNAP when it's stalled at a lesion during the elongation step of transcription. Mfd uses its translocase activity to push RNAP forward from the lesion site[19]. Both CarD and Mfd have Tudor like domains

through which they interact with RNAP β1 lobe[14]. In contrast to Mfd that pushes RNAP forward another helicase with Tudor like domain called UvrD (DNA Helicase II) has been shown to pull stalled RNAP backward[13]. *Mtb* has two UvrD's (namely UvrD1 and UvrD2) that are involved in DNA repair [20–22]. The C-terminal domain of *Mtb* UvrD1 also adopts a Tudor-like fold suggesting that CarD, UvrD1 and Mfd likely interact with RNAP on the same site at the beta-1 lobe of RNAP[23,24].

Since accessory proteins such as CarD, UvrD1 and Mfd bind to the same site on RNAP and carry out diverse functions during transcription, one question is how do multiple accessory proteins compete for binding to same site on RNAP? My hypothesis is that the change in conformation of RNAP during initiation, elongation and stalled phases of transcription drives the binding of various transcription factors to RNAP β1 lobe.

To test this hypothesis, I would measure binding affinity of RNAP complex with CarD and UvrD1 during initiation, elongation, and stalled phases of transcription. If the hypothesis is true then CarD would have more affinity for initiation complexes where it has been shown to stabilize RNAP promoter open complexes, and UvrD1 would have more affinity for the elongation/ stalled phases where it pulls RNAP backward stalled at the lesion. As controls, I would use mutants of CarD and UvrD1 lacking the Tudor domain that cannot interact with RNAP complexes. The first set of experiments that I performed to test this hypothesis is to determine if CarD and UvrD1 can interact with RNAP holoenzyme.

6.2 Results

6.2.1 Mtb UvrD1 interacts with RNAP in absence of DNA

During the process of elongation, RNAP can get stalled at the site of DNA lesions[17,25,26]. Previous studies have shown that *E. coli* UvrD helicase interacts with RNAP stalled at the site of a lesion[13]. It can then use its helicase or translocase activity to pull RNAP backward so repair enzymes can access and repair the lesion[27]. In addition, UvrD and *B. steraothermophilus* PcrA have been shown to interact with RNAP via the C-terminal Tudor domain[28]. *Mtb* has two UvrDs, UvrD1 and UvrD2 where UvrD1 shares a conserved C-terminal domain Tudor domain with *E. coli* UvrD and PcrA[29]. To test the hypothesis that *Mtb* UvrD1 backtracks RNAP and plays a role in DNA repair, I first examined if UvrD1 interacts with RNAP holoenzyme in the absence of DNA. I incubated *Mtb* UvrD1 with *Mtb* RNAP holoenzyme for 30 minutes in a tube with final buffer conditions of TRIS pH 8.0, 50mM NaCl and 10% glycerol, and loaded the reactions on the 3% native gel where UvrD1 or RNAP alone appears as a single peak shown in lanes 1 and 2 and the addition of increasing concentration of RNAP leads to the disappearance of free UvrD1 peak suggesting its incorporation in RNAP to form a complex.

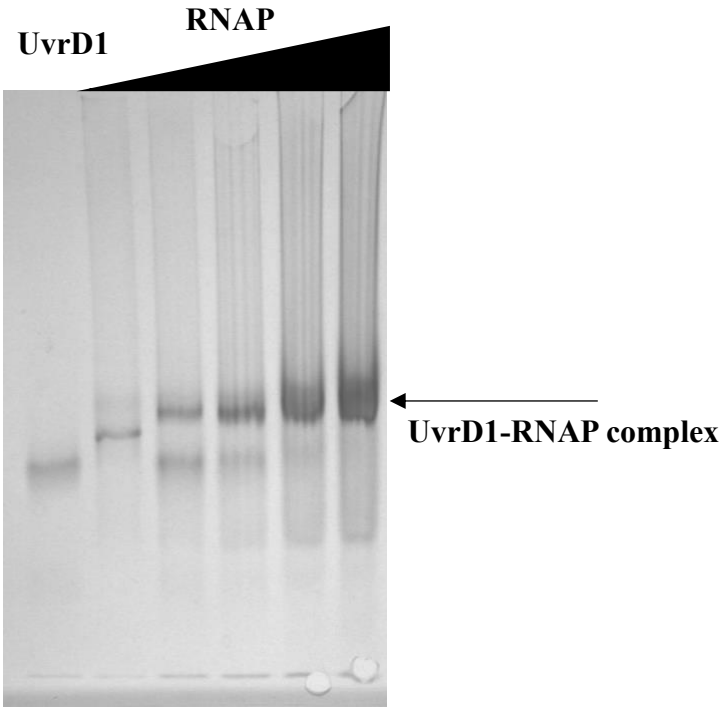


Figure 6.1**:** *Mtb* UvrD1 binds to RNAP in absence of DNA. Lane 1 from left shows 5µl of 5µM UvrD1, lane 2 is 1µl of 6µM RNAP. Titrating an increasing amount of RNAP from lane 3-6 shows the incorporation of free UvrD1 to RNAP suggesting an interaction between the two in absence of DNA. In the last lane 5µl of 5µM UvrD1 is fully incorporated with 7µl of 6µM RNAP.

In a native gel assay shown above it is difficult to determine if UvrD1 is interacting with an RNAP holoenzyme or if it's replacing any one of the subunits of RNAP since a multisubunit enzyme with two alpha, one beta, one beta prime, and one omega subunit depicted as $\alpha 2\beta\beta'\omega$ and binding with various sigma factors makes it a holoenzyme[30]. To determine if *Mtb* UvrD1 can bind in the presence of all subunits of RNAP holoenzyme, I used an alternative pull-down approach to test if UvrD1 interacts with multisubunit RNAP. The *Mtb* UvrD1 is purified using SUMO-his tag that is cleaved in a later step of purification[31]. However, the RNAP holoenzyme is purified with a 10X his-tag at the N-terminus. I performed a pull-down experiment where his-tagged RNAP holoenzyme (6µM) is incubated with UvrD1 (10µM) and after several washes with 10mM and 20mM imidazole containing buffer RNAP/UvrD1 or complex is eluted using a buffer containing 400mM imidazole with TRIS pH 8.0, 50mM NaCl and 10% glycerol. Since UvrD1 alone cannot bind to Ni-NTA column no band is seen in Figure 6.2, lane 4 starting from the left. The lane with RNAP holoenzyme RNAP that has his tag contains all subunits as shown in lane 7. The lane 5 that has both RNAP and UvrD1 shown all subunits of RNAP holoenzyme along with UvrD1 suggesting that UvrD1 interacts with RNAP with all subunits of RNAP.

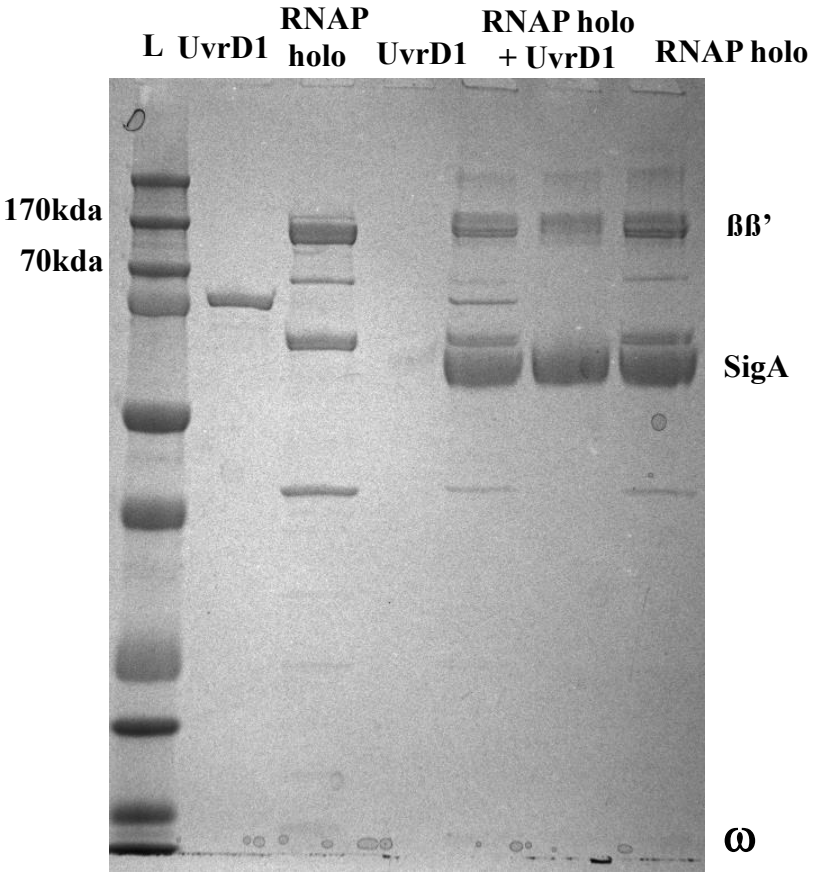


Figure 6.2: *Mtb* UvrD1 pull down assay with RNAP SigA holoenzyme. Lane 1 is ladder and lane 2 and 3 are UvrD1(10µM) and RNAP holoenzyme from storage (6µM). lane 4 is pull down of UvrD1 alone, lane 5 is pull down of UvrD1 with RNAP holoenzyme, and lane 7 is pull down of RNAP holoenzyme showing all subunits of RNAP. Since RNAP holoenzyme is his tagged UvrD1 bound to RNAP is pulled down with it.

6.2.2 CarD can bind with RNAP in the absence of DNA

CarD is a transcription factor that interacts with RNAP holoenzyme via its N-terminal Tudor domain[14]. Both UvrD helicase and CarD have Tudor-like domains and UvrD has been shown to interact with RNAP via C-terminal Tudor domains[23,32]. To determine the interaction of CarD with RNAP holoenzyme in the absence of DNA, I incubated CarD and RNAP holoenzyme and loaded the reactions on a native gel assay. As shown in the assay in Figure 6.3 when 5ul of 15µM CarD was incubated alone it appears as a diffuse band shown in lane 1 and 1ul of 6µM RNAP holoenzyme appears as a single band as shown in lane 8 and titrating in 6µM of RNAP shown in lane 2-5 leads to the formation of a shifted RNAP holoenzyme suggesting the formation

of RNAP holo-CarD complex. The concentration of RNAP holoenzyme shown in Figure 6.1 and 6.3 is the same however CarD shown in Figure 6.3 is 3X more than UvrD1 as lower concentrations of CarD were not visible on native gel. These results suggest that like UvrD1, CarD might also interact with RNAP holoenzyme in the absence of DNA.

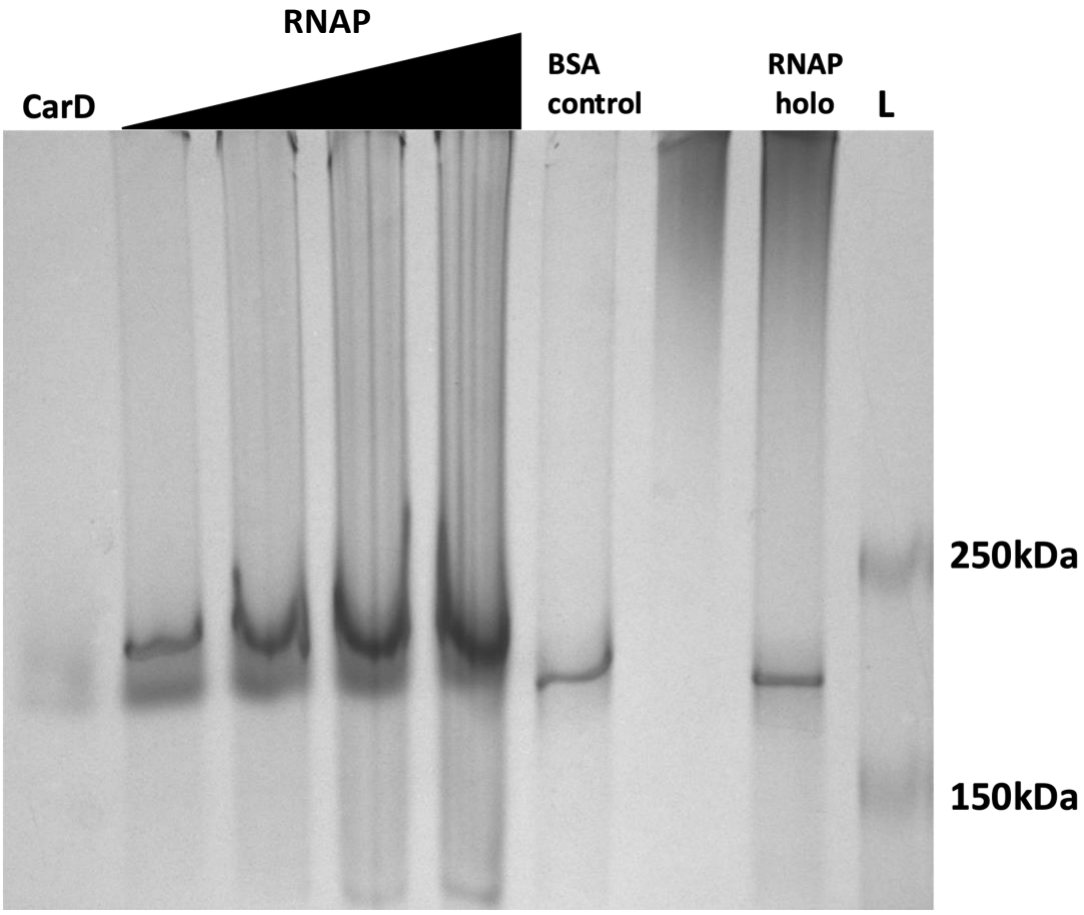


Figure 6.3: CarD can bind to RNAP in the absence of DNA. Lane 1 shows 5µl of 15µM CarD alone and lane 2-5 shows the titration of increasing concentration of RNAP with CarD such that the last lane contains 5µl of 15µM CarD with 7ul of 6µM RNAP and the formation of CarD-RNAP holoenzyme complex. Lane 6 is BSA alone as control and lane 8 is 1ul of 6µM RNAP holoenzyme.

6.2.3 Both *Mtb* UvrD1 and CarD can bind to RNAP holoenzyme in the absence of DNA

Previous studies have shown that both *Mtb* CarD and E. coli UvrD have Tudor domains through which they interact with RNAP. This suggests that both CarD and UvrD compete for binding to RNAP. Using native gel assays shown in Figures 6.1 and 6.3, I have shown that both *Mtb* CarD and *Mtb* UvrD1 can interact with RNAP holoenzyme in the absence of DNA. To determine if they both compete in binding to RNAP, I performed a Ni-NTA pull-down assay where RNAP holoenzyme is his-tagged (6µM) however his tags have been removed from both UvrD1 and CarD after purification. As shown in the gel in Figure 6.4 the pull-down fractions are shown from lane 4-9 from the left. Lane 5 shows pull down of RNAP with UvrD1 which shows bands corresponding to the molecular weight of RNAP holoenzyme subunits and UvrD1 band (see UvrD1 band from lane 4 for comparison). Lanes 6 and 7 do not show any band as UvrD1 and CarD both lack his tags and are not eluted in buffer containing imidazole. Lanes 8 and 9 show RNAP pull down with UvrD1 (10µM) or CarD (11µM) suggesting both can bind to RNAP. However, the last lane shows a pull-down with RNAP bound to both CarD and UvrD1 showing bands corresponding to *Mtb* UvrD1 and CarD both bound to RNAP holoenzyme. Since this is a pull-down assay and both CarD and UvrD1 can bind to RNAP this assay is not suited to answer the question that if CarD and UvrD1 compete for binding to same site on RNAP. One way to perform this assay would be to use fluorescently labeled UvrD1 and/or CarD and test their binding with initiation and elongation complexes. According to the model CarD will bind with higher affinity to initiation complexes and UvrD1 will bind to higher affinity with elongation or stalled complexes. If they compete for binding to RNAP then addition of unlabeled form of either one of those after formation of complex will lead to reduction in fluorescence intensity suggesting competition.

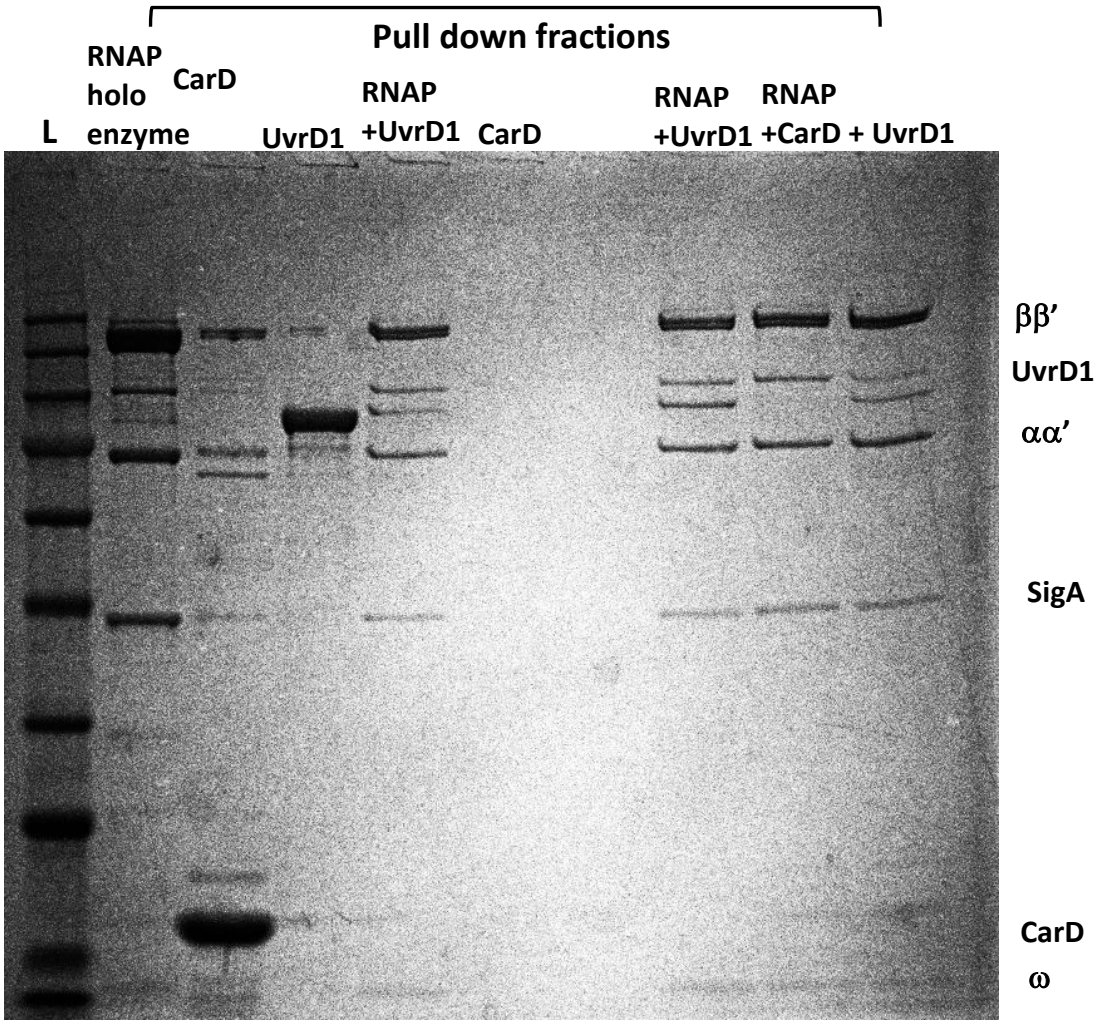


Figure 6.4: RNAP holoenzyme can bind both *Mtb* CarD and UvrD1 in the absence of DNA. Lane 1 shows a protein ladder, and lane 2-4 are stocks of RNAP holoenzyme (6uM), CarD (10µM), and UvrD1 (11µM). Pull-down assays using Ni-NTA resin show no bands in UvrD1 or CarD alone and the incubation of his-tagged RNAP holoenzyme along with CarD and Uvrd1 leads to pull down of complex containing RNAP holoenzyme bound to CarD or UvrD1 or both.

6.2.4 Open complex assay suggests cooperativity between *Mtb* UvrD1 and CarD bound to RNAP complex

CarD is a global transcriptional regulator in mycobacteria and has been shown to interact with RNAP holoenzyme functions by interacting directly with RNA polymerase (RNAP) holoenzyme-bound promoter complexes[15]. I used a fluorescently labeled promoter DNA with *Mtb* ribosomal promoter *rrnA*P3 permitted to monitor the events of the formation of promoter

opening during open complex formation before transcription initiation in real time[16]. In addition, I monitored this assay in presence of UvrD1 to determine if *Mtb* UvrD1 competes with CarD bound to open complex. In the stopped-flow assay described before, *Mtb* RNAP holoenzyme was mixed with CarD or an equivalent volume of transcription factor storage buffer such that the proteins were initially at twice the desired final reaction concentrations. Promoter DNA was also prepared at twice the desired final reaction concentration. Equal volume mixing was performed in a stopped-flow apparatus (Applied Photophysics SX-20, total shot volume 150 μl, dead time < 2 ms), so the initial protein and DNA solutions were each diluted by half to reach their final reaction concentrations[16]. All experiments were performed using a final *Mtb* RNAP holoenzyme concentration of 20 nM and a DNA concentration of 2 nM. I modified and performed this assay using a fluorimeter where the endpoints after 30 minutes of the reaction are shown instead of following the kinetics of open complex formation over 30 minutes at 25 degrees. Here when fluorescently labeled promoter DNA was added in one tube and the other tube consisted of RNAP with CarD or UvD1 or both at the concentrations mentioned below. The reaction was started by mixing the contents of both eppendorf and scans are collected from the wavelength range of 550-620nm. The data represents average fluorescence collected in the wavelength range that has been subtracted from the DNA alone signal and is represented in the form of a bar graph. Mixing *Mtb* RNAP holo with *rrnA*P3 promoter DNA resulted in a rapid increase in the fluorescence signal as the duplex DNA was unwound around the position of the Cy3 fluorophore, indicative of open complex formation[16]. The addition of CarD along with RNAP holoenzyme resulted in a further increase in fluorescence suggesting the stabilization of open complexes as shown before. However, the addition of *Mtb* UvrD1 after 20 minutes of adding CarD bound to

RNAP holoenzyme resulted in a further increase in fluorescence suggesting that both UvrD1 and CarD can bind to RNAP holoenzyme in presence of DNA (Figure 6.5).

This preliminary result suggests that both CarD and UvrD1 stabilize the open complex by cooperative interactions.

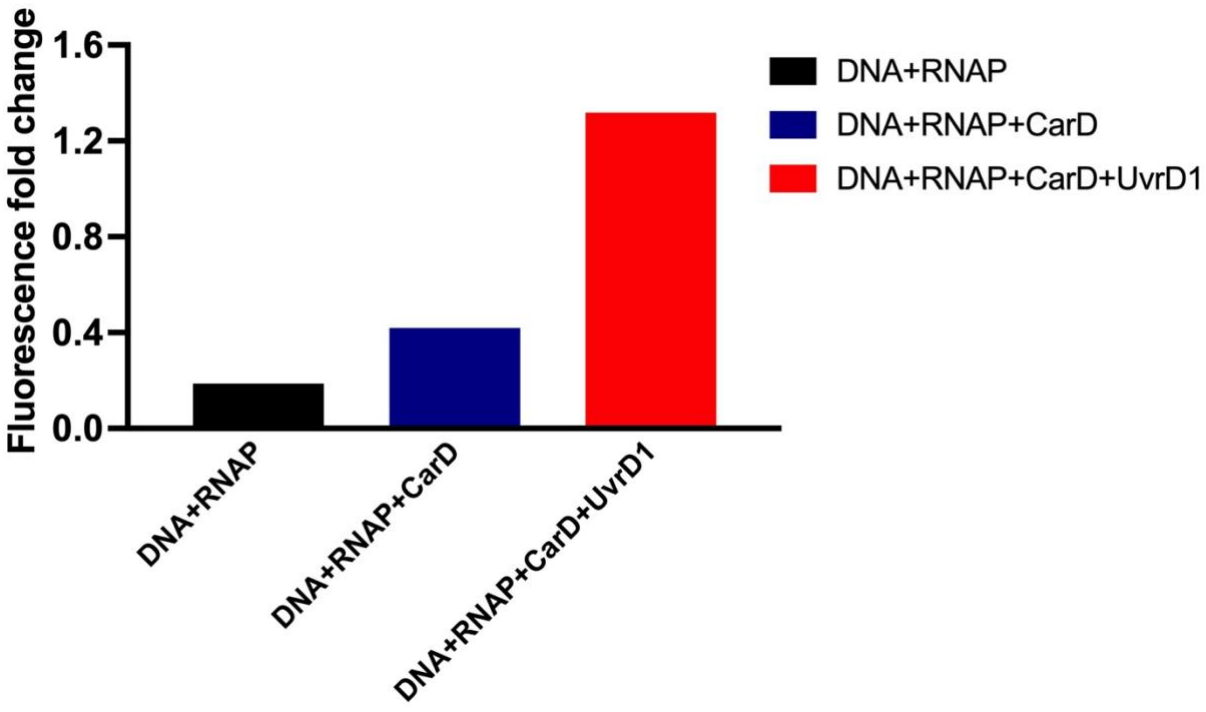


Figure 6.5.1: Increase in fluorescence upon addition of UvrD1 after CarD to open complex suggests cooperativity between CarD and UvrD1. The bar graph represents fold change in fluorescence values that have been subtracted from DNA alone signal. The plot in black is DNA with RNAP holoenzyme and plot in blue represents the addition of CarD (20nM) that leads to an increase in fluorescence signal and the plot in red shows the addition of UvrD1 (1000nM) after adding CarD suggesting cooperativity. The values represent end-point measurements after 30 minutes.

To determine if the increase in the fluorescence in Figure 6.5a is due to *Mtb* UvrD1 unwinding the DNA this assay was performed by adding fluorescently labeled DNA with *Mtb* UvrD1 and monitoring the fold change after 30 minutes. As shown in plot 6.5b the addition of *Mtb* UvrD1 does lead to a change in fluorescence but since the reactions were performed in the absence of ATP this change could indicate the interaction of UvrD1 with fluorophore or DNA melting in the absence of ATP as shown for other helicases[33]. Since this change does not represent a 1.4-fold increase in fluorescence as observed in the presence of CarD and RNAP suggests the cooperative interactions between UvrD1 and CarD stabilize the open complex.

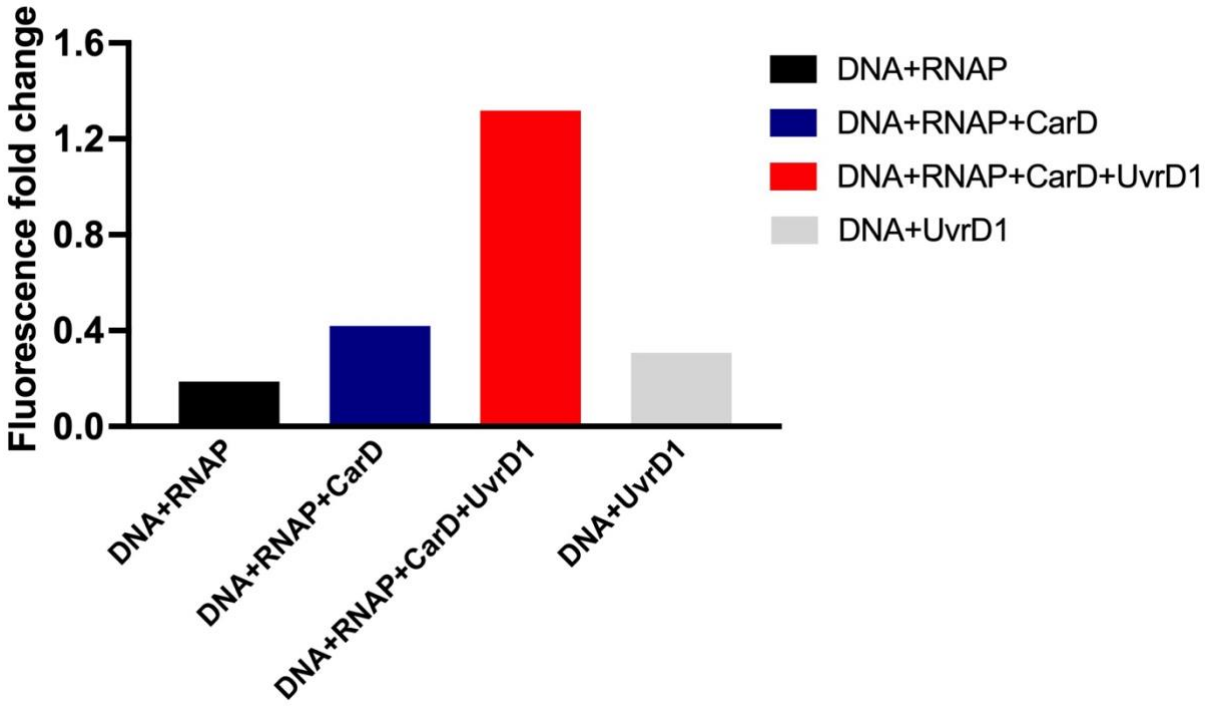


Figure 6.5.2: Increase in fluorescence upon addition of UvrD1 suggests the interaction of UvrD1 with fluorescently labeled DNA. The bar graphs shown in black, blue, and red represent DNA with RNAP or the addition of CarD and UvrD1 and the bar in grey shows fold change in fluorescence when UvrD1 is added to fluorescently labeled DNA suggesting binding or unwinding of promoter DNA.

6.3 Discussion and future directions

UvrD-like helicases belong to the SF1 family of helicases and use energy from ATP hydrolysis to translocate on single-stranded DNA and unwind double-stranded DNA to ssDNA strands[34]. UvrD can unwind DNA during nucleotide excision repair pathway and has been shown to backtrack RNAP stalled at a lesion suggesting a role in transcription-coupled repair (TCR)[13]. CarD a transcription factor in *Mtb* has been shown to bind RNAP and stabilize open complex formation during transcription initiation[16,30]. Both *Mtb* CarD and UvrD like helicases have Tudor domains through which they interact with RNAP. The goal of this study was to test the hypothesis that a change in conformation of RNAP drives the binding of CarD and UvrD1 to same on RNAP via Tudor domains. Since CarD has been shown to stabilize open complex during transcription it can bind to RNAP during the initiation step of transcription while UvrD1 can bind during elongation when RNAP is stalled at a lesion. The preliminary results from the native gel assay shown in Figure 6.1 suggest that *Mtb* UvrD1 interacts with RNAP in the absence of DNA. This interaction does not lead to the removal of any of the subunits of RNAP as shown in Figure 6.2 suggesting that *Mtb* UvrD1 can interact with an RNAP SigA holoenzyme. During characterization of *Mtb* UvrD1, I found that it exists as a monomer or dimer depending on the redox potential, therefore one future question will be to determine if UvrD1 monomer or dimer interacts with RNAP. Since the redox-dependent dimer of UvrD1 can be converted to a monomer in presence of a reducing agent or by mutation of a cysteine at position 451 to alanine one can use Fluorescence anisotropy assays with labeled UvrD1 or DNA in the presence of an initiation or an elongation complex to determine the affinity of a UvrD1 monomer or a dimer to RNAP. Previous studies have shown that CarD interacts with rrnAP3 promoter DNA to stabilize RNAP open complex[16]. The results in Figure 6.3 suggest that CarD can interact with RNAP in the absence

of DNA and based on the binding assay shown in Figure 6.4 both CarD and UvrD1 can bind to RNAP but if they compete for binding to the same site on RNAP is not clear from these experiments. To determine if CarD and UvrD1 compete for binding to RNAP one method is anisotropy where fluorescently labeled DNA, CarD and quantitatively fluorescently bound CarD-RNAP complex can be titrated with unlabeled UvrD1 to test if UvrD1 displaces CarD bound to the RNAP complex suggesting they compete with binding to RNAP. The open complex assay performed previously using stopped flow-based has shown that binding of CarD to RNAP complex led to an increase in fluorescence with time suggesting stabilization of open complex[16]. As shown in Figures 6.5a and 6.5b, I performed this assay in a fluorimeter and recorded final time points that are plotted in the form of a bar graph. These results suggest that the addition of UvrD1 after adding CarD further increases the fluorescence signal indication of stabilization of open complex showing that both CarD and UvrD1 act cooperatively to stabilize the open complex as has been shown previously for two factors of *Mtb* CarD and RbpA[16]. Since *Mtb* UvrD1 exists as a mixture of monomers and dimers according to redox status future experiments should be performed in the presence and absence of reducing agents and the kinetics of open complex assay should be monitored using stopped-flow kinetics.

6.4 References

1. G. Zhang, E.A. Campbell, L. Minakhin, C. Richter, K. Severinov, S.A. Darst, Crystal Structure of Thermus aquaticus Core RNA Polymerase at 3.3 Å Resolution, Cell. 98 (1999) 811–824. https://doi.org/10.1016/s0092-8674(00)81515-9.

2. R. Landick, A. Krek, M.S. Glickman, N.D. Socci, C.L. Stallings, Genome-wide mapping of the distribution of CarD, RNAP σA, and RNAP β on the Mycobacterium smegmatis chromosome using chromatin immunoprecipitation sequencing, Genom Data. 2 (2014) 110–113. https://doi.org/10.1016/j.gdata.2014.05.012.

3. K.S. Murakami, S. Masuda, E.A. Campbell, O. Muzzin, S.A. Darst, Structural Basis of Transcription Initiation: An RNA Polymerase Holoenzyme-DNA Complex, Science. 296 (2002) 1285–1290. https://doi.org/10.1126/science.1069595.

4. M. Djordjevic, R. Bundschuh, Formation of the Open Complex by Bacterial RNA Polymerase—A Quantitative Model, Biophys J. 94 (2008) 4233–4248. https://doi.org/10.1529/biophysj.107.116970.

5. R. Glyde, F. Ye, V.C. Darbari, N. Zhang, M. Buck, X. Zhang, Structures of RNA Polymerase Closed and Intermediate Complexes Reveal Mechanisms of DNA Opening and Transcription Initiation, Mol Cell. 67 (2017) 106-116.e4. https://doi.org/10.1016/j.molcel.2017.05.010.

6. J. Chen, C. Chiu, S. Gopalkrishnan, A.Y. Chen, P.D.B. Olinares, R.M. Saecker, J.T. Winkelman, M.F. Maloney, B.T. Chait, W. Ross, R.L. Gourse, E.A. Campbell, S.A. Darst, Stepwise Promoter Melting by Bacterial RNA Polymerase, Mol Cell. 78 (2020) 275-288.e6. https://doi.org/10.1016/j.molcel.2020.02.017.

7. S. Sengupta, R.K. Prajapati, J. Mukhopadhyay, Promoter Escape with Bacterial Two-component σ Factor Suggests Retention of σ Region Two in the Elongation Complex*, J Biol Chem. 290 (2015) 28575–28583. https://doi.org/10.1074/jbc.m115.666008.

8. J. Ko, T. Heyduk, Kinetics of promoter escape by bacterial RNA polymerase: effects of promoter contacts and transcription bubble collapse, Biochem J. 463 (2014) 135–144. https://doi.org/10.1042/bj20140179.

9. R.M. Saecker, M.T. Record, P.L. deHaseth, Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-Competent Open Complexes, and Initiation of RNA Synthesis, J Mol Biol. 412 (2011) 754–771. https://doi.org/10.1016/j.jmb.2011.01.018.

10. S. Borukhov, E. Nudler, RNA polymerase: the vehicle of transcription, Trends Microbiol. 16 (2008) 126–134. https://doi.org/10.1016/j.tim.2007.12.006.

11. J.W. Roberts, S. Shankar, J.J. Filter, RNA Polymerase Elongation Factors, Annu Rev Microbiol. 62 (2008) 211–233. https://doi.org/10.1146/annurev.micro.61.080706.093422.

12. C. Cagliero, D.J. Jin, Dissociation and re-association of RNA polymerase with DNA during osmotic stress response in Escherichia coli, Nucleic Acids Res. 41 (2013) 315–326. https://doi.org/10.1093/nar/gks988.

13. V. Epshtein, V. Kamarthapu, K. McGary, V. Svetlov, B. Ueberheide, S. Proshkin, A. Mironov, E. Nudler, UvrD facilitates DNA repair by pulling RNA polymerase backwards, Nature. 505 (2014) 372. https://doi.org/10.1038/nature12928.

14. D.B. Srivastava, K. Leon, J. Osmundson, A.L. Garner, L.A. Weiss, L.F. Westblade, M.S. Glickman, R. Landick, S.A. Darst, C.L. Stallings, E.A. Campbell, Structure and function of CarD,

an essential mycobacterial transcription factor, Proc National Acad Sci. 110 (2013) 12619–12624. https://doi.org/10.1073/pnas.1308270110.

15. C.L. Stallings, N.C. Stephanou, L. Chu, A. Hochschild, B.E. Nickels, M.S. Glickman, CarD Is an Essential Regulator of rRNA Transcription Required for Mycobacterium tuberculosis Persistence, Cell. 138 (2009) 146–159. https://doi.org/10.1016/j.cell.2009.04.041.

16. J. Rammohan, A. Ruiz Manzano, A.L. Garner, J. Prusa, C.L. Stallings, E.A. Galburt, Cooperative stabilization of Mycobacterium tuberculosis rrnAP3 promoter open complexes by RbpA and CarD, Nucleic Acids Res. 44 (2016) 7304–7313. https://doi.org/10.1093/nar/gkw577.

17. N.M. Haines, Y.-I.T. Kim, A.J. Smith, N.J. Savery, Stalled transcription complexes promote DNA repair at a distance, Proc National Acad Sci. 111 (2014) 4037–4042. https://doi.org/10.1073/pnas.1322350111.

18. B. Bae, J. Chen, E. Davis, K. Leon, S.A. Darst, E.A. Campbell, CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex, Elife. 4 (2015) e08505. https://doi.org/10.7554/elife.08505.

19. J.-S. Park, M.T. Marr, J.W. Roberts, E. coli Transcription Repair Coupling Factor (Mfd Protein) Rescues Arrested Complexes by Promoting Forward Translocation, Cell. 109 (2002) 757–767. https://doi.org/10.1016/s0092-8674(02)00769-9.

20. T. Saha, K. Shukla, R.S. Thakur, A. Desingu, G. Nagaraju, Mycobacterium tuberculosis UvrD1 and UvrD2 helicases unwind G-quadruplex DNA, Febs J. 286 (2019) 2062–2086. https://doi.org/10.1111/febs.14798.

21. J. Houghton, C. Townsend, A.R. Williams, A. Rodgers, L. Rand, K.B. Walker, E.C. Böttger, B. Springer, E.O. Davis, Important Role for Mycobacterium tuberculosis UvrD1 in Pathogenesis and Persistence apart from Its Function in Nucleotide Excision Repair, J Bacteriol. 194 (2012) 2916–2923. https://doi.org/10.1128/jb.06654-11.

22. C. Güthlein, R.M. Wanner, P. Sander, E.O. Davis, M. Bosshard, J. Jiricny, E.C. Böttger, B. Springer, Characterization of the Mycobacterial NER System Reveals Novel Functions of the uvrD1 Helicase, J Bacteriol. 191 (2009) 555–562. https://doi.org/10.1128/jb.00216-08.

23. K. Sanders, C.-L. Lin, A.J. Smith, N. Cronin, G. Fisher, V. Eftychidis, P. McGlynn, N.J. Savery, D.B. Wigley, M.S. Dillingham, The structure and function of an RNA polymerase interaction domain in the PcrA/UvrD helicase, Nucleic Acids Res. 45 (2017) gkx074-. https://doi.org/10.1093/nar/gkx074.

24. C.L. Stallings, M.S. Glickman, CarD, Biochem Soc Symp. 2 (2011) 15–18. https://doi.org/10.4161/trns.2.1.13628.

25. E. Lerner, S. Chung, B.L. Allen, S. Wang, J. Lee, S.W. Lu, L.W. Grimaud, A. Ingargiola, X. Michalet, Y. Alhadid, S. Borukhov, T.R. Strick, D.J. Taatjes, S. Weiss, Backtracked and paused

transcription initiation intermediate of Escherichia coli RNA polymerase, Proc National Acad Sci. 113 (2016) E6562–E6571. https://doi.org/10.1073/pnas.1605038113.

26. E. Nudler, RNA Polymerase Backtracking in Gene Regulation and Genome Instability, Cell. 149 (2012) 1438–1445. https://doi.org/10.1016/j.cell.2012.06.003.

27. T.M. Lohman, E.J. Tomko, C.G. Wu, Non-hexameric DNA helicases and translocases: mechanisms and regulation, Nat Rev Mol Cell Bio. 9 (2008) 391–401. https://doi.org/10.1038/nrm2394.

28. A.A. Kawale, B.M. Burmann, UvrD helicase–RNA polymerase interactions are governed by UvrD's carboxy-terminal Tudor domain, Commun Biology. 3 (2020) 607. https://doi.org/10.1038/s42003-020-01332-2.

29. K.M. Sinha, M.S. Glickman, S. Shuman, Mutational Analysis of Mycobacterium UvrD1 Identifies Functional Groups Required for ATP Hydrolysis, DNA Unwinding, and Chemomechanical Coupling, Biochemistry-Us. 48 (2009) 4019–4030. https://doi.org/10.1021/bi900103d.

30. B. Bae, A. Feklistov, A. Lass-Napiorkowska, R. Landick, S.A. Darst, Structure of a bacterial RNA polymerase holoenzyme open promoter complex, Elife. 4 (2015) e08504. https://doi.org/10.7554/elife.08504.

31. A. Chadda, D. Jensen, E.J. Tomko, A.R. Manzano, B. Nguyen, T.M. Lohman, E.A. Galburt, Mycobacterium tuberculosis DNA repair helicase UvrD1 is activated by redox-dependent dimerization via a 2B domain cysteine, Proc National Acad Sci. 119 (2022) e2114501119. https://doi.org/10.1073/pnas.2114501119.

32. A.A. Kawale, B.M. Burmann, UvrD helicase–RNA polymerase interactions are governed by UvrD's carboxy-terminal Tudor domain, Commun Biology. 3 (2020) 607. https://doi.org/10.1038/s42003-020-01332-2.

33. K.A. Reynolds, C.E. Cameron, K.D. Raney, Melting of Duplex DNA in the Absence of ATP by the NS3 Helicase Domain through Specific Interaction with a Single-Strand/Double-Strand Junction, Biochemistry-Us. 54 (2015) 4248–4258. https://doi.org/10.1021/acs.biochem.5b00214.

34. T.M. Lohman, K.P. Bjornson, Mechanisms of Helicase-Catalyzed DNA Unwinding, Annu Rev Biochem. 65 (1996) 169–214. https://doi.org/10.1146/annurev.bi.65.070196.001125.

Chapter 7: Summary of the research project

My project aims to study the activation mechanism of a helicase called UvrD1 from *Mycobacterium tuberculosis* (*Mtb)*. Helicases are proteins that use energy from ATP hydrolysis to translocate on ssDNA or unwind dsDNA to ssDNA[1]. *Mtb* causes tuberculosis disease in humans and is responsible for millions of deaths each year[2]. During infection, the bacterium is exposed to both reactive oxygen species (ROS) and nitrogen intermediates (RNI) due to host response to infection that can cause DNA damage[3]. UvrD1 helicase is one of the proteins that is involved in DNA damage response in *Mtb*[4]. UvrD helicases have been shown to unwind dsDNA during nucleotide excision repair (NER) and remove the damaged ssDNA strand, it can also translocate on ssDNA and remove protein bound to DNA. Recently UvrD was also shown to backtrack RNA polymerase stalled at a lesion playing a role in important cellular processes like replication, transcription, and repair[5–7]. The process of DNA unwinding needs to be regulated in the cell as an unregulated helicase can be harmful by destabilizing duplex DNA. It has been shown that the helicase and translocase activities of UvrD like helicases are regulated by its self-assembly state *in-vitro*[8]. Presence of accessory factors can also lead to activation of helicase activity[9–11]. Previous studies have reported that *Mtb* UvrD1 exists as a monomer and unwinds DNA over a timescale of minutes and Ku an accessory factor can activate the helicase activity of UvrD1[12]. However, the kinetics of unwinding and active species is unknown since purification and unwinding studies were performed in different solution conditions[13]. The results from my thesis show that *Mtb* UvrD1 exists as a mixture of monomers and dimers as a function of redox[14]. The critical cysteine residue responsible for the redox dependent dimer is the one located in the 2B domain of *Mtb* UvrD1 which suggests this domain to be the dimerization domain of UvrD like

helicases. In addition, the monomer of UvrD1 that is formed in reducing conditions or by mutating 2B domain cysteine to alanine cannot unwind the DNA on shorter time scales in single round conditions.

This suggests that translocation on ssDNA is not sufficient for unwinding a dsDNA. However, this monomer can be activated for unwinding by interaction with an accessory factor Ku via C-terminal Tudor domain of UvrD1. Unwinding by the UvrD1 monomer is slow over a timescale of 10 minutes, multiround and in presence of around 3 Ku dimers bound to DNA substrate that we tested(unpublished). These results suggest the role of UvrD1 in different repair pathways where a processive standalone dimer can initiate DNA repair during NER and an un-processive monomer can interact with Ku and participate in double strand break repair pathways.

7.1 Hypotheses

The premise of the hypotheses is based on extensive amounts of studies that have characterized the biochemical and biophysical properties of *E. coli* UvrD which can self-associate in solution in the absence and in the presence of DNA substrate[8,15]. The self-assembly state of UvrD plays a regulatory role for its translocase and helicase activities. A UvrD monomer can translocate processively in 3' to 5' direction along ssDNA however, processive DNA unwinding activity requires formation of at least a UvrD dimer or higher order oligomer in the absence of protein factors[15–17]. UvrD dimerization on the DNA substrate is also accompanied by rotation of the 2B sub-domain of the lead UvrD monomer to a more closed state resulting in helicase activation[18]. Furthermore, it was shown that the monomeric forms of *E. coli* UvrD and its close relatives, *E. coli* Rep and *B. stearothermophilus* PcrA, possess latent helicase activity that can be stimulated by the assistance of the pulling force onto the DNA substrate[19], relieving the auto-

inhibitory effect of the 2B sub-domain[20,21] or in the presence of an accessory protein[10,11,18,22]. While many studies have been reported focusing on UvrD family helicases from model bacteria, less is known about these enzymes in the distantly related human pathogen *Mtb*. Previous work on *Mtb* UvrD1 has shown that it is important for survival after UV and oxidative damage as well as for pathogenesis in mice[23]. In stark contrast to other UvrD family members, UvrD1 was reported to be monomeric and to either possess helicase activity on longer timescales or require activation via the binding of *Mtb* Ku. Based on previous studies it was also proposed that *Mtb* UvrD1 does not oligomerize in solution and is a weak helicase that can be activated by accessory factor Ku[12,13]. During characterization of *Mtb* UvrD1, I found that instead of a monomer UvrD1 exists as a mixture of monomers and dimers by change in redox potential and only the redox dependent dimer of UvrD1 unwinds DNA on shorter timescales. The previously published results of UvrD dimer being the unwinding species were further strengthened by the fact that the monomer in reducing conditions cannot unwind the DNA and 2B domain is the dimerization interface of UvrD like helicases. In addition, since Ku was shown to activate the helicase activity of UvrD1, I found that 3Ku dimers bound to DNA substrate can increase the helicase activity of a monomer of UvrD1 by the one bound near the junction of the duplex interacting with UvrD1 monomer via C-terminal Tudor domain.

7.2 Methods

The hypotheses above seek to determine the activating species of *Mtb* UvrD1. Therefore, experiments in this research project were designed to measure the assembly, conformational state in the absence and presence of ligands, and unwinding activity of UvrD1. Analytical ultracentrifugation experiments provided information on assembly states of UvrD1 and how the

conformation changes in presence of ligands and accessory factors. Information about the fast kinetics of unwinding in the range less than a second can be obtained from pre-steady state stopped-flow experiments by using a sequential-mixing mode to separate steps involving the formation of the active helicase species. They can be performed in both single round conditions where protein bound to DNA can unwind without being falling off and in multiround conditions where protein must bind and unbind to DNA multiple times before completing unwinding. This gives information about the processivity of the enzyme. Fluorescence anisotropy or quenching experiments allowed us to measure binding affinity of proteins with nucleic acids and determine stoichiometry of protein bound to DNA.

7.3 Results and Conclusions

During purification and characterization of *Mtb* UvrD1, I found that instead of a monomer, UvrD1 exists as a mixture of monomer, dimer, and oligomer in solution. Titration of increasing amount of salt changes the distribution to monomer and dimer population suggesting that the oligomer is formed due to electrostatic interaction however a dimer is redox dependent. Previous studies on *Mtb* UvrD1 have reported it to be a monomeric species and shown that it unwinds as a monomer. However, purification experiments were performed in presence of reducing agent and unwinding was performed in its absence[13]. Addition of a reducing agent changed the distribution of UvrD1 to a monomer suggesting that the dimer of UvrD1 is redox dependent. Since *Mtb* UvrD1 undergoes oxidative stress during its life cycle, I tested the hypothesis that change in redox potential may affect its function. I performed *in-vitro* unwinding assays under conditions when UvrD1 is a mixture of dimer and monomers and found that it can unwind DNA processively, that is in single round conditions when both dimers and monomers are present. In the presence of

reducing agent when UvrD1 is a monomer, it cannot unwind DNA. By measuring the translocation and binding activities of UvrD1 to DNA, I found that both monomers and dimers of UvrD1 can bind and translocate on DNA but only a dimer of UvrD1 can unwind DNA. Extensive work on UvrD like helicases have shown that they unwind as dimers but none of the studies have reported the dimerization interface of UvrD like helicases. Since *Mtb* UvrD1 resembles UvrD like helicases being 42% identical amino acid sequence and has two domains 1A and 2A RecA like domains with inserted domains 1B and 2B[1]. The predicted structure of *Mtb* UvrD1 showed that out of three cysteines two are partially buried in the 1A and 1B domain and one is solvent exposed in the 2B domain. When I mutated the cysteines in 1A and 1B domain *Mtb* UvrD1 retained both the monomer and the dimer species however mutating the cysteine in 2B domain to alanine *Mtb* UvrD1 changed to a monomer showing that the cysteine in 2B domain is responsible for dimerization and 2B domain likely is the dimerization interface of *Mtb* UvrD1 and for other helicases belonging to same subfamily. Previous studies have shown that Ku a factor involved in non-homologous end joining (NHEJ) increases the helicases activity of *Mtb* UvrD1 however since UvrD1 exists as a mixture of monomers and dimers it was unknown whether Ku can increase the helicase activity of monomers or dimers of UvrD1[12]. Prokaryotic Ku protein exists as a homodimer in solution unlike eukaryotic Ku that exists as a heterodimer[24,25]. I discovered that prokaryotic Ku that exists as dimer at lower micromolar concentrations binds to DNA only in presence of Mg^{+2} ion. The binding affinity of Ku is with higher nanomolar affinity and the partial duplex of 32bp with dT20 tail can bind up to 3 Ku homodimers. The binding of second or third Ku dimer on DNA is highly cooperative suggesting the one bound on junction interacts with UvrD1 monomer. The same partial duplex was used to determine the unwinding activity of UvrD1 monomers and dimers in presence of Ku. I found that the monomers of UvrD1 can unwind DNA

slowly over a timescale of 10 minutes and in multiround conditions in comparison to processive dimer that unwinds DNA within a second. Moreover, Ku can stimulate the unwinding of UvrD1 monomer under multiround conditions but did not affect the unwinding by dimer. To determine if Ku interacts with UvrD1 monomer via its C-terminal domain, I deleted the C-terminal Tudor domain of UvrD1 that is involved in interaction with RNAP and is a protein-protein interaction domain. Deleting the C-terminal domain increases the affinity of UvrD1 for DNA but abrogates the Ku dependent activation of UvrD1 monomer unwinding suggesting that Ku-UvrD1 monomer interaction occurs via Tudor domain. To summarize my experiments, suggest a model where the processive redox dependent dimer of UvrD1 can unwind DNA within a second and may play a role in nucleotide excision repair pathways while an unprocessive monomer can unwind DNA slowly and in multiround conditions in the presence of Ku and play a role in double strand break repair (DSBR). To understand the structural basis of unwinding by a UvrD1 dimer, I also determined the Cryo-EM structure of UvrD1 dimer alone in presence of crosslinker. By threading the PDB model of *Mtb* UvrD1 dimer in the electron density map gave a correlation coefficient of 0.7 and the cystine bond is visible verifying our *in-vitro* experiments. The structure also suggested that the two monomers of UvrD1 are not completely symmetric although in a closed conformation by rotation of 2B domain involved in dimerization but requires a higher resolution structure to say it with confidence.

7.4 Future directions

My thesis is focused on characterization of a helicase called UvrD1 from *Mtb*. It revealed aspects of helicase activation via dimerization by a novel mechanism of redox stress and revealed the dimerization interface of UvrD like helicases that was previously unknown. My results also

show the activation mechanism of UvrD1 monomer by means of accessory factor called Ku. Using Cryo-EM, I also obtained the structure of UvrD1 dimer in two conformations open and closed. These results also opened possibilities for many questions related to future research.

7.4.1 To determine if UvrD1 monomer or dimer can translocate on double stranded DNA

Previous studies have shown that UvrD helicase monomers are single stranded DNA translocases however certain helicases have been shown to translocate on double stranded DNA[26]. Translocation on double stranded DNA has been observed for helicases like TFIIH where the translocase unit SSl2 promotes DNA opening by tracking on one DNA strand and inserting DNA into Polymerase cleft to open DNA by unwinding[27]. For a helicase to interact with dsDNA it must interact with phosphodiester backbone of DNA and it's unknown if a monomer or dimer of UvrD1 can translocate on dsDNA. Ongoing work from UvrD and Rep like helicases has provided evidence that ATPase activity of both subunits of a dimer are required for unwinding. This suggests a mechanism of UvrD1 unwinding where one subunit interacts and translocate on ssDNA and other with the dsDNA during the process of unwinding. Since *Mtb* UvrD1 can form a redox dependent dimer, it is possible to design DNA substrates and test if the monomer/dimer of UvrD1 is a dsDNA translocase[28]. This measurement would help understand the mechanism of unwinding by a dimer of UvrD1.

7.4.2 To determine the role of UvrD1 in transcription-coupled repair (TCR)

UvrD like helicases have been shown to backtrack or pull RNA polymerases backwards during the elongation step of transcription when it is stalled at the site of lesion[29]. Another transcription factor in *Mtb* called CarD has been shown to bind and stabilize RNAP promoter open complexes during the initiation step of transcription[30–32]. Both CarD and UvrD bind to same site called beta-1 lobe of RNAP via Tudor domain[32,33]. This suggests that they both compete

for binding to RNAP. One hypothesis to test in future would be that a change in the conformation of RNAP during various steps of transcription controls binding and recruitment of correct factors. For instance, binding of initiation factor CarD stabilizes RNAP open complex for initiation of transcription whereas binding of UvrD stabilizes RNAP for repair[30,31]. Thus, measuring binding affinity during various stages of transcription (i.e. initiation, elongation, or stalling) RNAP will increase or decrease its preference for specific factors. Experiments like an *in-vitro* transcription assay can also be used to test the hypothesis that *Mtb* UvrD1 pulls stalled RNAP backward in TCR. Previous studies have shown *E. coli* UvrD pulls RNAP backward that has been stalled at the lesion[29]. Another question in the field is whether a dimer or a monomer of UvrD is involved in interaction with RNAP. Crosslinking studies suggest that a dimer of UvrD is involved in TCR however the C-terminal domain (CTD) called Tudor domain of only one monomer of UvrD1 has been shown to interact with RNAP[34]. I found that the Tudor domain of UvrD1 monomer is involved in interaction with Ku and deleting this domain abrogates the activation of unwinding by Ku. Since *Mtb* UvrD1 exists as a monomer under reducing conditions one can measure whether a monomer or a dimer of UvrD1 participates in RNAP backtracking and if the interaction occurs via Tudor domain.

7.4.3 Structural studies of *Mtb* UvrD1 with its interacting partners

I have determined the structure of UvrD1 dimer in the presence and absence of crosslinker but because of heterogeneity in one of the two monomeric subunits the structure is at low resolution. However, one can see the cystine bond and align the PDB model of the crystal structure of UvrD monomer into the electron density map. One of the future directions would be to determine the structure of UvrD1 dimer in the presence of DNA to understand the structural basis of UvrD1 dimer unwinding. I also found that 3 Ku dimers bound to partial duplex 32base pairs

with a dT20 tail and can activate the unwinding of UvrD1 monomer. In addition, Ku dimers can bind this DNA substrate only in the presence of Mg^{+2} ion(unpublished). Previous studies of Ku binding to DNA have shown that it binds in a sequence independent manner, in a cooperative mode and multiple Ku dimers can load on DNA, but the mechanism of binding is poorly understood[35–37]. One of the future studies would include getting the structure of Ku dimers bound to DNA substrate in presence of Mg^{+2} and understand the role of Mg^{+2} in cooperativity. It would provide information on how the conformation of Ku dimer changes in presence of Mg^{+2} that affects its ability to bind DNA. By including UvrD1 monomer in presence of Ku would provide information on the structural basis of unwinding by UvrD1 monomer and the conformation of UvrD1 bound to Ku on DNA. I also have preliminary data showing that *Mtb* UvrD1 interacts with RNAP so one potential study would be to get structure of UvrD1 bound to RNAP complex. This would help understand the structural basis of transcription-coupled repair in *Mtb* and will provide information on how a monomer or a dimer of UvrD1 interacts with RNAP.

7.4.4 Characterization of helicase activity of *Mtb* UvrD2 in the absence and presence of accessory factors

Mtb has two UvrD like helicases, UvrD1 and UvrD2. The N-terminal helicase domains of UvrD1 and UvrD2 are similar, but UvrD2 has a unique tetracysteine motif and a C-terminal HRDC domain[38]. Previous studies have shown that deleting the tetracysteine motif along with C-terminal domain leads to abrogation of the helicase activity of UvrD2 which is restored by addition of Ku[38]. However, replacing the cysteines with alanine in the tertacysteine module did not affect the helicase activity of UvrD2 suggesting that the arrangement of cysteines or the module and not the actual residues are responsible for unwinding. The role of CXXC motif of UvrD2 in Zinc binding and oligomerization for UvrD2 is not clear. Also, the role of C-terminal domain that

contains both the CXXC motifs and the HRDC domain is not clear. Future studies should focus on characterization of UvrD2 and how the mutations in the CXXC motifs or the C-terminal HRDC domain affect its unwinding and translocation activity. In addition, since unwinding of UvrD1 is affected by change in redox it is important to determine the function of UvrD2 as a function of change in redox and how it changes in presence of accessory factor Ku.

7.5 References

1. T.M. Lohman, E.J. Tomko, C.G. Wu, Non-hexameric DNA helicases and translocases: mechanisms and regulation, Nat Rev Mol Cell Bio. 9 (2008) 391–401. https://doi.org/10.1038/nrm2394.

2. S. Bagcchi, WHO's Global Tuberculosis Report 2022, Lancet Microbe. 4 (2022) e20. https://doi.org/10.1016/s2666-5247(22)00359-7.

3. M.I. Voskuil, I.L. Bartek, K. Visconti, G.K. Schoolnik, The Response of Mycobacterium Tuberculosis to Reactive Oxygen and Nitrogen Species, Front Microbiol. 2 (2011) 105. https://doi.org/10.3389/fmicb.2011.00105.

4. J. Houghton, C. Townsend, A.R. Williams, A. Rodgers, L. Rand, K.B. Walker, E.C. Böttger, B. Springer, E.O. Davis, Important Role for Mycobacterium tuberculosis UvrD1 in Pathogenesis and Persistence apart from Its Function in Nucleotide Excision Repair, J Bacteriol. 194 (2012) 2916–2923. https://doi.org/10.1128/jb.06654-11.

5. R.M. Brosh, DNA helicases involved in DNA repair and their roles in cancer, Nat Rev Cancer. 13 (2013) 542–558. https://doi.org/10.1038/nrc3560.

6. S.W. Matson, D.W. Bean, J.W. George, DNA helicases: Enzymes with essential roles in all aspects of DNA metabolism, Bioessays. 16 (1994) 13–22. https://doi.org/10.1002/bies.950160103.

7. W. Yang, Lessons Learned from UvrD Helicase: Mechanism for Directional Movement*, Biophysics. 39 (2010) 367–385. https://doi.org/10.1146/annurev.biophys.093008.131415.

8. N.K. Maluf, T.M. Lohman, Self-association Equilibria of Escherichia coli UvrD Helicase Studied by Analytical Ultracentrifugation, J Mol Biol. 325 (2003) 889–912. https://doi.org/10.1016/s0022-2836(02)01276-7.

9. B. Nguyen, M.K. Shinn, E. Weiland, T.M. Lohman, Regulation of E. coli Rep helicase activity by PriC, J Mol Biol. 433 (2021) 167072. https://doi.org/10.1016/j.jmb.2021.167072.

10. Y.A. Ordabayev, B. Nguyen, A.G. Kozlov, H. Jia, T.M. Lohman, UvrD helicase activation by MutL involves rotation of its 2B subdomain, Proc National Acad Sci. 116 (2019) 16320–16325. https://doi.org/10.1073/pnas.1905513116.

11. Y.A. Ordabayev, B. Nguyen, A. Niedziela-Majka, T.M. Lohman, Regulation of UvrD Helicase Activity by MutL, J Mol Biol. 430 (2018) 4260–4274. https://doi.org/10.1016/j.jmb.2018.08.022.

12. K.M. Sinha, N.C. Stephanou, F. Gao, M.S. Glickman, S. Shuman, Mycobacterial UvrD1 Is a Ku-dependent DNA Helicase That Plays a Role in Multiple DNA Repair Events, Including Double-strand Break Repair*, J Biol Chem. 282 (2007) 15114–15125. https://doi.org/10.1074/jbc.m701167200.

13. E. Curti, S.J. Smerdon, E.O. Davis, Characterization of the Helicase Activity and Substrate Specificity of Mycobacterium tuberculosis UvrD, J Bacteriol. 189 (2007) 1542–1555. https://doi.org/10.1128/jb.01421-06.

14. A. Chadda, D. Jensen, E.J. Tomko, A.R. Manzano, B. Nguyen, T.M. Lohman, E.A. Galburt, Mycobacterium tuberculosis DNA repair helicase UvrD1 is activated by redox-dependent dimerization via a 2B domain cysteine, Proc National Acad Sci. 119 (2022) e2114501119. https://doi.org/10.1073/pnas.2114501119.

15. N.K. Maluf, C.J. Fischer, T.M. Lohman, A Dimer of Escherichia coli UvrD is the Active Form of the Helicase In Vitro, J Mol Biol. 325 (2003) 913–935. https://doi.org/10.1016/s0022-2836(02)01277-9.

16. E.J. Tomko, C.J. Fischer, A. Niedziela-Majka, T.M. Lohman, A Nonuniform Stepping Mechanism for E. coli UvrD Monomer Translocation along Single-Stranded DNA, Mol Cell. 26 (2007) 335–347. https://doi.org/10.1016/j.molcel.2007.03.024.

17. J.A. Ali, T.M. Lohman, Kinetic Measurement of the Step Size of DNA Unwinding by Escherichia coli UvrD Helicase, Science. 275 (1997) 377–380. https://doi.org/10.1126/science.275.5298.377.

18. B. Nguyen, Y. Ordabayev, J.E. Sokoloski, E. Weiland, T.M. Lohman, Large domain movements upon UvrD dimerization and helicase activation, Proc National Acad Sci. 114 (2017) 12178–12183. https://doi.org/10.1073/pnas.1712882114.

19. M.J. Comstock, K.D. Whitley, H. Jia, J. Sokoloski, T.M. Lohman, T. Ha, Y.R. Chemla, Direct observation of structure-function relationship in a nucleic acid–processing enzyme, Science. 348 (2015) 352–354. https://doi.org/10.1126/science.aaa0130.

20. W. Cheng, K.M. Brendza, G.H. Gauss, S. Korolev, G. Waksman, T.M. Lohman, The 2B domain of the Escherichia coli Rep protein is not required for DNA helicase activity, Proc National Acad Sci. 99 (2002) 16006–16011. https://doi.org/10.1073/pnas.242479399.

21. S. Arslan, R. Khafizov, C.D. Thomas, Y.R. Chemla, T. Ha, Engineering of a superhelicase through conformational control, Science. 348 (2015) 344–347. https://doi.org/10.1126/science.aaa0445.

22. L.T. Chisty, C.P. Toseland, N. Fili, G.I. Mashanov, M.S. Dillingham, J.E. Molloy, M.R. Webb, Monomeric PcrA helicase processively unwinds plasmid lengths of DNA in the presence of the initiator protein RepD, Nucleic Acids Res. 41 (2013) 5010–5023. https://doi.org/10.1093/nar/gkt194.

23. J. Houghton, C. Townsend, A.R. Williams, A. Rodgers, L. Rand, K.B. Walker, E.C. Böttger, B. Springer, E.O. Davis, Important Role for Mycobacterium tuberculosis UvrD1 in Pathogenesis and Persistence apart from Its Function in Nucleotide Excision Repair, J Bacteriol. 194 (2012) 2916–2923. https://doi.org/10.1128/jb.06654-11.

24. R. Gupta, D. Barkan, G. Redelman-Sidi, S. Shuman, M.S. Glickman, Mycobacteria exploit three genetically distinct DNA double-strand break repair pathways, Mol Microbiol. 79 (2011) 316–330. https://doi.org/10.1111/j.1365-2958.2010.07463.x.

25. S. Zahid, M.S.E. Dahan, F. Iehl, P. Fernandez-Varela, M.-H.L. Du, V. Ropars, J.B. Charbonnier, The Multifaceted Roles of Ku70/80, Int J Mol Sci. 22 (2021) 4134. https://doi.org/10.3390/ijms22084134.

26. R.G. Lloyd, C.J. Rudolph, 25 years on and no end in sight: a perspective on the role of RecG protein, Curr Genet. 62 (2016) 827–840. https://doi.org/10.1007/s00294-016-0589-z.

27. J. Fishburn, E. Tomko, E. Galburt, S. Hahn, Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation, Proc National Acad Sci. 112 (2015) 3961–3966. https://doi.org/10.1073/pnas.1417709112.

28. C.J. Fischer, N.K. Maluf, T.M. Lohman, Mechanism of ATP-dependent Translocation of E.coli UvrD Monomers Along Single-stranded DNA, J Mol Biol. 344 (2004) 1287–1309. https://doi.org/10.1016/j.jmb.2004.10.005.

29. V. Epshtein, V. Kamarthapu, K. McGary, V. Svetlov, B. Ueberheide, S. Proshkin, A. Mironov, E. Nudler, UvrD facilitates DNA repair by pulling RNA polymerase backwards, Nature. 505 (2014) 372. https://doi.org/10.1038/nature12928.

30. B. Bae, J. Chen, E. Davis, K. Leon, S.A. Darst, E.A. Campbell, CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex, Elife. 4 (2015) e08505. https://doi.org/10.7554/elife.08505.

31. J. Rammohan, A. Ruiz Manzano, A.L. Garner, J. Prusa, C.L. Stallings, E.A. Galburt, Cooperative stabilization of Mycobacterium tuberculosis rrnAP3 promoter open complexes by RbpA and CarD, Nucleic Acids Res. 44 (2016) 7304–7313. https://doi.org/10.1093/nar/gkw577.

32. D.B. Srivastava, K. Leon, J. Osmundson, A.L. Garner, L.A. Weiss, L.F. Westblade, M.S. Glickman, R. Landick, S.A. Darst, C.L. Stallings, E.A. Campbell, Structure and function of CarD, an essential mycobacterial transcription factor, Proc National Acad Sci. 110 (2013) 12619–12624. https://doi.org/10.1073/pnas.1308270110.

33. K. Sanders, C.-L. Lin, A.J. Smith, N. Cronin, G. Fisher, V. Eftychidis, P. McGlynn, N.J. Savery, D.B. Wigley, M.S. Dillingham, The structure and function of an RNA polymerase interaction domain in the PcrA/UvrD helicase, Nucleic Acids Res. 45 (2017) gkx074-. https://doi.org/10.1093/nar/gkx074.

34. B. Martinez, B.K. Bharati, V. Epshtein, E. Nudler, Pervasive Transcription-coupled DNA repair in E. coli, Nat Commun. 13 (2022) 1702. https://doi.org/10.1038/s41467-022-28871-y.

35. R. Öz, J.L. Wang, R. Guerois, G. Goyal, S. KK, V. Ropars, R. Sharma, F. Koca, J.-B. Charbonnier, M. Modesti, T.R. Strick, F. Westerlund, Dynamics of Ku and bacterial non-homologous end-joining characterized using single DNA molecule analysis, Nucleic Acids Res. 49 (2021) gkab083-. https://doi.org/10.1093/nar/gkab083.

36. Y. Ma, M.R. Lieber, DNA Length-Dependent Cooperative Interactions in the Binding of Ku to DNA †, Biochemistry-Us. 40 (2001) 9638–9646. https://doi.org/10.1021/bi010932v.

37. J.R. Walker, R.A. Corpina, J. Goldberg, Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair, Nature. 412 (2001) 607–614. https://doi.org/10.1038/35088000.

38. K.M. Sinha, N.C. Stephanou, M.-C. Unciuleac, M.S. Glickman, S. Shuman, Domain Requirements for DNA Unwinding by Mycobacterial UvrD2, an Essential DNA Helicase, Biochemistry-Us. 47 (2008) 9355–9364. https://doi.org/10.1021/bi800725q.