November 2023

A Biochemical Approach to Characterize a Divergent Trypanosoma brucei Mitochondrial DNA Polymerase, POLIB

Stephanie B. Delzell
University of Massachusetts Amherst

A Biochemical Approach to Characterize a Divergent Trypanosoma brucei Mitochondrial DNA Polymerase, POLIB

Stephanie B. Delzell

A Biochemical Approach to Characterize a Divergent *Trypanosoma brucei* Mitochondrial DNA Polymerase, POLIB

A dissertation presented by

STEPHANIE BROOKLYN DELZELL

Submitted to The Graduate School of the

University of Massachusetts Amherst

in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2023

Department of Microbiology

A Biochemical Approach to Characterize a Divergent *Trypanosoma brucei* Mitochondrial DNA Polymerase, POLIB

A Dissertation Presented

By

STEPHANIE BROOKLYN DELZELL

Approved as to style and content by:

_____

Michele M. Klingbeil, Chair

_____

Yasu S. Morita, Committee member

_____

Scott W. Nelson, Committee member

_____

Steven J. Sandler, Committee member

_____

James F. Holden,
Department Head,
Department of Microbiology

DEDICATION

To my parents, for leading me into a love of life, learning, and science.

ACKNOWLEDGEMENTS

They say it takes a village to raise a Ph.D. student. I cannot imagine how I would have completed this project without my village:

Firstly, I would like to thank my advisor, Michele Klingbeil, for guiding me through the winding path that is grad school. You have given me the opportunities and means to gain so many tools in science, communication, and life that I know I will continue to use no matter what comes next.

I would also like to thank my dissertation committee: Scott Nelson, and his lab, for training me on some of the most critical assays I used in this project and always providing valuable insight along the way. Yasu Morita for always being a resource as an excellent GPD and committee member. And Steve Sandler for many useful pieces of advice and words of encouragement.

Klingbeil lab members past and present, Especially Jonathan Miller, Raveen Armstrong, David Anaguano, Yadi Bermudez, Matthew Frost, Brendan Maher, and Maira Ahmed for helpful technical advice, scientific insight, and a happy, supportive lab environment.

Thank you to my dear friends, Bechtold, Melzer, Lian, and Nate who made even the hardest parts of grad school (and a pandemic) amazing with good advice, better food, and lots of laughter.

To my family, Mom, Alexis, and Paul for all the fun interludes and adventures that help keep me sane (or close enough to sane). Thanks mom, for the mitochondrial DNA that powers everything I do.

I’d like to thank my cats, Clementine and Salem, for always giving me honest feedback, and occasionally trapping me on the couch.

Thank you Zach, for double-checking my math, giving the most practical suggestions, fixing every broken thing, and offering me endless support through this process.

ABSTRACT

A Biochemical Approach to Characterize a Divergent *Trypanosoma brucei* Mitochondrial DNA Polymerase, POLIB

September 2023

Stephanie B. Delzell

Directed by: Professor Michele M. Klingbeil

Trypanosoma brucei is a single-celled parasitic protist that causes African sleeping sickness in people and nagana in cattle in sub-Saharan Africa. *T. brucei* and related trypanosomatid parasites contain an unusual catenated mitochondrial genome known as kinetoplast DNA (kDNA) composed of dozens of 23 kb maxicircles and thousands of 1 kb minicircles. The kDNA structure and replication mechanism are divergent from other eukaryotes and essential for parasite survival. POLIB is one of three Family A DNA polymerases that are independently essential to maintain the kDNA network, and has been implicated in minicircle replication. However, the division of labor among the paralogs, particularly which might be a replicative, proofreading enzyme remains enigmatic. *De novo* modelling of POLIB suggested a structure that is divergent from all other Family A polymerases in which the thumb subdomain contains a 369 amino acid insertion with homology to DEDDh DnaQ family 3'-5' exonucleases. In chapter 2, we explore the polymerase and exonuclease activity of POLIB using purified his-tagged

recombinant variants that have been truncated and codon optimized for expression in *E. coli*. Using this recombinant protein variants we demonstrated that the 3'-5' exonuclease activity of recombinant POLIB prefers DNA vs. RNA substrates and prefers single-stranded vs. double-stranded substrates. POLIB exonuclease activity prevails over polymerase activity on DNA substrates at pH 8.0, while DNA primer extension is favored at pH 6.0. Mutations that ablate POLIB polymerase activity slow the exonuclease rate suggesting crosstalk between the domains. We show that POLIB extends an RNA primer more efficiently than a DNA primer in the presence of dNTPs but does not incorporate rNTPs efficiently using either RNA or DNA primers. Immunoprecipitation of Pol I-like paralogs from *T. brucei* corroborate the pH selectivity and RNA primer preferences of POLIB and revealed that the other paralogs efficiently extend a DNA primer. We also show that overexpression of the exonuclease-ablated variant of POLIB in *T. brucei* results in a loss of fitness and impacts kDNA replication. We postulate that this unique enzyme and the machinery associated with it in the process of kDNA replication could be excellent drug targets worthy of further study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Background

1.1 Importance of *Trypanosoma brucei*

African trypanosomes are flagellated parasitic protists that are endemic to sub-Saharan Africa, where they are a significant threat to public health as well as an enormous economic burden to the agricultural industry. *Trypanosoma brucei* subspecies cause Human African Trypanosomiasis (HAT) and a wasting disease in cattle, domestic pigs, and other farm animals called Animal African Trypanosomiasis (AAT). The parasite is transmitted to mammalian hosts through the bite of the tsetse fly (*Glossina* species). While it is encouraging that HAT has been declining in recent years livestock infections remain high and are an economically devastating problem because the farm animals in sub-Saharan Africa are used not only as a source of nutrition, but cattle are also beasts of burden, allowing the cultivation of a variety of crops [1]. There are no preventative vaccines and African trypanosome infections are always fatal without treatment. There are few established treatments for trypanosomiasis and overall, the drugs treatments are challenging to administer, associated with toxicity, and some are not effective against later stages of the disease when the parasites have entered the central nervous system (Table 1) [2]. Finding biological processes that are highly specific to the parasites will help uncover new drug targets to develop treatments with minimal side effects.

Drug	Disease Stage	Subspecies	Administration	Side Effects	In Use Since
Pentamidine	Stage 1	*T. b. gambiense*	Intramuscular or IV injection	Nausea, vomiting, loss of appetite, unusual taste/dryness in the mouth, dizziness, diarrhea	1940
Suramin	Stage 1	*T. b. rhodesiense*	IV	Nephrotoxicity, allergic reactions	1920
Melarsoprol	Stage 2	*T. b. rhodesiense*	IV	Encephalopathy	1949
Eflornithine	Stage 2	*T. b. gambiense*	IV	fever, pruritus, hypertension, nausea, vomiting, diarrhea, abdominal pain, headaches, myelosuppression (anemia, leucopenia, thrombocytopenia) and, more rarely, seizures	1990
Nifurtimox	Stage 2	*T. b. gambiense*	Oral	abdominal pain, nausea, vomiting and headache	2009
Fexinidazole	Stages 1 and 2	*T. b. gambiense*	Oral	vomiting, nausea, asthenia, decrease appetite, headache, insomnia tremor and dizziness	2019

Table 1.1: Approved chemotherapies for treatment of HAT [3,4].

1.2 Lifecycle of *Trypanosoma brucei*

Throughout their lifecycle, African trypanosomes undergo morphological and metabolic phenotypic shifts to adapt to their current host environment. Trypanosomes living in the midgut of the tsetse fly host are procyclic form cells. These are distinguished from other trypanosome lifecycle stages in part by a surface coat made up of procyclins and an upregulated set of mitochondrial metabolic enzymes that allow the cells to utilize oxidative phosphorylation to generate ATP [5]. Tsetse flies quickly metabolize available glucose as an energy source for flight, and so this lifecycle stage of *T. brucei* has adapted to use available proline as their carbon source [6].

In the proventriculus of the tsetse fly, *T. brucei* differentiates into via asymmetrical division into proventricular forms elongated cells which are thought to die off, and epimastigotes, which then migrate to the salivary gland and adhere to epithelial tissue there (Figure 1.1). In the salivary gland, epimastigotes differentiate into metacyclic form cells that are primed for transmission to the mammalian bloodstream by expression of VSGs from a limited set of metacyclic VSGs (mVSGs), but are not proliferative [5,7].

Once the tsetse fly takes a blood meal, trypanosomes enter the bloodstream of the mammalian host where they differentiate to long-slender bloodstream form (BSF) cells. These cells are actively replicating and have a reduced mitochondrial proteome, mainly relying on glycolysis for ATP generation via specialized organelles called glycosomes. These bloodstream form cells undergo antigenic variation of variant surface glycoproteins (VSGs) throughout the parasite population in order to evade the mammalian immune system [Reviewed in 8]. Long-slender BSF cells use a quorum sensing mechanism to trigger differentiation into short stumpy BSF cells [9].These short stumpy cells are transmissible to the tsetse fly during a blood meal and then differentiate into procyclic form cells in the midgut of the fly.

Recently, new metabolic forms of *T. brucei* have been described: those that can survive in dermis or adipose tissue of mammalian hosts. One of the subspecies of *T. brucei* that can infect humans, *T. b. gambiense*, has been found in dermal tissue of human patients, even when parasitemia was not detectable in the bloodstream [10]. In mice, a form of *T. b. brucei* has been found to be functionally adapted to living in adipose tissue by using fatty acids as a carbon source, as opposed to glucose used by BSF cells [11] (Figure 1.1). These discoveries lead to new questions about our understanding of the trypanosome lifecycle, such as: Are these metabolic forms necessary facets of the lifecycle and transmission cycle of these parasites, or are their specific circumstances in the mammalian hosts that lead to this differentiation? And finally: Will they need to be treated differently because of their differing metabolism?

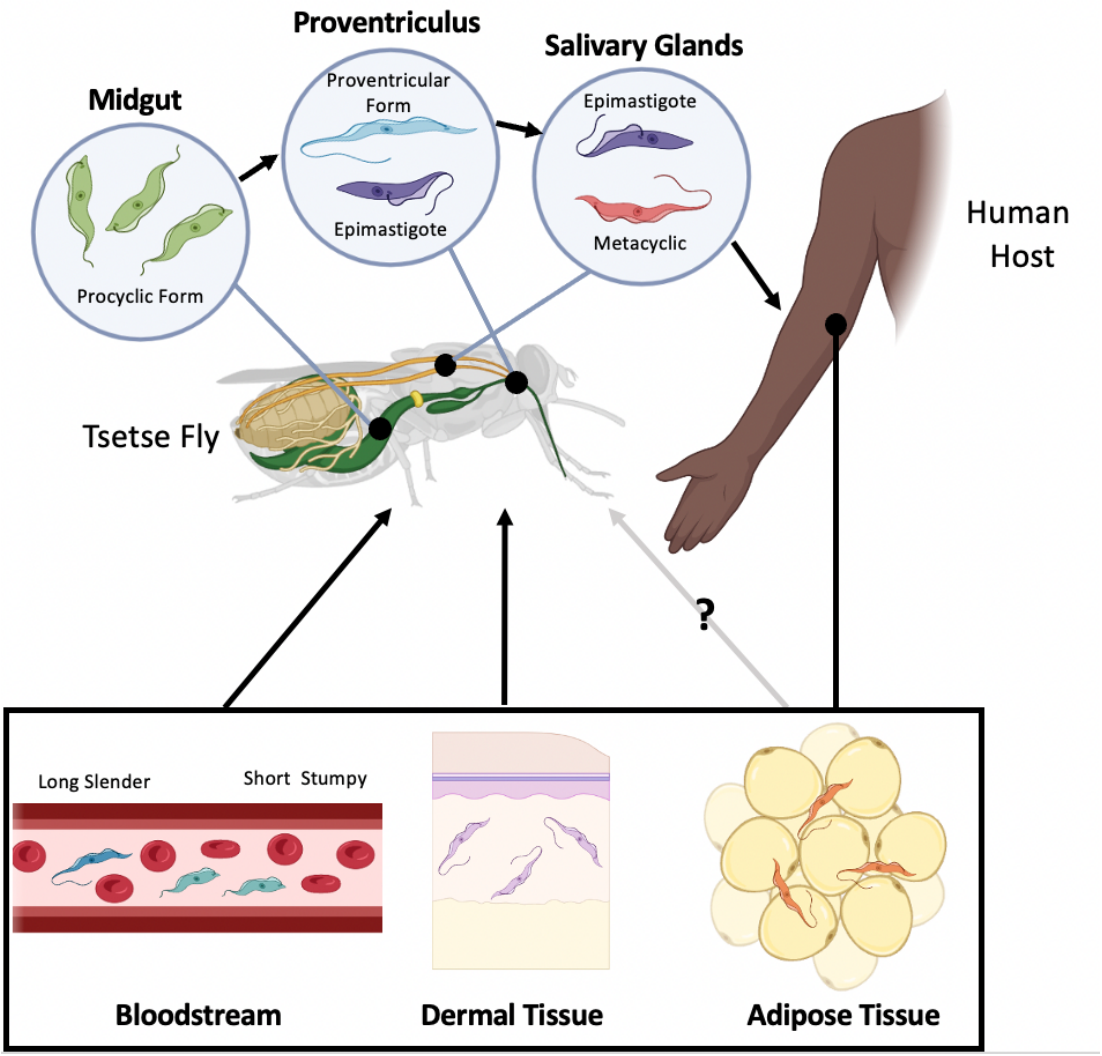


Figure 1.1 Lifecycle and Anatomical Reservoirs of human-infecting *T. brucei.* Procyclic parasites replicate in the midgut of the Tsetse fly, and migrate to the proventriculus where they differentiate into epimastigotes and metacyclic cells. In the salivary glands, epimastigotes adhere to epithelial tissue and differentiate into metacyclics, which are then able to infect a human host. In the bloodstream, long slender forms replicate until they differentiate into short stumpy form cells that can be taken up by a Tsetse fly. *T. brucei* can be found in dermal tissue in patients that may not have detectable loads of parasites in the bloodstream. These dermal parasites can be transmitted to tsetse flies and then new hosts. Adipose tissue forms may also represent an anatomical reservoir of trypanosomes, but more evidence is needed to understand the importance of this form for transmission (Biorender).

1.3 African Trypanosome Transmission Reservoirs

The subspecies of *Trypanosoma brucei* that are able to cause disease in humans, *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense,* are to establish infections in other mammalian hosts [12] These subspecies are adapted to evade the human Trypanosome lytic factors (TLF), a primate-specific innate immune response whereas *T. brucei brucei* remain susceptible [13]. It has been a long-term public health goal to eliminate *T. b. rhodesiense* and *T. b. gambiense* as a threat through treatment of disease in humans and passive monitoring for infections [14]. However, *Trypanosoma brucei* is capable of persisting in wildlife animal reservoirs making complete elimination incredibly challenging. Without strong surveillance, the parasite can and will likely be reintroduced into human populations [15].

The discovery of dermal and adipose-form trypanosomes represent a potential complication for treating human and animal trypanosomiasis. Trypanosomes in the dermis can successfully b, but there is very little data about whether successful established treatments are capable of clearing trypanosomes from these tissues [10].

African trypanosomes that cause disease in livestock include *Trypanosoma brucei brucei*, *T. congolense*, and *T. vivax*. Animal African Trypanosomiasis is estimated to cost African Agriculture 4.5 billion dollars annually [16]. These species of trypanosomes can be more challenging to control, as they have a broad host range, infecting a wide variety of mammals found in the wild in addition to domesticated livestock such as cattle and pigs. These trypanosome populations are able to persist in wildlife reservoirs and lead to a

source of infection for livestock, making AAT especially difficult to control. Strategies to address this problem include exciting innovation including efforts to clone Kenyan Boran that would be capable of clearing trypanosome infections [17]. Further effort should also be made to discover trypanosome specific drugs with low host toxicity. One avenue to pursue this goal is to study the divergent processes of these parasites to find new potential drug targets.

1.4 Mitochondrial DNA Replication in *T. brucei* is a Unique and Enigmatic Process.

Trypanosomes have one of the most complex mitochondrial DNA (mtDNA) structures in nature. In *T. brucei,* thousands of 1 kb minicircles and dozens of 23 maxicircles are catenated into a single network forming a single nucleoid that is condensed into a hockey-puck shaped structure *in vivo* and is called kinetoplast DNA (kDNA). Similar kDNA networks are found in other medically relevant trypanosomatid species including *Trypanosoma cruzi*, the causative agent of Chaga's disease, and *Leishmania* species that cause visceral, cutaneous and mucocutaneous Leishmaniasis. Studying the divergent kDNA replication proteins may lead to the discovery of highly specific drugs that can also target these neglected tropical diseases.

Replication of the kDNA network occurs once per cell cycle, coordinated with nuclear S phase, unlike in mammalian cells, where mtDNA replication occurs continuously throughout the cell cycle [Reviewed in 18, 19]. Minicircle replication is better understood than maxicircle replication, in part because minicircles are released from the catenated network to be replicated. This is a convenient mechanism for kDNA replication researchers as it allows for the tracking of minicircle replication intermediates using both

in vitro techniques such as Southern blot analysis as well as *in vivo* techniques such as FISH. The distinct process of minicircle replication includes a release and attachment mechanism that separates early replication events spatially and temporally. In *T. brucei*, 1 kb minicircles are released from the catenated network into the KinetoFlagellar Zone (KFZ) located between the kDNA disc and the mitochondrial membrane closest to the flagellar basal body [20]. Each minicircle contains a conserved 12-nucleotide sequence, the universal minicircle sequence (UMS), on the light DNA strand as well as a conserved 6-nucleotide sequence on the heavy strand. These regions are likely associated with replication initiation [21]. Universal minicircle sequence binding protein (UMSBP) is found in the KFZ and is thought to initiate minicircle replication by binding to these conserved DNA sequences [22]. Replication proceeds via unidirectional theta structures. Free minicircles localize to the antipodal sites, protein rich regions that are found 180 degrees from each other on either side of the kDNA disc [23, 24]. A suite of kDNA replication proteins isfound at the antipodal sites including DNA polymerases, primases, helicases, and topoisomerases, among others. (Figure 1.2). One of these topoisomerases, $TopoII_{mt}$, reattaches newly replicated minicircles back into the network at these sites [23]. Maxicircles are not released from the network for replication, and the mechanism for how these 23kb molecules are replicated while attached to the network remains largely undefined. The kDNA disc is physically linked to the flagellar basal body via the tripartite attachment complex (TAC), that crosses both the inner and outer mitochondrial membranes [25]. During kDNA replication, the flagellar basal body is duplicated and a second TAC attaches the new basal body to the network, allowing for the segregation of two daughter kDNA networks [25].

These properties of the kDNA structure and replication process in trypanosomes are highly divergent from those in animal cells. The complexity of kDNA replication is also reflected in the many additional factors not present in mammalian mitochondrial nucleoids [18,19]. In most well-studied eukaryotes, mtDNA is replicated and repaired by a single DNA polymerase (Pol), Pol γ. The animal mtDNA replisome consists of a minimal set of DNA replication proteins: TWINKLE DNA helicase, single-stranded binding protein (SSBP1), Three topoisomerases, and RNase H1 [Reviewed in 26]. In *T. brucei*, however, there are two Family X pols (TbPOLβ and TbPOLβ-PAK) and four Family A Pols that localize to the mitochondrion (TbPOLIA, IB, IC, ID). The mitochondrial DNA replisome is likely much more complex in trypanosomes than in mammalian cells, utilizing the large set of DNA transaction proteins in order to successfully replicate two different DNA substrates.

In addition to the multiplicity of polymerases, there are other families of DNA replication paralogs that localize to the Trypanosome mitochondrion whose precise roles have been difficult to extricate. Among these groups of paralogs are two ligases, (Ligkα and Ligkβ), two primases (PRI1 and PRI2) and six Pif1-like helicases [28,29, 30, 31].

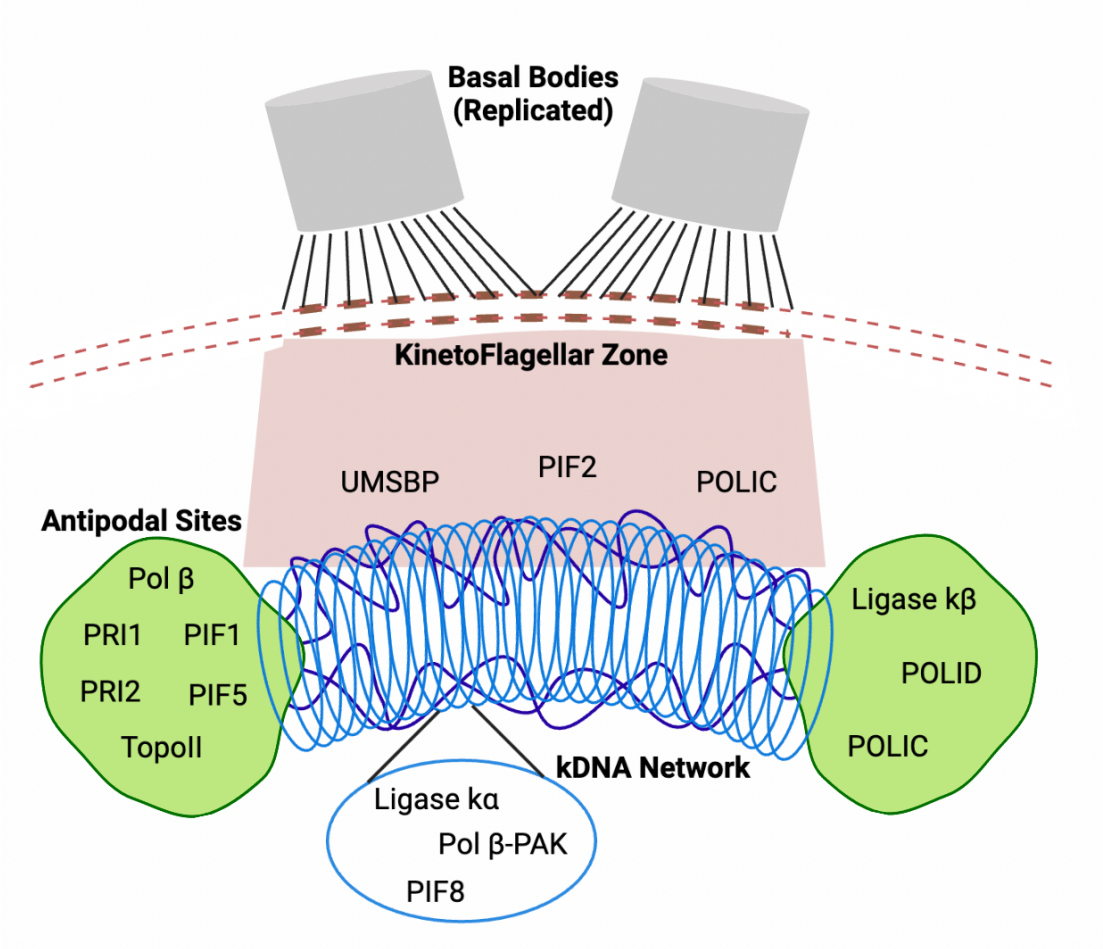


Figure 1.2 Localization of kDNA maintenance proteins during kDNA S-phase. kDNA minicirlces, light blue. kDNA maxicircles, dark blue. Proteins that localize to both antipodal sites shown in the green regions, protein that localize in the KFZ are in the pink region, while proteins localized within the network are within the light blue circle (Biorender, modified, from Michele Klingbeil).

1.5 kDNA is an Established Drug Target for *T. brucei*

kDNA maintenance is essential for the survival and successful transmission of African trypanosomes [32]. This structure represents an attractive drug target for trypanosomes, as its divergence from any structure in mammalian hosts could allow for highly selective targeting of parasites and minimize off-target effects. Drugs that impact kDNA maintenance have already been in use and have historically been effective ways to treat

trypanosome infections, even before the mechanism of action was well understood. Ethidium bromide, commonly known for its uses in molecular biology applications, was developed in 1952 as an anti-trypanosomal therapy to treat infected cattle and is still in use today [33, 34].

Ethidium bromide is an intercalating agent that impacts both nuclear and kDNA replication, but impacts kDNA at 10-fold lower concentrations than nuclear DNA by blocking minicircle replication initiation, indicating kDNA is the major target of this treatment [35]. Ethidium bromide provides us with an example of a drug that can clear *T. brucei* infections by interrupting kDNA replication, although its lack of selectivity is associated with a risk of mutagenesis in the host. A drug that targets kDNA more selectively could effectively kill *T. brucei* with minimal impact on a mammalian host.

Recently, Schnaufer and colleagues developed an assay comparing the impact of potential drugs on a *T. brucei* cell line that lacked kDNA (dyskinetoplastid) to a cell line that had intact kDNA networks [35]. These data were used to screen for drugs that specifically target kDNA maintenance. Ethidium bromide was used as a control in this study and remained the most kDNA selective agent tested [35]. Further screens of drug libraries or small molecules using this high-throughput assay may be able to identify other kDNA-selective compounds for use as pan trypanocidals to treat human and animal disease.

Pentamidine is one of six drugs used to treat HAT, particularly in the earlier stage of the disease. Pentamidine accumulates in *T. brucei* and is able to bind to the minor grooves of

DNA [36]. When treated with pentamidine, *T. brucei* loses mitochondrial membrane potential and kDNA is progressively depleted [37]. Loss of mitochondrial membrane potential could be due to the loss of kDNA and critical electron transport chain subunits in encodes. Although pentamidine may be target multiple pathways, there is evidence that this is another example of a successful trypanocidal therapy that targets kDNA maintenance, perhaps through blocking topoisomerase II function [37, 38]. These drugs provide a critical proof of concept that targeting kDNA associated processes is a viable strategy in treating both AAT and HAT. However, effort must be put into developing and discovering new drugs with fewer side effects and lower risk of resistance.

1.6 Understanding The Division of Labor in kDNA Replication Protein Paralogs

1.6.1 Pol Beta

Pol β is an X-family DNA polymerase involved in the gap-filling step of the base-excision repair (BER) pathway in nuclei of higher eukaryotes, including humans. This pathway is critical for the removal of oxidized, methylated, and deaminated bases, including damage caused by oxidative stress. In some trypanosomatid species, Pol β homologs localize exclusively to the mitochondrion and appear to be contributing to kDNA replication in addition to potential repair roles.

Pol β in *Crithidia fasciculata* was the first described mitochondrial Pol β like enzyme [39]. CfPol β was demonstrated to have low processivity and low fidelity when performing nucleotidyl incorporation, supporting the hypothesis that is a gap-filling (repair) enzyme [40]. Upon genome sequencing of *T. brucei*, a second Pol β homolog

was found that also localizes to the mitochondrion, dubbed Pol β-PAK for an N-terminal insertion rich in proline, alanine, and lysine residues [41] Biochemical studies have shown both in *T. brucei* and *T. cruzi* that Pol β and Pol β-PAK are capable of canonical Pol β BER repair activities as dRP-lyase activity and nucleotidyl incorporation. Interestingly, Pol β and Pol β-PAK have different optimal conditions for polymerization. In both, Pol β-PAK is more tolerant to a range of pH conditions than Pol β. In *T. cruzi* the primary structure of Pol β-PAK varies from *T. brucei* Pol β-PAK, as it as a smaller N-terminal insertion that does not include enrichment in the nominal, proline, alanine, and lysine residues. [41,42]. The role of Pol β may not be conserved across kinetoplastids- in *Leishmania infantum*, Pol β is a nuclear enzyme indicating it is playing a distinct role from its kDNA gap-filling/repair role described in the *Trypanosoma* species [43].

While Pol β is found in the antipodal sites during replication in *T. brucei* (Figure 1.2), Pol β-PAK is found within the kDNA network. This localization pattern contributes to the model in which Pol β is involved in gap-filling between Okazaki fragments, while Pol β-PAK fills in the last remaining gap in each minicircle once network replication is completed. In an investigation designed to identify components of the TAC segregation machinery, Pol β PAK was enriched in an immunoprecipitation of TAC102 suggesting an association of these proteins *in vivo* [44]. It is unclear what the functional mechanism of this interaction may be without further study.

Pol β in *T. cruzi* is also localized to the antipodal sites of the kDNA has a role in repairing oxidative damage of kDNA. *T. cruzi* cells overexpressing Pol β survive better

than a control when treated with peroxide [45]. Interestingly, in these peroxide-treated cells a Pol β foci forms in the KFZ, suggesting the enzyme translocates to participate in a repair role vs a replication role at the antipodal sites [45]. Treatment of *T. cruzi* cells with peroxide results in higher levels of Pol β and an increase in a phosphorylated, active form of the enzyme perhaps suggesting a mechanism for translocation to the KFZ for repair [46].

More than a decade after Pol β enzymes were first described in kinetoplastid organisms, Pol Beta has now been described as playing a repair role in mammalian mitochondria in addition to its nuclear role [47]. Data regarding the contributions of Pol β to kDNA replication in trypanosomatid organisms may inform hypotheses for how Pol β works in other systems.

1.6.2 kDNA Helicases

In mammalian mitochondria, TWINKLE is the only helicase required for mtDNA replication, a hexamer that is part of the SF4 superfamily [48]. No helicase that belongs to this superfamily is present in the mitochondrion of trypanosomatids, and so the current model for kDNA replication relies on the use of Pif1-like helicases. Pif1-like helicases are ATP-dependent helicases found across domains of life, from bacterial species through eukaryotes, belonging to the SF1 superfamily of helicases. These helicases are well studied in *Saccharomyces cerevisiae*, where there are two paralogs: ScRrm3 and the founding member of the protein family, ScPif1 [Reviewed in 49]. Other well-studied eukaryotes encode for a single Pif1-like helicase including humans (HsPif1). These

proteins localize to both the nucleus and mitochondria to participate in DNA repair and recombination. *Trypanosoma brucei* and *Trypanosoma cruzi* each encode for 8 Pif1- like helicases, while *Leishmania major* encodes for 7 Pif1 like paralogs [50]. These helicases are best studied in *Trypanosoma brucei* which has eight different Pif1-like genes, denoted as TbPif1-8. It has been established that six of these proteins localize to the mitochondrion: TbPIF1, 2, 4, 5, 7, and 8 [31].

TbPif1 localizes to the antipodal sites in *T. brucei*, and RNAi of Pif1blocks minicircle replication, loss of fitness, and results in buildup of fraction U [31]. This data is consistent with a role for Pif1 at the minicircle replisome. Experiments with TbPif2 clearly demonstrate its role in regulating the quantity of maxicircles in the network. The amount of Pif2 expressed is positively correlated to the number of maxicircles as demonstrated through RNAi, where maxicircle copy number decreases, and ectopic overexpression of TbPif2, where maxicircle copy number increases [31]. These experiments represent the most definitive data of a kDNA replication enzyme that is dedicated solely to maxicircle replication. Pif2 should be considered a tool for finding functionally related proteins to better understand the maxicircle replisome.

Pif5 is predicted to facilitate Okazaki fragment processing by unwinding RNA primers after lagging strand synthesis, consistent with its antipodal site localization [31]. However, no loss of fitness is seen when Pif5 is depleted by RNAi. Pif8 localized to distal side of the network or throughout the kDNA network, potentially in a cell cycle-stage dependent manner [51].

An alternative pathway for Okazaki fragment removal described in *S. cerevisiae* involves the lengthening of DNA flaps to then be cleaved, facilitated by Pif1 unwinding [52]. Okazaki fragments in minicircle replication could demand such an alternative processing pathway, and may be the role of one or more kinetoplastid Pif1-like helicases.

Aglycone-NDI 7, a G quadruplex ligand that has been proposed for use as an anti-trypanosomatid therapy, localizes in both the nucleus and kinetoplast of *T. brucei,* indicating that G4 structures are likely present in the DNA of both organelles [53]. In *S. cerevisiae* Pif1 is important for suppressing genome instability at G4 sites and for progression of lagging strand synthesis through G4 sequences [54, 55]. It is therefore possible that one or more of the trypanosomatid Pif1-like proteins is involved in maintenance of G-quadruplex structures in the kinetoplast, but this has not yet been investigated. More detailed studies of these helicases would reveal whether they are participating in these proposed roles for the proteins.

1.6.3 kDNA Ligases

In *C. fasciculata* as well as *T. brucei*, two paralogous ligases, Lig kα and Lig kβ localize to the mitochondrion. In both organisms, Lig kβ localizes to the antipodal sites [28, 56]. In *T. brucei*, Lig kα localizes throughout the kDNA network. Knockdown of Lig kβ does not result in a loss of fitness phenotype, however knockdown of Lig kα does result in loss of fitness and a kDNA replication defect that includes shrinkage of the network [28]. Based on these results it is predicted that Lig kα is sealing newly replicated minicircles after they are reattached into the network.

1.6.4 kDNA Primases

Two mitochondrial primases both localize to the antipodal sites in *T. brucei,* PRI1 and PRI2 [29,30]. When PRI1 is knocked down, cells have a loss of fitness phenotype and lose maxicircles over time [29]. Cells also lose fitness when PRI2 is knocked down, and lose both minicircles and maxicircles [30]. Bases on these studies, PRI1 is likely the primase involved in maxicircle replication while PRI2 is likely involved in minicircle replication.

1.7 Variations on a Theme: Family A Polymerase Structures

In order to make hypotheses and mechanistic models for how trypanosomatid organisms are replicating their complex mitochondrial nucleoid, we look to well-characterized organellar DNA polymerases. In animals there is a single family A DNA polymerase (Pol) essential for mtDNA replication, Pol γ, a heterotrimer made up of catalytic subunit Pol γA and a dimeric subunit that confers processivity of the holoenyzme, Pol γB. Pol γA has a typical structure for Family A Pols that includes finger, palm, and thumb subdomains. In the thumb subdomain, there is a large spacer region (310 residues) that is divergent from other polymerases and aids in increasing the processivity of the enzyme through the intrinsic processivity subdomain of the spacer and the accessory interacting determinant subdomain that serves as an interface with Pol γB [57]. Another well-characterized Family A pol, T7 DNA polymerase, has a 76 aa domain that is located at the end of a thumb subdomain helix and allows for binding to thioredoxin. This interaction improves the processivity of the enzyme [58].

Another example of an organellar DNA replication system occurs in flowering plants. *Arabidopsis thaliana* has two family A polymerases for organellar DNA replication and repair, the plant organellar polymerases or POPs. One of these pols, AtPOLIA, has higher fidelity for replication, while the other *A. thaliana* POP, AtPOLIB, has lower fidelity and insertions within its polymerase domain that allow repair activities [Reviewed in 59]. Based on this enzymatic activity, it is expected that AtPOLIA is acting as the replicative pol while AtPOLIB is playing a support role in repair. Both of these POPs also have insertions in their thumb domains that facilitate the ability to participate in microhomology mediated end joining repair of double strand breaks [60]. Variations in the thumb subdomain of Family A pols is well-documented to impact the functions of these enzymes.

Recent advances in protein structure prediction allows us to use tools such as Alphafold2 and Rosettafold to gain high quality structural information about proteins without crystal structures. Predicted protein structures for kDNA replication proteins, including the Pol I-like enzymes of both *T. cruzi* and *T. brucei*, are deposited in the Alphafold database [61, 62]. Family A polymerases typically have a highly conserved right hand structure, but the protein structure predictions of the trypanosomatid Pol I-like enzymes revealed striking divergence from this canonical structure. Alphafold models of POLIB from *T. brucei* reveal a large insertion that projects from the helices of the thumb subdomain. The conserved catalytic active sites of a DEDDh exonuclease are found in this thumb insertion, a feature not seen in any other characterized DNA polymerase (Figure 1.3).

When the pol domains of POLIB from *Leishmania donovani* and *Crithidia fasciculata* are modelled using a similar protein structure prediction tool Rosettafold, they too have this exonuclease thumb insertion (Figure 1.4). The conservation of this unique structure across trypanosomatid organisms suggests that it may be a key feature of POLIB for its role in kDNA replication. Interestingly, the UCR domains of the *T. cruzi* and *T. brucei* POLIB Alphafold models are the part of the homologs that are most structurally different from each other, suggesting the roles of this domain may vary between the two species. The protein structure predictions for *T. brucei* POLIC and POLID also reveal large thumb insertions in these proteins ~489 aa in POLID (Figure 1.3). However, the insertions in these structures have no homology to enzyme active sites as in POLIB. Structure-function studies will need to be executed in the future to understand the role of the insertions in POLIC and POLID. As in other DNA polymerases with thumb insertions, these domains could contribute to protein-protein interactions, increasing processivity, or conferring specialized repair activities.

These predictions have provided a foundation for better understanding the divergent enzymes of trypanosmatid organisms, although predictions for these species are typically lower quality than those for model organisms [63]. Richard Wheeler demonstrated improvement of trypanosomatid protein structure predictions when using Colabfold, an open tool for Alphafold, supplemented with an HMMER search of the Discoba database [63]. Implementing this strategy will help improve the quality of models used to study the divergent proteins of these organisms.

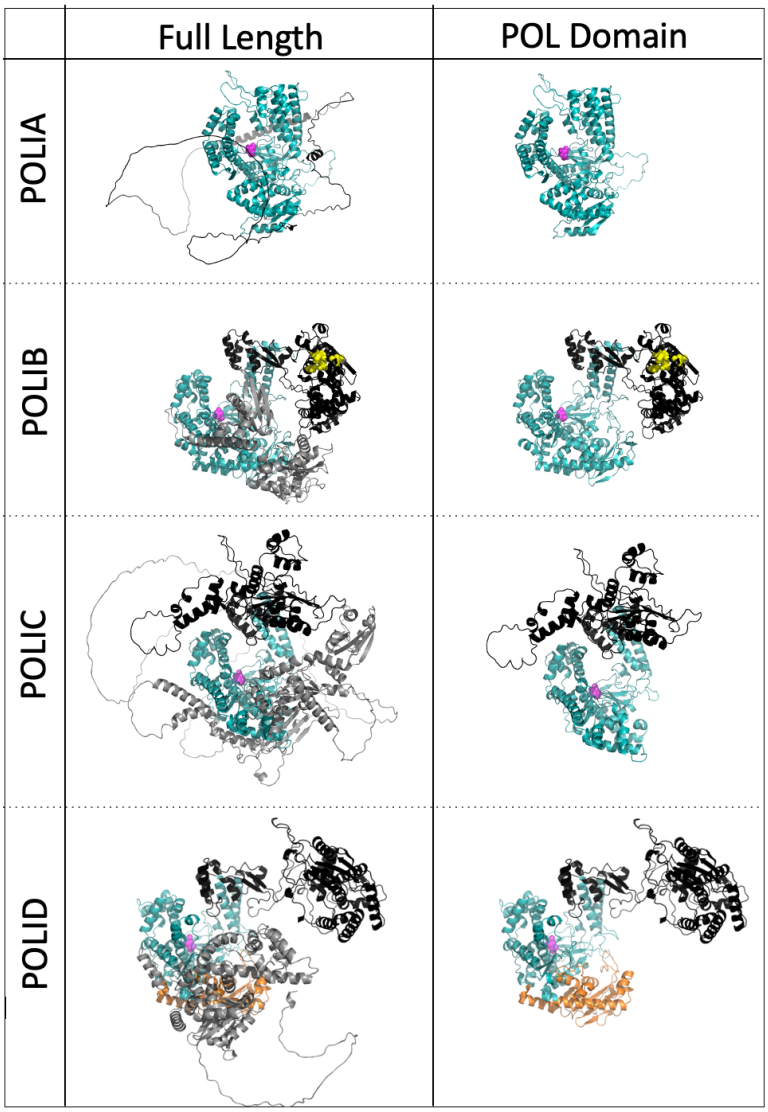


Figure 1.3: Alphafold structure prediction of the right-hand structure of the *Trypanosoma brucei* PolI-like paralogs. Left: Full length protein structure prediction, Right: N-terminal UCRs not displayed (POLIA amino acids 253-958 POLIB amino acids 413-1404, POLIC amino acids 760-1646, POLID amino acids 501-1629) Teal: right-hand structures of the conserved polymerase domains. Black: insertions found in pol domains. Yellow: conserved catalytic residues of a DEDDh exonuclease found in the POLIB pol insertion. Magenta: conserved catalytuc aspartic acid residues of pol domain. Orange: POLID 3'-5' exonuclease domain.

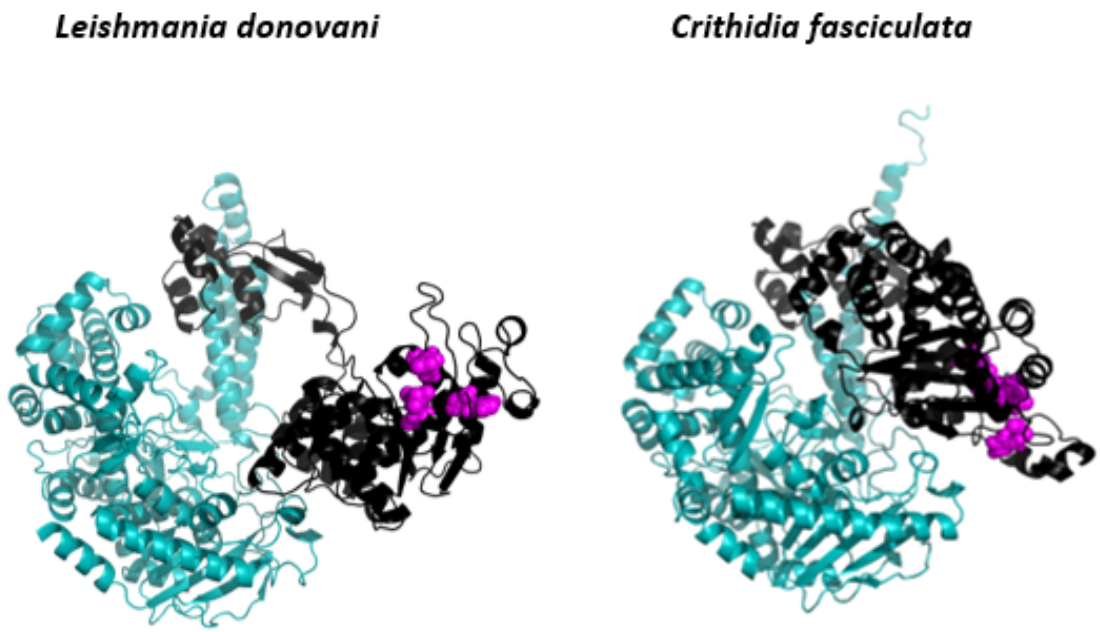


Figure 1.4: Rosettafold Predicted Protein Structures of Leishmania donovani and Crithidial fasciculata POLIB pol domains. N-terminal UCRS not shown. Teal: Right-hand structure of the pol domain, Black: thumb insertion. Magenta: conserved exonuclease DEDDh residues.

1.8 Developments in Understanding *T. brucei* Pol I-Like Polymerases

1.8.1 Past

There are four Pol I-like proteins that localize to the mitochondrion in *T. brucei*, POLIA, POLIB, POLIC, POLID [64]. Homologs of these four polymerases can be found in other medically relevant trypanosomatids such as *T. cruzi* and *Leishmania* species. The trypanosomatid mitochondrial Pol I-like proteins each contain a C-terminal POLA like domain and N-terminal regions that have no homology to any other characterized enzymes, termed the uncharacterized region (UCR). Two of these proteins, POLIB and POLID also have predicted 3′-5′ exonuclease domains that are hypothesized to be

proofreading domains. Typically, a proofreading domain in a pol confers higher fidelity, one of the hallmarks of a replicative pol.

Following the model of the POPs, one could predict that one of these proteins is playing a replicative role, while the others are contributing to repair and maintenance. This hypothesis is complicated by RNAi studies that demonstrate three of the PolI-like proteins POLIB, POLIC, and POLID are essential for successful network replication and are non-redundant in function [64, 65, 66, 67]. Knockdown of each of these proteins in procyclic cells results in loss of fitness alongside canonical minicircle replication defects. The impact of RNAi of each of these pols is characterized by distinct phenotypic hallmarks depending on which transcript was depleted.

POLIB knockdown in procyclic cells results in an increase in cells with no or small kDNA, and an increase in the ratio of covalently closed minicircles to newly replicated nicked/gapped minicircles. The formation of Fraction U, a heterogenous mix of multiply interlocked, covalently closed minicircles was also observed when POLIB is depleted. There was also an increase in the amount of linearized kDNA molecules in DNA isolated from cells that were induced for POLIB RNAi, although it is unclear if this was due to the DNA being more fragile in these conditions or some other mechanism [66].

RNAi of POLIC resulted in cells with a small kDNA phenotype, but only a small portion of cells that had no detectable kDNA, and the quantity of minicircle replication intermediates increases through the course of the induction [64]. RNAi of POLID results in cells with small or no kDNA, and a characteristic phenotype in which both

nicked/gapped and covalently closed minicircle species transiently increase at days 2 and 3 of induction and then dramatically decrease [65]. POLIB, POLIC, and POLID are also essential for kDNA replication in bloodstream form parasites. Knockdown in these cells, the disease-causing form of the parasite, also results in loss of fitness and depletion of kDNA, providing good evidence for the potential of these proteins as drug targets [67]. These studies indicate that each of these enzymes is independently essential for kDNA replication, but do not give us enough mechanistic information to understand how.

1.8.2 Present

In the last several years, elegantly designed studies and the advent of new techniques have provided more detailed data the Pol I-like kDNA replication proteins. Immunofluorescence studies revealed cell cycle dependent localization for two of these proteins: POLID and POLIC. POLID localizes to the antipodal sites very early in kDNA S phase and redistributes throughout the mitochondrial matrix in the late stages of kinetoplast division, suggesting it is recruited to these sites for its role in kDNA replication [68]. Recently, POLID was appeared in a screen for TAC components using TAC102 via BioID proximity-labeling experiment, suggesting there may be a functional interaction between replication and segregation machinery [69].

POLIC also exhibits cell cycle localization, appearing in the KFZ as kDNA S phase begins, then co-localizing with POLID in the antipodal sites in mid to late kDNA S-phase [70]. Recently it was reported that a truncation of POLIC lacking the N-terminal UCR does not form these foci, and is instead distributed throughout the mitochondrial matrix.

This study provided the first evidence that the UCR of one of these POLI-like proteins is key to successful localization. POLIC was also demonstrated to be involved in the correct distribution of the kDNA network in addition to its role in minicircle replication [71].

POLIB is predicted to physically interact with DRBD10 (a putative RNA binding protein) and putative kinetoplast associated protein 3 (KAP3) based on a machine-learning study of protein complexes that co-purified in size-exclusion chromatography (SEC) and strong anion exchange (SAX) liquid chromatography protocols [72]. POLIB and KAP3 both have sites of arginine methylation [73] which could provide a mechanism to regulate the association of these proteins. There are many open questions related to how these Pol I like enzymes contribute to kDNA replication. This study is particularly interested in answering some of these questions by addressing the enzymatic features of one of these critical kDNA enzymes, POLIB.

1.9 Contribution and Importance of This Study

In order to investigate exonuclease and polymerase activities of the divergent structures of POLIB recombinant protein was used to determine the optimal conditions and substrates protein, and how its unique thumb insertion participates in its activity. Understanding the enzymatic capabilities of recombinant POLIB has also led to new models for how the three essential kDNA polymerases are each contributing to kDNA replication. Evidence for the enzymatic activities of this enzyme also allows for the possibility of strategically targeting these activities with drugs. Similar biochemical analyses for POLIC and POLID will provide better evidence for their enzymatic

capabilities that could also lead to more specific models of kDNA replication and repair and could also lead to these pols serving as valuable drug targets.

1.10 Goals of this Study

In this study, we set out to define biochemical characteristics of *Trypanosoma brucei* mitochondrial DNA polymerase POLIB. Based on the RNAi study of this gene, we hypothesized that POLIB is playing a role in minicircle replication. To understand more in depth how POLIB may be contributing to this process, we tailored experiments to understand how it may function *in vivo*, with an emphasis on understanding whether POLIB was likely a replicative kDNA pol. Our goals to address these questions were to:

1) Test whether the divergent insertion predicted to be within the thumb subdomain of POLIB is an active exonuclease domain, and if so, whether is acts preferentially on mismatched nucleotides for proofreading.

2) Determine optimal conditions for POLIB pol and exo activities including optimal pH, salt concentrations, and substrates.

3) Determine whether the enzymatic qualities of POLIB are compatible with the hypothesis that it is a replicative DNA polymerase involved in kDNA replication. To fit this role, we expected that DNA polymerase activity of POLIB would have high processivity and fidelity. We also expected that the fidelity of POLIB would be improved by its predicted 3'-5' exonuclease domain activity.

1.11 References

1. Eshetu, E.; Begejo, B. The Current Situation and Diagnostic Approach of Nagana in Africa: A Review. **2015**.
2. Kennedy, P. G. Clinical Features, Diagnosis, and Treatment of Human African Trypanosomiasis (Sleeping Sickness). *Lancet Neurol* **2013**, *12* (2), 186–194. https://doi.org/10.1016/S1474-4422(12)70296-X.
3. *WHO Interim Guidelines for the Treatment of Gambiense Human African Trypanosomiasis*; WHO Guidelines Approved by the Guidelines Review Committee; World Health Organization: Geneva, 2019.
4. Álvarez-Rodríguez, A.; Jin, B.-K.; Radwanska, M.; Magez, S. Recent Progress in Diagnosis and Treatment of Human African Trypanosomiasis Has Made the Elimination of This Disease a Realistic Target by 2030. *Front Med (Lausanne)* **2022**, *9*, 1037094. https://doi.org/10.3389/fmed.2022.1037094.
5. Rotureau, B.; Van Den Abbeele, J. Through the Dark Continent: African Trypanosome Development in the Tsetse Fly. *Front Cell Infect Microbiol* **2013**, *3*, 53. https://doi.org/10.3389/fcimb.2013.00053.
6. Lamour, N.; Rivière, L.; Coustou, V.; Coombs, G. H.; Barrett, M. P.; Bringaud, F. Proline Metabolism in Procyclic Trypanosoma Brucei Is Down-Regulated in the Presence of Glucose. *J Biol Chem* **2005**, *280* (12), 11902–11910. https://doi.org/10.1074/jbc.M414274200.
7. Graham SV, Barry JD. Transcriptional regulation of metacyclic variant surface glycoprotein gene expression during the life cycle of Trypanosoma brucei. *Mol Cell Biol*. 1995;15: 5945–5956. doi: 10.1128/MCB.15.11.5945
8. Sima, N.; McLaughlin, E. J.; Hutchinson, S.; Glover, L. Escaping the Immune System by DNA Repair and Recombination in African Trypanosomes. *Open Biol.* **2019**, *9* (11), 190182. https://doi.org/10.1098/rsob.190182.
9. MacGregor P, Savill NJ, Hall D, Matthews KR. Transmission stages dominate trypanosome within-host dynamics during chronic infections. Cell Host Microbe. 2011 Apr 21;9(4):310-8. doi: 10.1016/j.chom.2011.03.013. PMID: 21501830; PMCID: PMC3094754.
10. Capewell, P.; Cren-Travaillé, C.; Marchesi, F.; Johnston, P.; Clucas, C.; Benson, R. A.; Gorman, T.-A.; Calvo-Alvarez, E.; Crouzols, A.; Jouvion, G.; Jamonneau, V.; Weir, W.; Stevenson, M. L.; O'Neill, K.; Cooper, A.; Swar, N.-R. K.; Bucheton, B.; Ngoyi, D. M.; Garside, P.; Rotureau, B.; MacLeod, A. The Skin Is a Significant but Overlooked Anatomical Reservoir for Vector-Borne African Trypanosomes. *Elife* **2016**, *5*, e17716. https://doi.org/10.7554/eLife.17716.
11. Trindade, S.; Rijo-Ferreira, F.; Carvalho, T.; Pinto-Neves, D.; Guegan, F.; Aresta-Branco, F.; Bento, F.; Young, S. A.; Pinto, A.; Van Den Abbeele, J.; Ribeiro, R. M.; Dias, S.; Smith, T. K.; Figueiredo, L. M. Trypanosoma Brucei Parasites Occupy and Functionally Adapt to the Adipose Tissue in Mice. *Cell Host Microbe* **2016**, *19* (6), 837–848. https://doi.org/10.1016/j.chom.2016.05.002.
12. Balmer, O.; Beadell, J. S.; Gibson, W.; Caccone, A. Phylogeography and Taxonomy of Trypanosoma Brucei. *PLoS Negl Trop Dis* **2011**, *5* (2), e961. https://doi.org/10.1371/journal.pntd.0000961.

13. Pays, E.; Vanhollebeke, B.; Uzureau, P.; Lecordier, L.; Pérez-Morga, D. The Molecular Arms Race between African Trypanosomes and Humans. *Nat Rev Microbiol* **2014**, *12* (8), 575–584. https://doi.org/10.1038/nrmicro3298.
14. Longbottom, J.; Wamboga, C.; Bessell, P. R.; Torr, S. J.; Stanton, M. C. Optimising Passive Surveillance of a Neglected Tropical Disease in the Era of Elimination: A Modelling Study. *PLoS Negl Trop Dis* **2021**, *15* (3), e0008599. https://doi.org/10.1371/journal.pntd.0008599.
15. Venturelli, A.; Tagliazucchi, L.; Lima, C.; Venuti, F.; Malpezzi, G.; Magoulas, G. E.; Santarem, N.; Calogeropoulou, T.; Cordeiro-da-Silva, A.; Costi, M. P. Current Treatments to Control African Trypanosomiasis and One Health Perspective. *Microorganisms* **2022**, *10* (7), 1298. https://doi.org/10.3390/microorganisms10071298.
16. Swallow, B. M. Impacts of Trypanosomiasis on African Agriculture. **1999.**
17. Yu, M.; Muteti, C.; Ogugo, M.; Ritchie, W. A.; Raper, J.; Kemp, S. Cloning of the African Indigenous Cattle Breed Kenyan Boran. *Anim Genet* **2016**, *47* (4), 510–511. https://doi.org/10.1111/age.12441.
18. Jensen, R. E.; Englund, P. T. Network News: The Replication of Kinetoplast DNA. *Annu Rev Microbiol* **2012**, *66*, 473–491. https://doi.org/10.1146/annurev-micro-092611-150057.
19. Amodeo, S.; Bregy, I.; Ochsenreiter, T. Mitochondrial Genome Maintenance - the Kinetoplast Story. *FEMS Microbiol Rev* **2022**, fuac047. https://doi.org/10.1093/femsre/fuac047.
20. Drew, M. E.; Englund, P. T. Intramitochondrial Location and Dynamics of Crithidia Fasciculata Kinetoplast Minicircle Replication Intermediates. *J Cell Biol* **2001**, *153* (4), 735–744. https://doi.org/10.1083/jcb.153.4.735.
21. Milman, N.; Motyka, S. A.; Englund, P. T.; Robinson, D.; Shlomai, J. Mitochondrial Origin-Binding Protein UMSBP Mediates DNA Replication and Segregation in Trypanosomes. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (49), 19250–19255. https://doi.org/10.1073/pnas.0706858104.
22. Abu-Elneel, K.; Robinson, D. R.; Drew, M. E.; Englund, P. T.; Shlomai, J. Intramitochondrial Localization of Universal Minicircle Sequence-Binding Protein, a Trypanosomatid Protein That Binds Kinetoplast Minicircle Replication Origins. *J Cell Biol* **2001**, *153* (4), 725–734. https://doi.org/10.1083/jcb.153.4.725.
23. Melendy, T.; Sheline, C.; Ray, D. S. Localization of a Type II DNA Topoisomerase to Two Sites at the Periphery of the Kinetoplast DNA of Crithidia Fasciculata. *Cell* **1988**, *55* (6), 1083–1088. https://doi.org/10.1016/0092-8674(88)90252-8.
24. Gluenz, E.; Shaw, M. K.; Gull, K. Structural Asymmetry and Discrete Nucleic Acid Subdomains in the Trypanosoma Brucei Kinetoplast. *Mol Microbiol* **2007**, *64* (6), 1529–1539. https://doi.org/10.1111/j.1365-2958.2007.05749.x.
25. Schneider, A.; Ochsenreiter, T. Failure Is Not an Option – Mitochondrial Genome Segregation in Trypanosomes. *Journal of Cell Science* **2018**, *131* (18), jcs221820. https://doi.org/10.1242/jcs.221820.
26. Kasiviswanathan, R.; Collins, T. R. L.; Copeland, W. C. The Interface of Transcription and DNA Replication in the Mitochondria. *Biochim Biophys Acta* **2012**, *1819* (9–10), 970–978. https://doi.org/10.1016/j.bbagrm.2011.12.005.

27. Shapiro, T. A. Mitochondrial Topoisomerase II Activity Is Essential for Kinetoplast DNA Minicircle Segregation. *Mol Cell Biol* **1994**, *14* (6), 3660–3667. https://doi.org/10.1128/mcb.14.6.3660-3667.1994.
28. Downey, N.; Hines, J. C.; Sinha, K. M.; Ray, D. S. Mitochondrial DNA Ligases of Trypanosoma Brucei. *Eukaryot Cell* **2005**, *4* (4), 765–774. https://doi.org/10.1128/EC.4.4.765-774.2005.
29. Hines, J. C.; Ray, D. S. A Mitochondrial DNA Primase Is Essential for Cell Growth and Kinetoplast DNA Replication in Trypanosoma Brucei. *Mol Cell Biol* **2010**, *30* (6), 1319–1328. https://doi.org/10.1128/MCB.01231-09.
30. Hines, J. C.; Ray, D. S. A Second Mitochondrial DNA Primase Is Essential for Cell Growth and Kinetoplast Minicircle DNA Replication in Trypanosoma Brucei. *Eukaryot Cell* **2011**, *10* (3), 445–454. https://doi.org/10.1128/EC.00308-10.
31. Liu, B.; Wang, J.; Yaffe, N.; Lindsay, M. E.; Zhao, Z.; Zick, A.; Shlomai, J.; Englund, P. T. Trypanosomes Have Six Mitochondrial DNA Helicases with One Controlling Kinetoplast Maxicircle Replication. *Mol Cell* **2009**, *35* (4), 490–501. https://doi.org/10.1016/j.molcel.2009.07.004.
32. Dewar, C. E.; MacGregor, P.; Cooper, S.; Gould, M. K.; Matthews, K. R.; Savill, N. J.; Schnaufer, A. Mitochondrial DNA Is Critical for Longevity and Metabolism of Transmission Stage Trypanosoma Brucei. *PLoS Pathog* **2018**, *14* (7), e1007195. https://doi.org/10.1371/journal.ppat.1007195.
33. Watkins, T. I.; Woolfe, G. Effect of Changing the Quaternizing Group on the Trypanocidal Activity of Dimidium Bromide. *Nature* **1952**, *169* (4299), 506–506. https://doi.org/10.1038/169506a0.
34. Roy Chowdhury, A.; Bakshi, R.; Wang, J.; Yildirir, G.; Liu, B.; Pappas-Brown, V.; Tolun, G.; Griffith, J. D.; Shapiro, T. A.; Jensen, R. E.; Englund, P. T. The Killing of African Trypanosomes by Ethidium Bromide. *PLoS Pathog* **2010**, *6* (12), e1001226. https://doi.org/10.1371/journal.ppat.1001226.
35. Miskinyte, M.; Dawson, J. C.; Makda, A.; Doughty-Shenton, D.; Carragher, N. O.; Schnaufer, A. A Novel High-Content Phenotypic Screen To Identify Inhibitors of Mitochondrial DNA Maintenance in Trypanosomes. *Antimicrob Agents Chemother* **2022**, *66* (2), e01980-21. https://doi.org/10.1128/AAC.01980-21.
36. Carter, N. S.; Berger, B. J.; Fairlamb, A. H. Uptake of Diamidine Drugs by the P2 Nucleoside Transporter in Melarsen-Sensitive and -Resistant Trypanosoma Brucei Brucei. *Journal of Biological Chemistry* **1995**, *270* (47), 28153–28157. https://doi.org/10.1074/jbc.270.47.28153.
37. Thomas, J. A.; Baker, N.; Hutchinson, S.; Dominicus, C.; Trenaman, A.; Glover, L.; Alsford, S.; Horn, D. Insights into Antitrypanosomal Drug Mode-of-Action from Cytology-Based Profiling. *PLoS Negl Trop Dis* **2018**, *12* (11), e0006980. https://doi.org/10.1371/journal.pntd.0006980.
38. Shapiro, T. A.; Englund, P. T. Selective Cleavage of Kinetoplast DNA Minicircles Promoted by Antitrypanosomal Drugs. *Proc Natl Acad Sci U S A* **1990**, *87* (3), 950–954. https://doi.org/10.1073/pnas.87.3.950.
39. Torri, A. F.; Englund, P. T. Purification of a Mitochondrial DNA Polymerase from Crithidia Fasciculata. *J Biol Chem* **1992**, *267* (7), 4786–4792.

40. Torri, A. F.; Kunkel, T. A.; Englund, P. T. A Beta-like DNA Polymerase from the Mitochondrion of the Trypanosomatid Crithidia Fasciculata. *J Biol Chem* **1994**, *269* (11), 8165–8171.
41. Saxowsky, T. T.; Choudhary, G.; Klingbeil, M. M.; Englund, P. T. Trypanosoma Brucei Has Two Distinct Mitochondrial DNA Polymerase Beta Enzymes. *J Biol Chem* **2003**, *278* (49), 49095–49101. https://doi.org/10.1074/jbc.M308565200.
42. Lopes, D. de O.; Schamber-Reis, B. L. F.; Regis-da-Silva, C. G.; Rajão, M. A.; DaRocha, W. D.; Macedo, A. M.; Franco, G. R.; Nardelli, S. C.; Schenkman, S.; Hoffmann, J.-S.; Cazaux, C.; Pena, S. D. J.; Teixeira, S. M. R.; Machado, C. R. Biochemical Studies with DNA Polymerase β and DNA Polymerase β-PAK of Trypanosoma Cruzi Suggest the Involvement of These Proteins in Mitochondrial DNA Maintenance. *DNA Repair* **2008**, *7* (11), 1882–1892.
43. Taladriz, S.; Hanke, T.; Ramiro, M. J.; García-Díaz, M.; García De Lacoba, M.; Blanco, L.; Larraga, V. Nuclear DNA Polymerase Beta from Leishmania Infantum. Cloning, Molecular Analysis and Developmental Regulation. *Nucleic Acids Res* **2001**, *29* (18), 3822–3834. https://doi.org/10.1093/nar/29.18.3822.
44. Baudouin, H. C. M.; Pfeiffer, L.; Ochsenreiter, T. A Comparison of Three Approaches for the Discovery of Novel Tripartite Attachment Complex Proteins in Trypanosoma Brucei. *PLoS Negl Trop Dis* **2020**, *14* (9), e0008568. https://doi.org/10.1371/journal.pntd.0008568.
45. Schamber-Reis, B. L. F.; Nardelli, S.; Régis-Silva, C. G.; Campos, P. C.; Cerqueira, P. G.; Lima, S. A.; Franco, G. R.; Macedo, A. M.; Pena, S. D. J.; Cazaux, C.; Hoffmann, J.-S.; Motta, M. C. M.; Schenkman, S.; Teixeira, S. M. R.; Machado, C. R. DNA Polymerase Beta from Trypanosoma Cruzi Is Involved in Kinetoplast DNA Replication and Repair of Oxidative Lesions. *Mol Biochem Parasitol* **2012**, *183* (2), 122–131.
46. Maldonado, E.; Rojas, D. A.; Urbina, F.; Solari, A. T. Cruzi DNA Polymerase Beta (Tcpolβ) Is Phosphorylated in Vitro by CK1, CK2 and TcAUK1 Leading to the Potentiation of Its DNA Synthesis Activity. *PLoS Negl Trop Dis* **2021**, *15* (7), e0009588. https://doi.org/10.1371/journal.pntd.0009588.
47. Sykora, P.; Kanno, S.; Akbari, M.; Kulikowicz, T.; Baptiste, B. A.; Leandro, G. S.; Lu, H.; Tian, J.; May, A.; Becker, K. A.; Croteau, D. L.; Wilson, D. M.; Sobol, R. W.; Yasui, A.; Bohr, V. A. DNA Polymerase Beta Participates in Mitochondrial DNA Repair. *Mol Cell Biol* **2017**, *37* (16), e00237-17. https://doi.org/10.1128/MCB.00237-17.
48. Peter, B.; Falkenberg, M. TWINKLE and Other Human Mitochondrial DNA Helicases: Structure, Function and Disease. *Genes (Basel)* **2020**, *11* (4), 408. https://doi.org/10.3390/genes11040408.
49. Muellner, J.; Schmidt, K. H. Yeast Genome Maintenance by the Multifunctional PIF1 DNA Helicase Family. *Genes (Basel)* **2020**, *11* (2), 224. https://doi.org/10.3390/genes11020224.
50. Bochman, M. L.; Sabouri, N.; Zakian, V. A. Unwinding the Functions of the Pif1 Family Helicases. *DNA Repair (Amst)* **2010**, *9* (3), 237–249. https://doi.org/10.1016/j.dnarep.2010.01.008.
51. Wang, J.; Englund, P. T.; Jensen, R. E. TbPIF8, a Trypanosoma Brucei Protein Related to the Yeast Pif1 Helicase, Is Essential for Cell Viability and Mitochondrial

Genome Maintenance. *Mol Microbiol* **2012**, *83* (3), 471–485. https://doi.org/10.1111/j.1365-2958.2011.07938.x.
52. Pike, J. E.; Henry, R. A.; Burgers, P. M. J.; Campbell, J. L.; Bambara, R. A. An Alternative Pathway for Okazaki Fragment Processing: Resolution of Fold-Back Flaps by Pif1 Helicase. *J Biol Chem* **2010**, *285* (53), 41712–41723. https://doi.org/10.1074/jbc.M110.146894.
53. Belmonte-Reche, E.; Martínez-García, M.; Guédin, A.; Zuffo, M.; Arévalo-Ruiz, M.; Doria, F.; Campos-Salinas, J.; Maynadier, M.; López-Rubio, J. J.; Freccero, M.; Mergny, J.-L.; Pérez-Victoria, J. M.; Morales, J. C. G-Quadruplex Identification in the Genome of Protozoan Parasites Points to Naphthalene Diimide Ligands as New Antiparasitic Agents. *J Med Chem* **2018**, *61* (3), 1231–1240. https://doi.org/10.1021/acs.jmedchem.7b01672.
54. Paeschke, K.; Bochman, M. L.; Garcia, P. D.; Cejka, P.; Friedman, K. L.; Kowalczykowski, S. C.; Zakian, V. A. Pif1 Family Helicases Suppress Genome Instability at G-Quadruplex Motifs. *Nature* **2013**, *497* (7450), 458–462. https://doi.org/10.1038/nature12149.
55. Dahan, D.; Tsirkas, I.; Dovrat, D.; Sparks, M. A.; Singh, S. P.; Galletto, R.; Aharoni, A. Pif1 Is Essential for Efficient Replisome Progression through Lagging Strand G-Quadruplex DNA Secondary Structures. *Nucleic Acids Res* **2018**, *46* (22), 11847–11857. https://doi.org/10.1093/nar/gky1065.
56. Sinha, K. M.; Hines, J. C.; Downey, N.; Ray, D. S. Mitochondrial DNA Ligase in Crithidia Fasciculata. *Proc Natl Acad Sci U S A* **2004**, *101* (13), 4361–4366. https://doi.org/10.1073/pnas.0305705101.
57. Lee, Y.-S.; Kennedy, W. D.; Yin, Y. W. Structural Insight into Processive Human Mitochondrial DNA Synthesis and Disease-Related Polymerase Mutations. *Cell* **2009**, *139* (2), 312–324. https://doi.org/10.1016/j.cell.2009.07.050.
58. Tabor, S.; Huber, H. E.; Richardson, C. C. Escherichia Coli Thioredoxin Confers Processivity on the DNA Polymerase Activity of the Gene 5 Protein of Bacteriophage T7. *J Biol Chem* **1987**, *262* (33), 16212–16223.
59. Brieba, L. G. Structure-Function Analysis Reveals the Singularity of Plant Mitochondrial DNA Replication Components: A Mosaic and Redundant System. *Plants (Basel)* **2019**, *8* (12), 533. https://doi.org/10.3390/plants8120533.
60. García-Medel, P. L.; Baruch-Torres, N.; Peralta-Castro, A.; Trasviña-Arenas, C. H.; Torres-Larios, A.; Brieba, L. G. Plant Organellar DNA Polymerases Repair Double-Stranded Breaks by Microhomology-Mediated End-Joining. *Nucleic Acids Res* **2019**, *47* (6), 3028–3044. https://doi.org/10.1093/nar/gkz039.
61. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.
62. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool,

K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res* **2022**, *50* (D1), D439–D444. https://doi.org/10.1093/nar/gkab1061.

63. Wheeler, R. J. A Resource for Improved Predictions of Trypanosoma and Leishmania Protein Three-Dimensional Structure. *PLoS One* **2021**, *16* (11), e0259871. https://doi.org/10.1371/journal.pone.0259871.
64. Klingbeil, M. M.; Motyka, S. A.; Englund, P. T. Multiple Mitochondrial DNA Polymerases in Trypanosoma Brucei. *Mol Cell* **2002**, *10* (1), 175–186. https://doi.org/10.1016/s1097-2765(02)00571-3.
65. Chandler, J.; Vandoros, A. V.; Mozeleski, B.; Klingbeil, M. M. Stem-Loop Silencing Reveals That a Third Mitochondrial DNA Polymerase, POLID, Is Required for Kinetoplast DNA Replication in Trypanosomes. *Eukaryot Cell* **2008**, *7* (12), 2141–2146. https://doi.org/10.1128/EC.00199-08.
66. Bruhn, D. F.; Mozeleski, B.; Falkin, L.; Klingbeil, M. M. Mitochondrial DNA Polymerase POLIB Is Essential for Minicircle DNA Replication in African Trypanosomes. *Mol Microbiol* **2010**, *75* (6), 1414–1425. https://doi.org/10.1111/j.1365-2958.2010.07061.x.
67. Bruhn, D. F.; Sammartino, M. P.; Klingbeil, M. M. Three Mitochondrial DNA Polymerases Are Essential for Kinetoplast DNA Replication and Survival of Bloodstream Form Trypanosoma Brucei. *Eukaryot Cell* **2011**, *10* (6), 734–743. https://doi.org/10.1128/EC.05008-11.
68. Concepción-Acevedo, J.; Luo, J.; Klingbeil, M. M. Dynamic Localization of Trypanosoma Brucei Mitochondrial DNA Polymerase ID. *Eukaryot Cell* **2012**, *11* (7), 844–855. https://doi.org/10.1128/EC.05291-11.
69. Baudouin, H. C. M.; Pfeiffer, L.; Ochsenreiter, T. A Comparison of Three Approaches for the Discovery of Novel Tripartite Attachment Complex Proteins in Trypanosoma Brucei. *PLoS Negl Trop Dis* **2020**, *14* (9), e0008568. https://doi.org/10.1371/journal.pntd.0008568.
70. Concepción-Acevedo, J.; Miller, J. C.; Boucher, M. J.; Klingbeil, M. M. Cell Cycle Localization Dynamics of Mitochondrial DNA Polymerase IC in African Trypanosomes. *Mol Biol Cell* **2018**, *29* (21), 2540–2552. https://doi.org/10.1091/mbc.E18-02-0127.
71. Miller, J. C.; Delzell, S. B.; Concepción-Acevedo, J.; Boucher, M. J.; Klingbeil, M. M. A DNA Polymerization-Independent Role for Mitochondrial DNA Polymerase I-like Protein C in African Trypanosomes. *J Cell Sci* **2020**, *133* (9), jcs233072. https://doi.org/10.1242/jcs.233072.
72. Crozier, T. W. M.; Tinti, M.; Larance, M.; Lamond, A. I.; Ferguson, M. A. J. Prediction of Protein Complexes in Trypanosoma Brucei by Protein Correlation Profiling Mass Spectrometry and Machine Learning. *Mol Cell Proteomics* **2017**, *16* (12), 2254–2267. https://doi.org/10.1074/mcp.O117.068122.
73. Fisk, J. C.; Li, J.; Wang, H.; Aletta, J. M.; Qu, J.; Read, L. K. Proteomic Analysis Reveals Diverse Classes of Arginine Methylproteins in Mitochondria of

Trypanosomes. *Mol Cell Proteomics* **2013**, *12* (2), 302–311. https://doi.org/10.1074/mcp.M112.022533.

Chapter 2: Trypanosoma brucei Mitochondrial DNA Polymerase POLIB Contains a Novel Polymerase Domain Insertion That Confers Dominant Exonuclease Activity

2.1 Abstract

Trypanosoma brucei and related parasites contain an unusual catenated mitochondrial genome known as kinetoplast DNA (kDNA) composed of maxicircles and minicircles. The kDNA structure and replication mechanism are divergent and essential for parasite survival. POLIB is one of three Family A DNA polymerases independently essential to maintain the kDNA network. However, the division of labor among the paralogs, particularly which might be a replicative, proofreading enzyme remains enigmatic. *De novo* modelling of POLIB suggested a structure that is divergent from all other Family A polymerases in which the thumb subdomain contains a 369 amino acid insertion with homology to DEDDh DnaQ family 3'-5' exonucleases. Here we demonstrate recombinant POLIB 3'-5' exonuclease prefers DNA vs. RNA substrates and degrades single- and double-stranded DNA non-processively. Exonuclease activity prevails over polymerase activity on DNA substrates at pH 8.0, while DNA primer extension is favored at pH 6.0. Mutations that ablate POLIB polymerase activity slow the exonuclease rate suggesting crosstalk between the domains. We show that POLIB extends an RNA primer more efficiently than a DNA primer in the presence of dNTPs but does not incorporate rNTPs efficiently using either primer. Immunoprecipitation of Pol I-like paralogs from *T. brucei* corroborate the pH selectivity and RNA primer preferences of POLIB and revealed that the other paralogs efficiently extend a DNA

primer. The enzymatic properties of POLIB suggest this paralog is not a replicative kDNA polymerase and the noncanonical polymerase domain provides another example of exquisite diversity among DNA polymerases for specialized function.

2.2 Introduction

Family A polymerases (pols) have a wide variety of functions in DNA replication and repair across all domains of life. Some of these pols, including *Homo sapiens* θand *Escherichia coli* Pol I are critical to DNA repair pathways. In eukaryotes, organellar DNA replication is also performed by Family A DNA pols. Examples include *H. sapiens* Pol γ and *Saccharomyces cerevisiae* Mip1 which are replicative DNA pols for mitochondrial DNA. In *Arabidopsis thaliana*, two Family A pols (At PolIA and AtPolIB) replicate both the chloroplast and mitochondrial DNA. Replication in these organelles likely uses a scheme where one of these proteins, PolIA, is the replicative enzyme while PolIB plays a supporting repair role [1]. Even in divergent systems of organellar DNA replication, such as within the quadruple-membrane bound organelle, the apicoplast, found in the malaria-causing protist *Plasmodium falciparum*, a Family A polymerase is the replicative enzyme (PREX) [2].

Another example of a highly divergent organellar DNA replication system is found in the group of protists named trypanosomatids. This group includes several species of parasites that are notorious human pathogens, including *Leishmania donovani* which causes Leishmaniasis, *Trypanosoma cruzi* which causes Chagas disease, and *Trypanosoma brucei* which causes African sleeping sickness in humans and Nagana in cattle. Trypanosomatids are characterized by their mitochondrial DNA structure, kinetoplast DNA (kDNA). The

kinetoplast is composed of catenated circular kDNA molecules in two categories: minicircles and maxicircles. In the well-studied trypanosomatid organism, *Trypanosoma brucei,* minicircles are 1 kb each and maxicircles are 23 kb [3]. These circular DNA molecules are condensed into a disc-shaped network found in close proximity to the flagellar basal body of these organisms.

In *T. brucei,* six DNA polymerases localize exclusively to the mitochondrion. These polymerases include two paralogs belonging to Family X, Pol β and Pol β-PAK and four Family A Pol I-like paralogs, POLIA, POLIB, POLIC, and POLID [4,5]. This multiplicity of DNA pols is likely the result of a gene duplication event since the four Pol I-like paralogs are conserved in all kinetoplastid organisms sequenced to date [6,7]. RNAi studies have shown that three of the Pol I-like proteins, POLIB, POLIC, and POLID, are essential for successful replication of the complex kDNA nucleoid found in the mitochondrion [5,8,9].

RNAi of POLIB or POLID resulted in loss of fitness, disrupted the balance of minicircle replication intermediates, and caused progressive loss of the kDNA network; hallmarks of a kDNA replication defect. Yet, a role in maxicircle replication was never addressed. Structure function studies of POLIC using RNAi complementation revealed that POLIC is a dual-functioning DNA Pol with essential roles in nucleotidyl incorporation and a non-catalytic role in kDNA distribution [10]. However, the precise division of labor among the paralogs, particularly which protein(s) is a processive, proofreading enzyme at a replication fork remains enigmatic.

Family A DNA polymerases have a conserved structure resembling a right hand with subdomains labeled, fingers, palm, and thumb. This structure of Family A DNA

polymerases are often compared to the founding member of the family, *E. coli* Pol I. Variations found in other Family A DNA pols include insertions that modulate their functions. An insertion in T7 DNA pol allows for binding to *E. coli* thioredoxin and enhanced processivity of the enzyme [11]. In plant organellar polymerases (POPs), small insertions allow for the enzymes to perform translesion synthesis [12]. Human mitochondrial DNA polymerase (Pol γ) has a 310 amino acid spacer region in its thumb subdomain that improves the intrinsic processivity of the enzyme and allows for it to bind to the non-catalytic subunit [13].

Based on homology modeling and *de novo* predicted structures, POLIB presents an unusually large insertion in the conserved Family A pol domain of POLIB. This insertion houses the catalytic residues typical of an exonuclease (exo) domain. Using recombinant purified protein, we confirmed this domain is indeed an active 3'-5' exonuclease that is capable of degrading both DNA and RNA oligos, with a preference for single stranded substrates. We also demonstrate nucleotidyl incorporation activity of POLIB for the first time and demonstrate a lower pH tolerance for this activity than for POLIB exo activity. POLIB also incorporates nucleotides more rapidly from an RNA primer than a DNA primer, suggesting an RNA primed template may be its more natural substrate. Our data suggest that POLIB is likely not contributing to kDNA replication as the replicative polymerase but is contributing to nucleic acid metabolism in the mitochondrion through exonuclease activity and short extension from an RNA primer.

Materials and Methods

Materials

All chemicals were molecular biology grade or better. Ultrapure NTPs were from New England Biolabs and dNTPs were from Invitrogen. Primers containing ribonucleotides, fluorescein (6-FAM), or a 2-aminopurine (2AP) were synthesized by Integrated DNA Technologies. 5' hexachlorofluorescein (HEX)-labeled RNA primer was gel-purified prior to use. All other primers were synthesized by Invitrogen. A full listing of primers used in this study can be found in Table 2.1, Table AI.1 and Table AI.2.

DNA substrate preparation

Two separate reactions were prepared in annealing buffer (10 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA, pH 7.8) One reaction contained unlabeled primer (22mer, RNA or DNA) and template (42mer) in a 1:1.25 ratio, while the other contained 5' HEX-labeled 22mer (RNA or DNA) and unlabeled 42mer in a 1:1.25 ratio. Both reactions were heated to 95°C for 5 min and cooled to room temperature before being combined in a ratio of 1:20 labeled:unlabeled substrate (8.4 μM).

	Length	Primer Sequence	Assay
Primer (5' HEX) **Template**	22 42	5' GCTACCGTGGTTGAGTCAGCGT 3' CGATGGCACCAACTCAGTCGCAGAATCAGTTGAGATGAGTAC	**Recessed DNA**
Primer (5' HEX) **Template**	22 42	5' GCTACCGTGGTTGAGTCAGCG**T** 3' CGATGGCACCAACTCAGTCGC**G**GAATCAGTTGAGATGAGTAC	**Mismatch**
Primer (5' HEX) **Template**	22 22	5' GCTACCGTGGTTGAGTCAGCGT 3' CGATGGCACCAACTCAGTCGCA	**Blunt DNA**
Primer (5' HEX)	22	5' GCTACCGTGGTTGAGTCAGCGT	**ssDNA**
Primer (5' HEX) **Template**	22 42	5' GCUACCGUGGUUGAGUCAGCGU 3' CGATGGCACCAACTCAGTCGCAGAATCAGTTGAGATGAGTAC	**Recessed Hybrid**
Primer (5' HEX) **Template**	22 22	5' GCUACCGUGGUUGAGUCAGCGU 3' CGATGGCACCAACTCAGTCGCA	**Blunt Hybrid**
Primer (5' HEX)	22	5' GCUACCGUGGUUGAGUCAGCGU	**ssRNA**
Primer (2AP position 1)	23	5' GCAGCGGATCTTAATGGATGGC**P**	**Processivity**
Primer (2AP position 10)	23	3' GCAGCGGATCTTA**P**TGGATGGCA	
Template (overhang)	44	5' CGTCGCCTAGAATTACCTACCGTCCACTGGGTAGATATTAGT	
Template (blunt)	23	3' CGTCGCCTAGAATTACCTACCGT	
Primer (3' FAM) **Template**	22 42	5' GTCTTAGTCAACTCTACTCATG 3' CGATGGCACCAACTCAGTCGCAGAATCAGTTGAGATGAGTAC	**Recessed DNA**
Primer (3' FAM)	22	5' GTCTTAGTCAACTCTACTCATG	**ssDNA**

Table 2.1. Annealed substrates for exonuclease and primer extension assays. Primer sequences displayed 5'-3' and template sequences are displayed 3'-5'. Primers contain a 5' hexachlorofluorescein (HEX) label, a 5' 2-aminopurine (2AP) or a 3' fluorescein (6-FAM) label. Black lettering indicated DNA sequence, red letters indicate RNA sequence, **bold** indicates mismatched nucleotides, and **blue P** indicates 2AP at either the 1st or 10th position.

Sequence Alignments and de novo Modelling

Representative Kinetoplastid POLIB protein sequences were retrieved from tritrypdb.org (release 56, 15 Feb 2022) that is hosted through the Eukaryotic Pathogen, Vector and Host Informatics Resource (veupathdb.org) and representative family A DNA polymerase domain sequences were aligned using NCBI COBALT multiple alignment tool [14,15,16]. Alignments were viewed in Snapgene (version 6.0.2). Structure based alignment of

TbPOLIB with *E. coli* Klenow (PDBid = kln1) defined residues 414-1400 as the POLA domain and were subsequently used for the design of truncated TbPOLIB for recombinant protein expression. DEDD family exonuclease motif sequences were aligned using the NCBI COBALT. De novo modeling of TbPOLIB (414-1400) was accomplished using the standard RoseTTAfold submission accessed via the RosettaCommons web-interface Robetta server on March 21st, 2022 (https://robetta.bakerlab.org/) [17,18].

Plasmid Preparation for Bacterial Expression

TbPOLIB (UniProt Accession Q8MWB4) coding region (residues 414-1400) was codon optimized for expression in *E. coli*, synthesized and cloned into the pET100/D-TOPO vector (Geneart, Thermofisher) (Table AI.1A), resulting in the addition of an N-terminal 6x His-tag. This truncated version of TbPOLIB is termed IBWT for this study. Quikchange Lightning Multi Site-directed mutagenesis kit (Agilent Technologies) was used to introduce catalytic site point mutations (Table AI.1B) to IBWT resulting in IBPol- and IBExo- variants. Mutations were verified by sequencing, and all truncated POLIB variants were transformed into BL21(DE3) chemically competent cells (Invitrogen) for protein expression.

Protein Purification

Purification scheme was identical for the three POLIB variants. Four liters of cells were grown in Luria Broth with shaking at 37˚C until OD600 reached 0.4-0.8. Protein expression was induced with 40 μM final concentration IPTG and incubated 16-18 hours at room temperature (~22°C) with shaking. Cells were harvested by centrifugation, washed with

phosphate buffered saline (PBS, 137 mM NaCl, 2.683 mM KCl, 10.144 Na_2HPO_4, 1.764 mM KH_2PO_4, pH 7.4) and stored at -20˚C. Frozen pellets (8.7-11.6 g) were resuspended in Lysis Buffer (10 mM NaH_2PO_4, 10 mM Na_2HPO_4, 500 mM NaCl, 10 mM imidazole, pH 8.0) and lysed by passing 2-3 times through a microfluidizer (Microfluidics). After lysis, all steps were performed at 4˚C. Lysate was centrifuged at 17090 x g for 1 hour, then the cleared bacterial lysate was passed through a 0.22 µm PES filter (Genessee Scientific) and then applied to a 5 mL packed Ni-NTA agarose beads (Goldbio) equilibrated with Lysis buffer. The beads were washed with 100 mL lysis buffer followed by 50 mL of Wash Buffer (10 mM NaH_2PO_4, 10 mM Na_2HPO_4, 500 mM NaCl, 40 mM imidazole, pH 8.0). Protein was eluted using 20 mL Elution Buffer (10 mM NaH_2PO_4, 10 mM Na_2HPO_4, 20 mM NaCl, 500 mM imidazole, pH 8.0) diluted 10-fold in Q Buffer (20 mM Tris, 1 mM DTT, pH 8.0) and applied to a pre-equilibrated (Q Buffer) 5 mL HiTrap Q-column (General Electric) using an ÄKTA pure chromatography system. Protein was eluted with a 0-500 mM NaCl linear gradient in Q buffer (75 ml). 1.5 mL fractions were collected. Fractions containing POLIB were identified by SDS-PAGE followed by staining with Coomassie blue. Fractions with active enzyme were pooled, concentrated (5 µM - 35 µM), and buffer exchanged into storage buffer (20 mM Tris, pH 8.0, 100 mM NaCl, 1 mM DTT, 20% glycerol) using Millipore Amicon Ultra Centrifugal Filter Units (50 kDa cutoff).

Protein concentration was determined using A280 (Nanodrop, Thermofisher), the molecular weight and the extinction coefficients of the POLIB variants predicted by Expasy ProtParam (https://web.expasy.org/protparam/). Small aliquots (10-20 µL) of protein were flash frozen in liquid nitrogen and stored at -80°C. Replicates of all assays were performed with three different preparations of the relevant POLIB variant.

Gel-Based Exonuclease Activity Assays

The incubation mixture contained, in a final volume of 50 μL, 50 mM Tris-HCl (pH 7.0), 1 mM DTT, 0.1% BSA. As substrate, 840 nM of single-stranded (ssDNA) or 840 nM of double-stranded DNA (dsDNA) was used (Table 1). The amount of each POLIB variant (25 nM or 50 nM) was adjusted to obtain data points for initial reaction rates, as noted.

Reactions were initiated with the addition of 5 mM metal ion (Mg^{2+} or Mn^{2+}) and incubated at 37˚C for the indicated times. Time points of reactions were taken by removing 10 μL and quenching by adding an equal volume of quench buffer (0.1 M EDTA, 80% formamide, 0.1% Orange G) at designated time points. Quenched reactions were heated to 95˚C for 10 min. Reactions were analyzed by electrophoresis in 7.5 M urea/20% polyacrylamide gels, visualized on a Typhoon™ FLA 9500 imager (GE) and band intensity was determined using Image J [19]. Percent of 22mer remaining in exonuclease assays was calculated using Equation 1:

$$22mer\% = \frac{22mer}{22mer + degradation\ products} * 100 \qquad \textbf{Equation 1}$$

where *22mer* represents the total signal coming from the 22mer and *degradation products* represent the total signal coming from all the bands below the 22mer substrate.

Exonuclease rates were calculated using Equation 2:

$$rate = \frac{product\ (nM)}{time\ (\text{sec}) * [\text{enzyme (nM)}]} \qquad \textbf{Equation 2}$$

where *product* represents the concentration of all product bands in nM, *time* represents the time point at which the reaction was quenched and *enzyme* represents the concentration of enzyme in nM.

Gel-Based Primer Extension Assays

Standard exonuclease reaction mixtures (50 μL) were used with the addition of 500 μM nucleotides (dNTP or rNTP) and increased POLIB variant concentration (200 nM). Immunoprecipitation of epitope-tagged Pol I-like paralogs from *T. brucei* cell extracts were performed at previously described [10]. To assay immunoprecipitated protein for activity, 5 μL protein bound to IgG beads was added to each reaction, and the whole reaction was quenched after 60 minutes with the addition of 50 μL quench buffer. Quenched reactions were heated to 95˚C for 10 min. Reactions were analyzed by electrophoresis and ImageJ using the same protocols as exonuclease Urea-PAGE assays. Percent extension product was calculated using Equation 3:

$$\%extension = \frac{extension\ products}{22mer + extenstion\ products} * 100 \qquad \textbf{Equation 3}$$

where *extension products* is the total signal from all the bands above the 22mer and *22mer* represents the signal from the 22mer substrate.

2-Aminopurine Exonuclease Processivity Assays

Each reaction (25 μL) consisted of 5 μM DNA substrate (Table 1), 50 mM Tris, 1 mM DTT, 0.1% BSA at pH 7.0. Reactions were preincubated at 37°C for 5 minutes before initiating with the addition of IBWT and were analyzed in a black 384 well plate. Relative fluorescence was measured using a SpectraMax M2e Microplate reader (excitation wavelength 310 nm and emission wavelength 375 nm) every 10 seconds. Free 2AP was calculated using a standard curve of 2AP (Alfa Aesar) measured using the same conditions as the processivity reactions. Each reaction was run in triplicate. Dynafit was used for global fitting of the full progress curves to determine the ratio of the lower or upper boundaries of the rate constants for 1st position 2AP substrates and 10th position 2AP substrates, the two scripts used to fit the data can be found in Appendix I Figures 1&2.

Trypanosoma brucei DNA Constructs

POLIA was amplified from gDNA from TREU927 cells using primers POLIA F and POLIA R (Table AI.2) and cloned into pLEW79-MHTAP using the BamHI site. The protein encoding sequence of POLIB was amplified from *T. brucei* procyclic form Lister 427 gDNA using primers POLIB F and POLIB R adding XhoI and XbaI cut sites (Table S2) and ligated into pLew100-PTP^{Puro} digested with the same restriction enzymes, generating pLew100-FLPOLIB-PTP^{Puro} [19,10]. A POLIB gene fragment (1-1053 bp) was obtained from Genescript in pUC57 to replace the natural corresponding POLIB sequence in pLew100-FLPOLIB-PTP^{Puro}. The gene fragment included a recoded region (249-801 bp) to use in unrelated RNAi complementation experiments. The gene fragment was amplified with primers POLIB Rc F and POLIB Rc R (Table AI.2) for subsequent Gibson assembly

with pLew100-FLPOLIB-PTP^{Puro} digested with XhoI and BamHI. Resulting vector and gene fragment were combined via Gibson Assembly (Takara In-fusion Kit) generating pLew100-FLPOLIB$_{Rc}$-PTP^{Puro}. Polymerase and exonuclease catalytic site mutations were separately introduced to this construct using mutagenic primers POLIB M1, POLIB M2, and POLIB M3 (Table AI.2) and the QuikChange Lightning Multi-Site Mutagenesis Kit (Agilent) according to manufacturer's protocol.

The *POLID* open reading frame (4887 bp), excluding the stop codon, was PCR-amplified from *T. brucei* 927 genomic DNA using primers POLID F and POLID R (Table AI.2). The PCR product was cloned into pCR-Blunt-II-TOPO (Invitrogen) to generate pTOPO-FLID-WT. Site-directed mutagenesis of *POLID* was performed on pTOPO- FLID-WT using the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent) with primers POLID M1 and POLID M2 to generate pTOPO-FLID-Poldead. Due to the relatively large plasmid size of pTOPO-FLID-WT, non-mutagenic primers POLID M3 and POLID M4 (Table AI.2) which anneal to the pCR-Blunt-II-TOPO backbone, were also included in each mutagenesis reaction. TbPOLID variants wild type and polymerase-dead were cloned into pLew100-PTP^{Puro} via MfeI and BamHI sites by GENEWIZ, Inc. generating pLew100ID-WT-PTP^{Puro} and pLew100ID-poldead-PTP^{Puro}.

Trypanosoma brucei Culture and Transfection

Procyclic *Trypanosoma brucei brucei* 29-13 cells expressing T7 RNA polymerase and tetracycline repressor were maintained at 27°C in SDM-79 medium supplemented with heat-inactivated fetal bovine serum (15%), G418 (15 μg/ml), and hygromycin (50 μg/ml)

[21]. Transfected cell lines were additionally supplemented with the appropriate selectable drug and the resulting transgenic variants used in this study are listed in Table AI.2B.

Generation of overexpressing POLIC variant cell lines was previously described [10]. 10 μg pLEW79IA-MHTAP was NotI linearized and transfected into 29-13 procyclic cells by electroporation [8]. 7.5 μg pLew100-PTP^{Puro} plasmids encoding full length POLIB and POLID variants were NotI linearized and transfected into 29-13 cells via Amaxa Nucleofection (Parasite Kit, Lonza). Following selection with 1 μg/mL puromycin for pLew100 based constructs and 2.5 μg/mL phleomycin for cells transfected with pLEW79IA-MHTAP, clonal cell lines were obtained via limiting dilution as previously described [8].

Overexpression of all variants was induced by addition of 1 μg/mL tetracycline to the cultures and supplemented with 0.5 μg/mL on days dilutions were not performed. Clones were selected for use based on highest expression of inducible protein. Cultures were induced for 48 hours before cells were collected for immunoprecipitation. No impact on fitness was detected for any of the overexpression cell lines during the induction period.

Western Blot and SYPRO™ Ruby Protein Detection

Proteins were separated via SDS-PAGE on 8% polyacrylamide gels, followed by 16-hour transfer onto a PVDF membrane. Membranes were blocked in Tris-buffered saline (TBS) containing 5% non-fat dry milk. PTP-tagged or MHTAP-tagged protein variants were detected using the peroxidase-anti-peroxidase soluble complex (PAP) (1:2000, Sigma) as previously described [10]. Detection of 6x His-tagged recombinant POLIB variants was

performed with primary antibody Penta•His (1:1000, Qiagen) for one hour, and subsequently incubated with secondary rabbit anti-mouse IgG horseradish peroxidase conjugate (1:4000, Zymed) for one hour. Signal was detected using Pierce™ ECL Western Blotting Substrate with an Amersham ImageQuant 800 (Cytiva). Sypro™ Ruby (Thermofisher) detection of total protein was performed following SDS-PAGE according to manufacturer's protocol.

2.4 Results

Kinetoplastid organisms are the only group of eukaryotes that possess four Family A DNA polymerase paralogs exclusively targeted to the mitochondrion [5,7]. Among these paralogs, the Family A Pol domains (PolA) share only 25-35% amino acid identity, suggesting specialized roles for each [5]. In addition to the conserved C-terminal PolA domain, only the POLIB and POLID paralogs contain predicted exonuclease (exo) domains. Although the POLID exonuclease domain shares similarity with those of other Family A pols, the POLIB exo domain is markedly different in sequence and position.

POLIB presents a Noncanonical Polymerase domain containing a large insertion

Predictions of POLIB domains using the SuperFamily database show two stretches of amino acids that have homology to Family A polymerases (414-660 and 1031-1400 aa) (Figure 2.1A) [22]. Embedded within the PolA domain is a region with homology to the DnaQ 3'-5' exo superfamily (763-956 aa). The first 30 aa are predicted to be a mitochondrial targeting sequence (TargetP-2.0) but the remainder of the N-terminal region (through aa 413) does not have homology to any characterized proteins and is referred to

as the uncharacterized region (UCR) [23]. Attempts to gain structural insight into the unusual POLIB domain arrangement using homology modeling of amino acids 414-1400 with the *E. coli* Pol I structure revealed a canonical right-hand Pol domain structure (fingers, palm, thumb subdomains) with a large unstructured insertion in the thumb domain that spanned the C-terminal amino acids 631-1030. When modeled separately, the insertion had homology to *E. coli* RNaseT, a DnaQ exo superfamily member.

Recent developments in *de novo* protein structure prediction using deep learning now allow for predicted protein structures with atomic accuracy [17, 24]. We used the *de novo* modelling platform RoseTTAfold to gain insight into how the two POLIB annotated enzymatic domains would fold as part of the same structure. Figure 1B depicts a model with a local Distance Difference Test (lDDT) confidence score of 0.75, indicating a good quality model.[25] In the plot of predicted error of position of the residues in the model, the majority falls under 5 ångströms with the main exception being the C terminus (Figure AI.3A). The Pol domain right-hand structure contains the previously identified invariant aspartic acids (D1117 and D1309) essential for nucleotidyl incorporation in the palm subdomain as expected. An amino acid alignment of various PolA domains highlights the conserved Motifs A, B and C and a key difference for POLIB, which is the presence of HDS instead of the invariant HDE found in Motif C in all other Family A pols insertion (Figure AI.4).

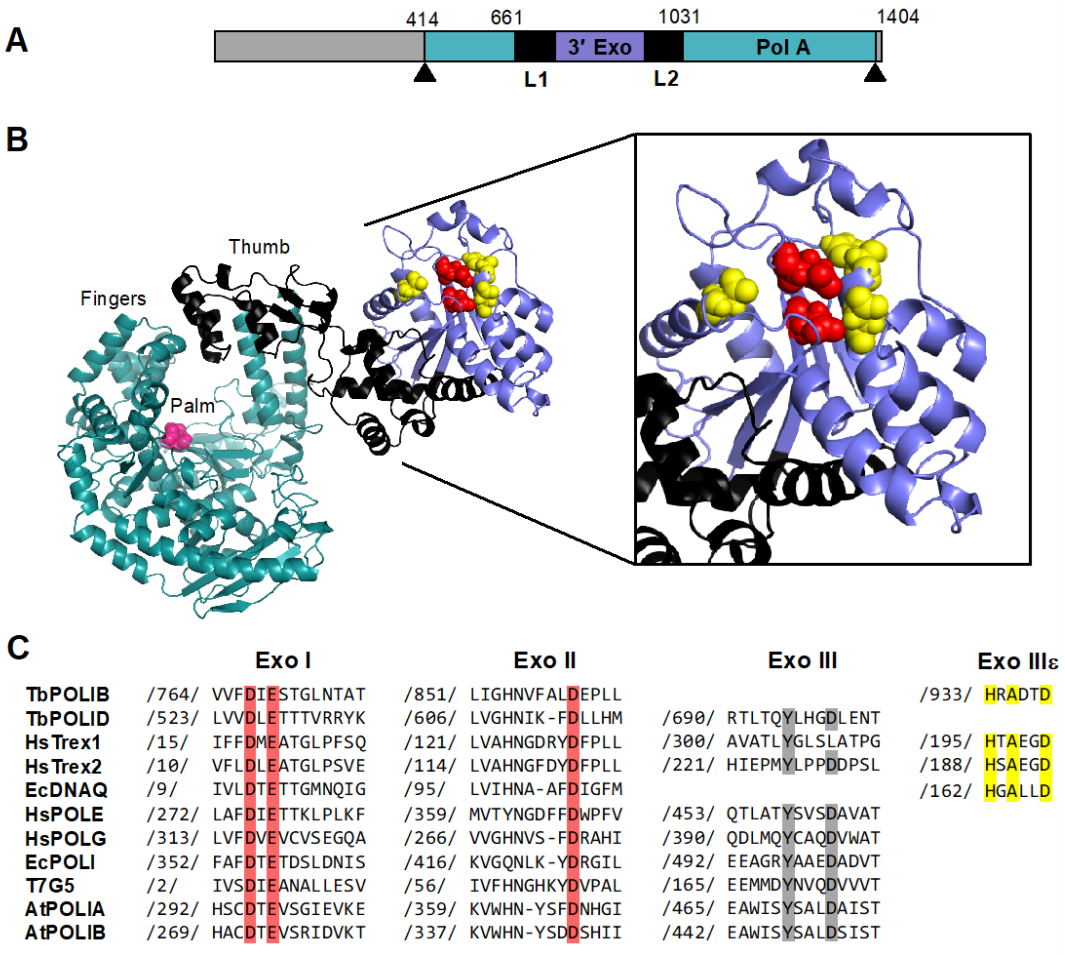


Figure 2.1: Divergent Structure of TbPOLIB. (A) Schematic representation of TbPOLIB with predicted polymerase (Pol) and exonuclease (Exo) domains. Mitochondrial targeting sequence, 1-30; N-terminal uncharacterized region, 31–413; Linker region 1 (L1), 661–761; Exo, 762–955; Linker region 2 (L2), 956–1030 and pol domain, 413–660 and 1031–1364. Arrowheads mark the region used for de novo modeling (414-1400). (B) RoseTTAfold model of TbPOLIB. Teal, pol domain shows as canonical ‘‘right-hand’’ configuration with fingers, palm and thumb, fingers subdomains; purple, thumb insertion with homology to exonuclease; black, linker regions. Magenta, Pol domain active sites (only one active site residue is seen in this orientation); Red, Exo domain Asp767 and Glu769; Yellow, additional residues associated with DEDDh exonucleases. (C) Alignment of conserved Exo motifs for selected DEDD 3’-5’ exonucleases. Red, three active site carboxylates; Yellow, DEDDh consensus motif (HxAxxD); Grey, DEDDy consensus motif (YxxxD).

The RoseTTAfold model depicts the large 369 aa insertion as a projection from the tip of the thumb. Residues 763-956 fold into a separate domain that is bounded by flexible linker regions (L1 and L2). Smaller thumb domain insertions have been described in other Family

A Pols that contribute to increased processivity (T7 Pol, human Pol θ) or translesion DNA synthesis (*A. thaliana* PolIA) [11,12,26]. Additionally, proofreading Family A members have the exo domain located under the palm subdomain (*E. coli* Pol I, Pol γ) [27,28]. Thus far, POLIB is divergent from all other Family A proteins containing a unique substitution within Motif C, and the largest thumb insertion with similarity to an exo domain. Interestingly, these divergent features are conserved among all the POLIB sequences from Kinetoplastid organisms sequenced to date (Figure AI.5). A similar conformation was predicted for the full-length POLIB model retrieved from the Alphafold database depicted with (Fig AI.3B) and without the UCR present in the model (Figure AI.3C). The position of the UCR relative to the other domains has low confidence based on the predicted aligned error so its position relative to the rest of the protein is unreliable in this model (Figure AI.3D).

The DnaQ exo superfamily contains three conserved motifs Exo I, Exo II and Exo III that are clustered around the active site and participate in a two-metal ion catalysis. This superfamily is also called the DEDD based on four negatively charged invariant residues present in the conserved motifs and can be further classified into DEDDy or DEDDh proteins depending on the variation found in motif III (YxxxD or HxAxxD, respectively) [29]. POLIB contains the essential active site residues characteristic of the DEDDh subfamily of exos (Figure 2.1B, C). DEDDh exos includes *E. coli* DnaQ, the proofreading subunit for *E. coli* Pol III, and human TREX1 and TREX2 which non-processively degrade cytosolic ssDNA and dsDNA [30]. The 3'-5' exo domain of POLID, a paralog of POLIB, is more similar to the DEDDy proofreading domains of human Pol γ and Pol ε.

POLIB Polymerase Domain Insertion is an Active 3'-5' Exonuclease

To evaluate whether the POLIB insertion was an active 3'-5' exo, we generated codon optimized His-tagged recombinant POLIB variants for expression and purification from *E. coli*. The variants were truncated to eliminate the UCR and include IBWT, IBExo- (D767N, E769Q), and IBPol-, (D1117A, D1309A). All variants were purified to near homogeneity with the expected molecular weight of 115.6 kDA using Ni affinity and anion exchange chromatography. SYPRO ruby staining demonstrated that the purity of the POLIB variants was higher than 96% (Figure AI.6). Typical yields were 0.575-1.625 mg of pure protein per liter of induced culture.

The 3'-5' exonuclease activity of POLIB variants was measured using a 5' hexachlorofluorescein (HEX)-labeled 22mer primer annealed to an unlabeled 42mer template in the absence of dNTPs. Degradation products were visualized on a Urea-PAGE gel. IBWT exhibits 3'-5' exo activity that is completely ablated in the IBExo- variant, while IBPol- displayed reduced degradation activity on a recessed dsDNA template (Figure 2.2A). IBWT degraded 15% of the primer to 21mer or smaller products within the first minute of the reaction, whereas IBPol- only degraded 5% during the same time. At longer time points, from 2 to 8 min, intermediate degradation products accumulated but products smaller than 10mer were only observed for IBWT.

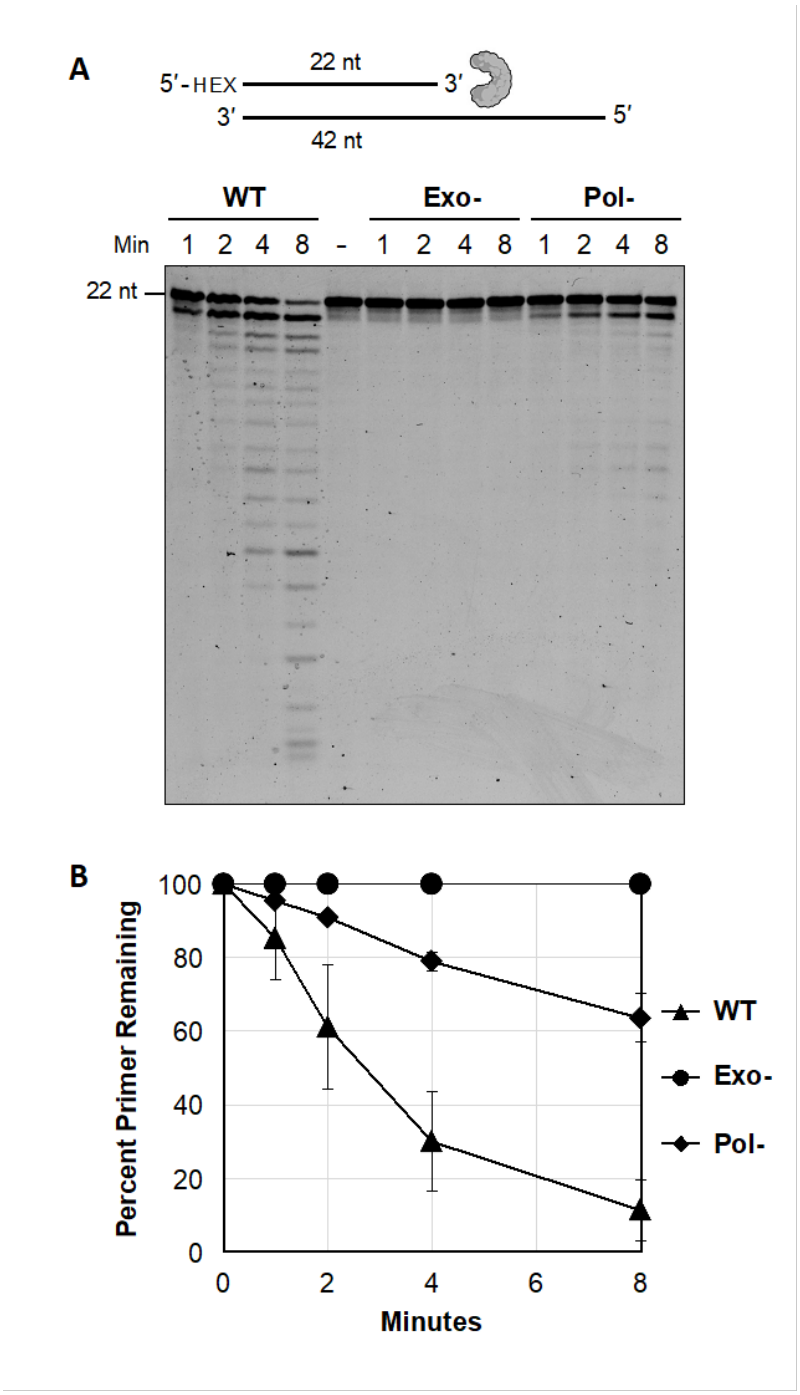


Figure 2.2: Exonuclease activity of POLIB variants. (A) Examination of 3'-5' exonuclease activity using 840 nM recessed 5' HEX-labeled primer/template dsDNA substrate, 50 nM POLIB of each variant in the absence of dNTPs. Reactions stopped at the indicated times to evaluate progression of reactions. Representative image from multiple experiments. No protein control Time=0, N. (B) Quantification of dsDNA degradation of POLIB variants. N=3, error bars represent standard deviation.

Quantification of primer degradation for the POLIB variants demonstrated that catalytic mutations in IBPol- impacted the rate of DNA degradation about 2-fold, suggesting that there may be crosstalk between the Pol and Exo domains of POLIB, a characteristic described in other well-characterized DNA polymerases such as human Pol γ [28]. No 5'-3' exonuclease activity was detected when using 3' fluorescein (6-FAM)-labeled single stranded or double stranded DNA substrates compared to the T5 exo control (Figure AI.7).

Substrate Preference for POLIB Exonuclease Activity

Gel-based assays using the HEX-labeled DNA and RNA substrates were also used to determine substrate preference for IBWT exo activity (Figure AI.8). Some proofreading DNA polymerases contain an additional 3'-5' exonuclease activity that can be used to remove an incorrectly added nucleotide during DNA extension. Proofreading domains typically display a preference for removing mismatched nucleotides over correctly paired matches, including those domains in human Pol γ and *E. coli* Pol I [31,32]. We tested whether POLIB displays a preference for mismatch removal by observing the relative rate of removal of a mismatched terminal nucleotide versus a correctly terminal nucleotide on a recessed 3' DNA substrate (Figure 2.3A). No significant difference was observed in the relative rates of these reactions, suggesting POLIB does not have a preference for mismatched substrates.

We also assayed the exonuclease activity on different DNA templates to determine what may be the preferred substrates of POLIB. Due to the higher rate of degradation of POLIB on single-stranded oligos, IBWT concentration was reduced to 25 nM in reactions with these substrates. IBWT successfully degraded DNA from the 3' end and was capable of degrading both ssDNA and dsDNA with a 5' overhang or blunt dsDNA. IBWT degraded ssDNA at a faster rate than either dsDNA substrate, with an average rate of $0.9914 \pm 0.0560\ s^{-1}$ for ssDNA, compared to $0.0626 \pm 0.0136\ s^{-1}$degradation of the recessed DNA substrate and $0.09133 \pm 0.0055\ s^{-1}$ for blunt dsDNA (Figure 2.3B). On average, IBWT degraded blunt dsDNA more rapidly, and appeared to remove the terminal nucleotide of the blunt DNA substrate more rapidly than the terminal nucleotide of the recessed DNA substrate. However, POLIB degraded subsequent nucleotides more rapidly in the recessed DNA substrate than the blunt substrate, as seen by the increase of shorter degradation products. (Figure AI.8). This could be due to an affinity of IBWT to longer DNA templates. POLIB is also capable of degrading ssRNA and RNA/DNA hybrids, although at lower rates than DNA degradation (Figure 2.3C). No detectable degradation occurred in the first 30 seconds of the ssRNA degradation reaction, suggesting IBWT has a strong preference for degrading ssDNA. On average, the rate of degradation for the ssRNA oligo was $0.1591 \pm 0.0216\ s^{-1}$, as compared to a rate of 0.0369 ± 0.0103 for the recessed RNA and 0.0198 ± 0.0151 for the blunt RNA/DNA hybrid.

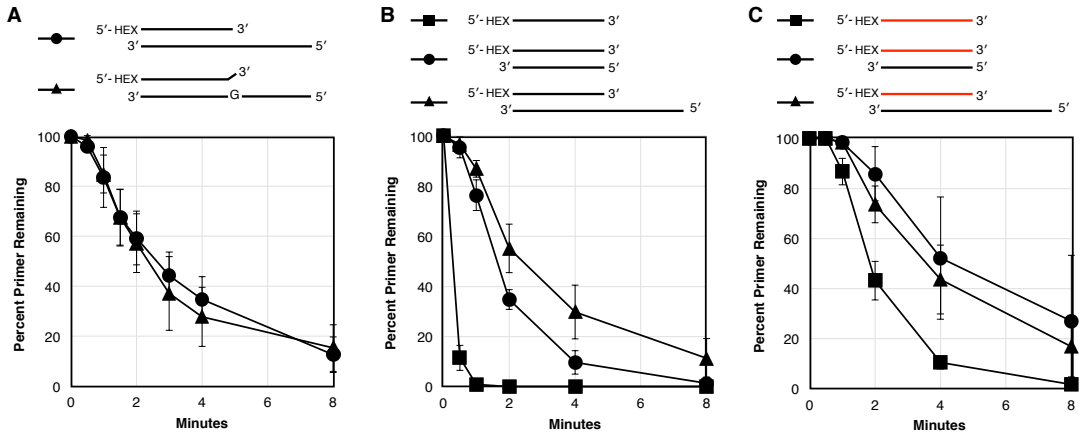


Figure 2.3: POLIB exonuclease template preference. Quantification of gel-based exonuclease assays. Standard reactions contained 25 nM IBWT for single stranded substrates or 50 nM for double stranded substrates in order to observe the linear range of these reactions. Reaction time points were run on Urea-PAGE gels and quantified using image J. Red oligo: RNA, Black oligo: DNA. (A) Relative rates of primer degradation with either 3' terminal matched or mismatched basepairs. (B) Relative rates of degradation of DNA oligos. ssDNA reaction used 25 nM POLIB and 840 nM DNA, dsDNA reactions used 50 nM POLIB and 840 nM DNA. (C) Relative rates of degradation of 840 nM ssRNA with 25 nM POLIB, or RNA in an RNA-DNA hybrid with 50 nM POLIB.

POLIB degrades DNA substrates in a Non-processive Manner

DNA substrates containing 2-aminopurine (2AP) at either the 3′ end (1st position) or at the 10th internal nucleotide (10th position) were used to evaluate the processivity of IBWT on three different substrates: ssDNA, recessed dsDNA, or blunt dsDNA. Single-hit conditions were used, where the concentration of the DNA substrate was much higher than IBWT (25x higher for dsDNA substrates, 200x higher for ssDNA substrates) [33]. For all the three substrates, a significant delay between the removal of the 1st position 2AP and the 10th position 2AP was observed. The lag period that is observed for the 10th position DNA

substrate indicates that multiple binding and nucleotide excision events must occur prior to the enzyme reaching the 10th position (Figure 2.4).

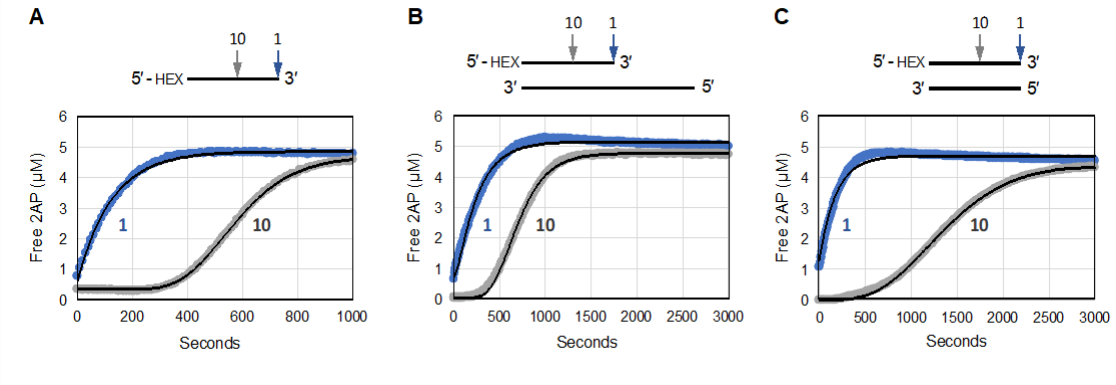


Figure 2.4: Exonuclease Processivity of POLIB. Standard reactions contained 5 μM of DNA substrate with 2AP either at the 1st position (1AP) or the 10th position (10AP). Progress curves were fitted using Dynafit to determine upper or lower boundaries of k_{exo} and k_{off}. Blue circles, 1AP data points; grey circles, 10AP data points; black line, fitted progress curve. (A) Representative progression curves for ssDNA substrate degradation by 50 nM IBWT. (B) Representative progression of reactions for recessed dsDNA substrate degradation by 200 nM IBWT. (C) Representative progression curves for blunt dsDNA substrate degradation by 200 nM IBWT.

The lower or upper limit of the k_{exo} and k_{off} values were determined by global fitting of the progress curves for the 1st and 10th 2AP positions with Dynafit (Figures AI.1, AI.2). This ratio of these two rate constants can be used to determine the likelihood of whether POLIB will dissociate from a DNA substrate or remain bound to excise another nucleotide. This ratio was 95.92 ± 8.9 for ssDNA, indicating very low processivity. For the recessed dsDNA substrate, the ratio of the upper limit of k_{exo} and k_{off} is 28.949 ± 30.988, also demonstrating low processivity of IBWT (Figure 2.4B). For blunt dsDNA, this ratio is 2.429 ± 0.462, which also indicates POLIB is more likely to dissociate from this substrate

after removing a single nucleotide than removing another in a single binding event (Figure 2.4C).

For ssDNA substrates, the excision of the 1st position 2AP occurred at a slower rate than subsequent nucleotides, which could indicate a lower affinity for the 2AP nucleotide or the presence of a slow activation step prior to the onset of activity (i.e., enzyme hysteresis) (Figure 2.4A). This is also the case for blunt dsDNA but was not observed in the case of the recessed dsDNA substrate. This possibly indicates a reduced affinity of IBWT to the shorter products of degradation from the blunt 22mer in comparison to the longer 42mer template of the recessed substrate. These substrate-dependent changes in rate as the enzyme proceeds further into the substrate were also observed on gel-based exo assays, as noted (Figure 2.3).

pH Impacts Competition Between DNA Extension and Degradation Activities of POLIB

Two other pols involved in kDNA maintenance, Pol β and Pol β-PAK, incorporate nucleotides most efficiently at pH 9 [4]. We therefore predicted POLIB would catalyze more efficiently at a basic pH. Notably, we observed a change in IBWT preference between degradation and extension activities that was dependent upon pH during characterization of recombinant POLIB (Figure 2.5A).

IBWT was used to demonstrate the relative rate of extension and degradation of a 5′ Hex-labeled recessed DNA template at pH conditions of 6.0, 7.0, and 8.0. At pH 8.0, minimal nucleotidyl incorporation was observed but shorter exonuclease products build up through

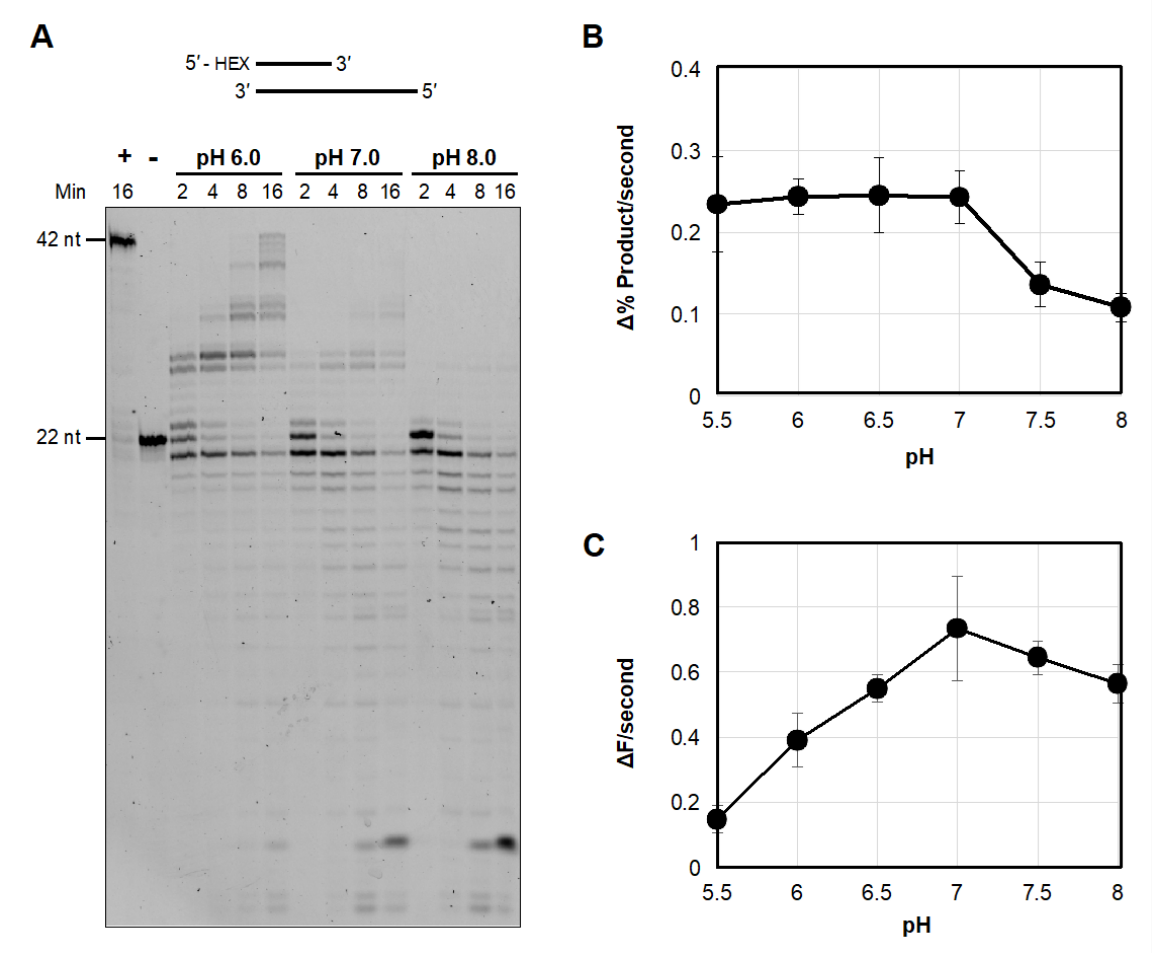


Figure 2.5: Impact of pH on POLIB exonuclease and extension activities. A) Representative image of IBWT extension and degradation products assayed using standard reaction buffer with varying pH conditions and 500 μM dNTPs. 1U Klenow, +; Time=0, pH 7.0, N. (B) Rates of extension determined using a gel-based assay. Standard reaction buffer was used with varying pH conditions, 200 nM IBExo- 840 nM oligo substrate, and 500 μM dNTPs. Rate determined using quantification of original primer and extension products. N=3, error bars represent standard deviation. (C) Rates of 2AP excision from recessed dsDNA substrate varying pH conditions, 200 nM IBWT and 5 μM substrate in each reaction.

the course of the reaction. At pH 7.0, both extension products and degradation products are being formed through the course of the reaction. At pH 6.0, more extension products are formed than degradation products (Figure 2.5A). We hypothesized that this pattern was due to separate pH optimums for the nucleotidyl incorporation and exonuclease activities

of POLIB. To test this, we used IBExo- (lacking competing exonuclease activity) to observe the rate of extension in varying pH conditions using the gel-based assay. POLIB extension activity was relatively similar from pH 5.5 – 7.0 but decreased in pH 7.5 and 8.0 by about half (Figure 2.5B).

To test the rate of nucleotide excision across pH conditions, we used a primer with a terminal 3′ 2-aminopurine (2AP), an adenine analog that is quenched by base-stacking and fluoresces when released from the primer [33]. Using a recessed dsDNA substrate, POLIB was able to degrade DNA most rapidly at a pH of 7.0. The rate decreased under more acidic pH conditions, demonstrating a 3-fold lower rate of excision at pH 5.5 compared to pH 7.0. However, at pH 8.0 exonuclease activity only declines 23% from the rate at pH 7.0 (Figure 2.5C). Similar patterns were seen for the change in efficiency of exonuclease activity across pH conditions for a blunt dsDNA oligo and ssDNA (Figure AI.9). These results demonstrate that the pH tolerance for the nucleotidyl incorporation and degradation activities of POLIB are distinct and impact the activity preference of the enzyme.

We also investigated other factors that could have an impact of the rate of primer extension by IBExo-. Increasing the salt concentration of either NaCl or KCl, from 0 mM to 50 mM reduced the extension activity of IBExo- (Figure AI.10). A concentration of 150 mM of either salt eliminated any detectable extension. Another variable we tested was an exchange from using Mg^{2+} as the divalent cation to Mn^{2+}. These reactions demonstrated that IBExo- incorporates dCTP at a similar rate with either divalent cation but was more likely to misincorporate dCTP past the correct basepair on the template with Mn^{2+}. Furthermore,

with Mg^{2+} as a divalent cation IBExo- did not extend the DNA primer with CTP, but was able to incorporate CTP in the presence of Mn^{2+} (Figure AI.11 A, B).

POLIB Exhibits Preference for Extension from RNA primers

The invariant Asp in Motif C that is conserved in other Family A pols is replaced with a Ser in the Kinetoplastid lineage (Figure AI.3). A serine is present in this motif in some RNA pols such as T3 bacteriophage RNA Pol and *Saccharomyces cerevisiae* mitochondrial RNA Pol [34,35]. Based on the Ser substitution, we hypothesized that the pol domain of POLIB may have a role in RNA metabolism. We evaluated the three POLIB variants in 60-minute extension reactions for extension activity using either an RNA or DNA primer is the presence of deoxyribonucleotides or ribonucleotides (Figure 2.6A).

IBWT was capable of extending from an RNA primer utilizing dNTPs, and this extension activity was improved in IBExo- that lacks formation of degradation products and results in an increase in fully extended product (Figure 2.6A). When the reaction using IBWT included rNTPs instead of dNTPs, exonuclease activity outcompeted extension. IBExo- was capable of extending an RNA primer with rNTPS, although the majority of the products formed in this case are the result of a single nucleotide addition. Any extension of the RNA primer is ablated in reactions with IBPol-, although IBPol- is capable of degradation of both RNA and DNA primers.

With a template primed with DNA and the addition of dNTPs, IBWT formed more degradation products than with an RNA primer and fewer extension products. IBExo- produced more extension products than IBWT in these conditions as well. No IB variants

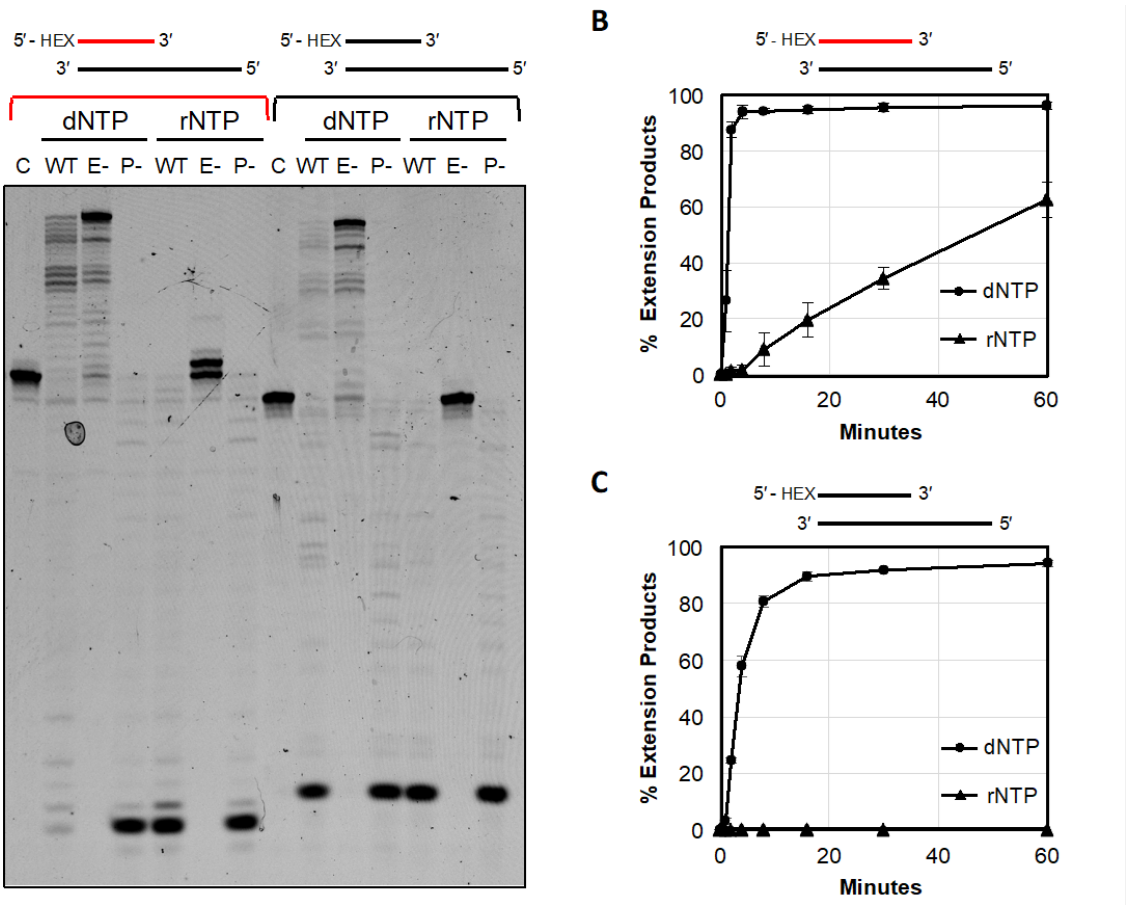


Figure 2.6: POLIB extension from RNA and DNA primers. Gel-based assays for primer extension, each reaction included 840 nM oligo, 200 nM POLIB variant, 500 µM dNTPs or rNTPs in standard assay buffer. Red oligo: RNA, Black oligo: DNA. (A) Representative gel-based assay for POLIB variant extension of DNA or RNA primers. and was quenched at 60 minutes. No protein control Time=0, C. (B) Quantification of gel-based assay time course of IBExo- extension from RNA primer, averages from 3 gels (C) Quantification of gel-based assay time course of IBExo- extension from DNA primer, averages from 3 gels, error bars represent standard deviation.

were able to extend the DNA primer with rNTPs. In order to determine if primer composition impacts the rate of the POLIB nucleotidyl incorporation, we quantified the progress of extension reactions using Urea-PAGE over a 60-minute time course. IBExo- extended an RNA primer with dNTPs more efficiently than with rNTPs. (Figure

2.6B). Progress curves of IBExo- extension of a DNA primer with dNTPs were slower than extension from an RNA primer, with 3% of the DNA primer extended in the first minute of the reaction compared to 26% of the RNA primer (Figure 2.6B,C). IBExo- was not able to incorporate rNTPs from a DNA primer at all (Figure 2.6C). This primer-dependent activity may suggest a conformational change of the IB pol domain active site when bound to an RNA primer.

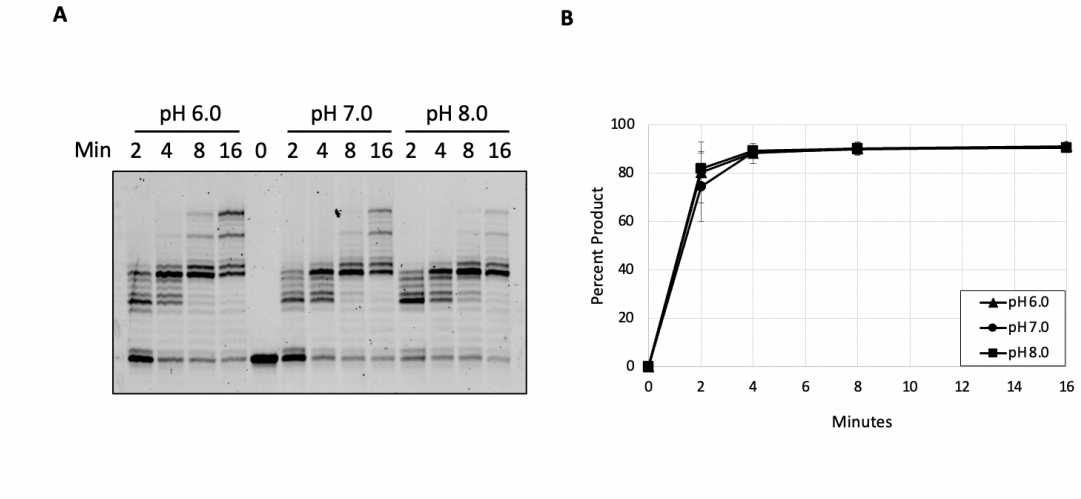


Figure 2.7: pH Impact on Extension Activity of RNA primer. A) Representative gel from two replicates of RNA extension assays in varying pH conditions. Reactions contained 200 nM IBexo- along with Hex-labeled RNA primer annealed to a DNA template prepared as described for assays in chapter 2. Reactions were identical to previous extension reactions, briefly: 840 nM RNA primed DNA substrate, 200 nM IBexo-, 0.1% BSA, 1 mM DTT, 500 mM dNTPs, with variation of pH of the 50 mM Tris buffer to 6.0, 7.0, or 8.0. B) Quantfication of percent product formed in from two experiments. Error bars= standard deviation. N=2

Extension from an RNA primer is Stable across pH Conditions

IBexo- extension from an RNA primer using dNTPs represents the most robust extension activity seen by POLIB. Unlike with a DNA primer, changing the pH conditions with an RNA primer to a pH of 6.0, 7.0, or 8.0 do not seem to have a significant impact on the rate of extension by POLIB (Figure 2.7A, B). However, the extension of the RNA primer may be out of the linear range of the reaction before the first time point is taken at two minutes. The assay used to determine changes in extension rate of IBexo- from an RNA primer in varying pH conditions should be repeated with earlier time points included, to make a more robust conclusion.

Activity of Immunoprecipitated Mitochondrial POLI-like proteins

To evaluate whether truncation of the POLIB UCR impacts POLIB activities we immunoprecipitated POLIB variants and the other mitochondrial Pol I-like paralogs from *T. brucei* cell lines overexpressing full-length C-terminally PTP epitope-tagged proteins. Successful immunoprecipitation of variants was verified by SDS-PAGE and Western blot (Figure AI.12A). Equal volumes of PTP-tagged proteins still bound to IgG sepharose beads were assayed using standard buffer conditions, the HEX-labeled recessed DNA substrate. Initially pH 8.0 was used to mimic the predicted basic environment of the mitochondrial matrix [36].

In addition to the Klenow control, POLIA, POLIC, and POLID wildtype proteins exhibited nucleotidyl incorporation while precipitates from untagged parental 29-13 *T. brucei* and beads only lacked any detectable activity. POLIC and POLID variants with alanine

substitutions in the Pol domain conserved aspartic acid catalytic residues (ICpol- and IDpol-) ablate all detectable nucleotidyl incorporation activity (Figure 2.7A). The results for POLIC and POLICpol- agree with previously reported data [10]. POLIBwt, POLIBpol-, and POLIDwt exhibited 3′-5′ exonuclease activity. POLIDpol- demonstrated significantly more degradation products than POLIDwt. Notably, POLIBwt exhibited only exonuclease activity even in the presence of dNTPs that is lost in the POLIBexo- variant. However, weak nucleotidyl incorporation was detected (Figure 2.8A). This suggests that the full-length POLIBwt exonuclease activity outcompeted nucleotidyl incorporation activity at pH 8.0.

We next asked whether pH and primer composition might also impact extension activity of immunoprecipitated POLIB variants similar to that of truncated recombinant protein. Successful immunoprecipitations were verified by SDS-PAGE and Western blot (Figure AI.12B). POLIBwt and POLIBexo- immunoprecipitate was assayed under standard primer extensions conditions, with the pH varied as noted. Activity on RNA or DNA primers was compared using either a HEX-labeled recessed DNA substrate or a HEX-labeled recessed RNA/DNA hybrid. When immunoprecipitated POLIBwt was assayed at pH 8.0, more degradation products were detected than at lower pH conditions

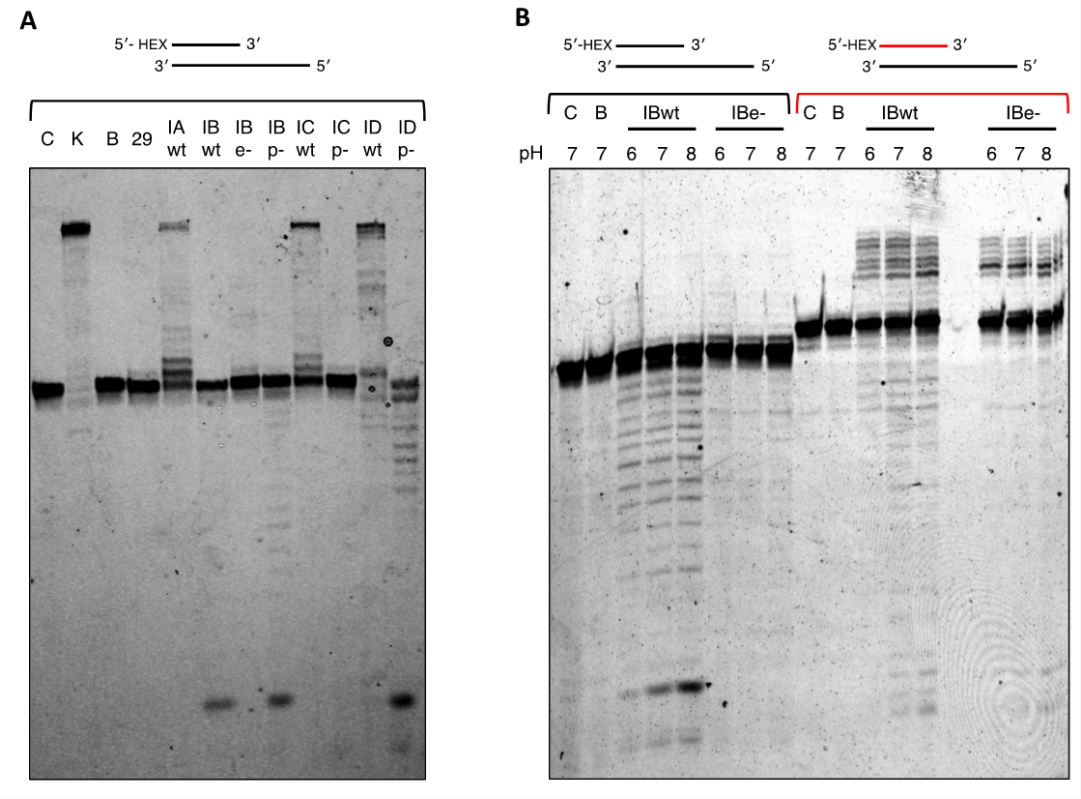


Figure 2.8: Pol and Exo activity of immunoprecipitated *T. brucei* Pol I-like paralogs. Gel-based extension assays were used to demonstrate activity of immunoprecipitated protein. (A) Each reaction contained 840 nM DNA primed substrate, 5 µL IgG beads bound to immunoprecipitated protein, using standard reaction buffer at pH 8.0 with 500 µM dNTPs for 60 minutes. No protein control Time=0, C; 1U Klenow fragment positive control, K; IgG beads control, B; 29-13 parental cells with IgG beads, 29. (B) POLIB variants in pol reactions at pH 6.0, 7.0, or 8.0, including 500 µM dNTPs with 840 nM either DNA primed template or RNA primed template for 60 minutes. Red oligo: RNA, Black oligo: DNA.

for both the DNA and RNA substrates similar to the results for truncated IBWT (Figure 2.8B). Degradation activity was favored over primer extension on a DNA primed substrate, but extension from an RNA primer was favored for both the POLIBwt and POLIBexo- variants. Some degradation products are detected in reactions with POLIBexo- possibly

due to incomplete ablation of exonuclease activity in the full-length protein or a contaminating exonuclease in the immunoprecipitated protein sample (Figure 2.8B).

2.5 Discussion

The predicted domain arrangement from POLIB models is highly divergent from all other characterized Family A Pols. In other well-characterized proteins in this family including *E. coli* Pol I, T7 DNA Pol, and Mammalian Pol γ, the exonuclease domain located underneath the pol domain. Homology and *de novo* modeling of POLIB predict the 3'-5' exonuclease domain to be inserted within the thumb of the pol domain (Figure 2.1, Figure AI.1). We have confidence in this arrangement of catalytic domains due to the generation of similar predictions in both RoseTTAfold and Alphafold (Figure 2.1B, AI.3). Based on the robust exonuclease activity of this enzyme that outcompetes pol activity under physiologically relevant conditions, it is possible the DnaQ-like exonuclease insertion interferes with the ability of the pol domain to carry out its DNA polymerase activity. This structural feature of POLIB was likely developed after the gene duplication event that gave rise to four Pol I-like proteins, giving rise to the specialized function of POLIB in kDNA replication [6].

When we investigated the activity of the POLIB exo domain using recombinant protein, IBPol- consistently showed reduced exonuclease activity compared to wild-type protein (Figure 2.2). This suggests that there may be crosstalk between the two domains, where an amino acid subsitution in one could impact the activity of the other. This is a phenomenon described in other pols such as human Pol γ [28] and may be a more dramatic effect in POLIB due to the divergent arrangement of the two functional domains.

In recombinant IBWT, exo activity unexpectedly prevails over pol activity at higher pH conditions (Figure 2.5). These more basic conditions are likely closer to the physiological pH of the mitochondrial matrix, based on what has been reported in other systems. For example, in *Saccharomyces cerevisiae,* the cytosolic pH is ~7.2 under typical growth conditions while the mitochondrial matrix pH is more basic, ~7.5 [37]. Similarly, in HeLa cells, the pH of the cytosol is ~7.4 and the mitochondrial matrix pH is ~8.0 [38,39]. While the pH of the mitochondrial matrix of trypanosomes has not been experimentally determined, the pH of the cytosol has been reported to be 7.47± 0.06 [40]. We expect that the environment of the mitochondrial matrix in trypanosomes follows a similar pattern seen in the other organisms and is more basic than the cytosol. This suggests that POLIB exo activity on DNA substrates is dominant over pol activity in the alkaline *in vivo* environment.

Interestingly, Gluenz and colleagues defined discrete regions around the kDNA disk via electron microscopy and ethanolic phosphotungstic acid staining where basic proteins accumulated during kDNA replication [41]. These regions likely correspond with the previously described antipodal sites and kinetoflagellar zones defined by immunofluorescence microscopy [42,43]. If POLIB is recruited to these sites, it would likely be an optimal climate for exo activity based on the differential regulation of POLIB activities we demonstrated (Figure 2.5). This localization to particular environmental conditions could be a mechanism of regulating the balance between exo and pol activities for POLIB.

Trypanosomatids also contain two kDNA Pol β-like repair paralogs that differ in their kDNA associated localizations and have distinct optimal conditions, suggesting that they are also specialized for the conditions to which they localize.[4] Additionally, the Pol I-like paralogs POLIC and POLID display spatiotemporal localization to the antipodal sites only during kDNA replication stages of the cell cycle representing well-documented examples of functional regulation of kDNA proteins via localization [44,45]. POLIB also localizes near the kDNA, however a detailed study including the precise kDNA-associated region, and a cell cycle localization pattern has not yet been established [5]. Similar to POLIC and POLID, POLIB may spatiotemporally localize to areas around the disc that have ideal conditions for one of the enzymatic activities or the other.

Another mechanism that could be altering POLIB activity is the use of different available divalent cations. Regulation by divalent cations has been described for other DNA Pols, such as human mitochondrial Pol γ, that has an enhanced ability to perform translesion synthesis (TLS) in the presence of Mn^{2+} compared to Mg^{2+} [46]. In another example, Pol η shifts its activity to ribonucleotide incorporation through an increased affinity to rNTPs in the presence of Mn^{2+} [47]. This modulation in activity for both proteins is likely due to a conformational change of the active site that is allowed due to the more relaxed coordination requirement of Mn^{2+} [48]. We show that POLIB extends more rapidly and with less fidelity when using Mn^{2+} compared to $Mg^{2+,}$ suggesting the active site of POLIB may also be more accommodating to structural rearrangements in the presence of Mn^{2+} (Figure AI.11). For both Pol γ and Pol η, even a relatively low concentration of Mn^{2+} compared to Mg^{2+} in a reaction impacted the activity of the enzyme. Although the concentrations of Mg^{2+} and Mn^{2+} in the *T. brucei* mitochondrion have not been

experimentally determined, both Mg^{2+} and Mn^{2+} are canonically found in mitochondria of other systems [49,50,51,52]. We therefore expect that Mn^{2+} is present in the trypanosome mitochondrion, and that even in relatively small amounts, could impact POLIB *in vivo* to more rapidly extend DNA.

It is possible that POLIB activity is modulated *in vivo* by an accessory subunit, protein-protein interaction, or post-translational modification to improve pol activity. This kind of regulation could be mediated by two arginine methylation sites that have been identified in the N-terminal UCR domain of POLIB (Figure AI.1) [53]. These sites are not present in the truncated recombinant protein purified for the assays in this study. However, it is possible that the presence and methylation status of these residues play a part in regulating the activity of POLIB, either directly through some interdomain interaction on the protein or indirectly through facilitating protein-protein interactions or localization.

This work provides the first demonstration of enzymatic activity of all four mitochondrial Pol I-like proteins of *T. brucei*. One model to explain the presence of four A-Family pols in the mitochondrion of this organism postulates that one protein acts as a replicative polymerase while the other three are playing support roles in kDNA maintenance. However, POLIB, POLIC, and POLID have been shown to be essential for kDNA replication by RNAi [5,8]. In our assay of immunoprecipitated proteins, reactions containing POLIA, POLIC, and POLID resulted in primers that were extended to the end of the DNA template in higher proportions than the reaction products of POLIB (Fig 2.7A). Based on these data, the other three Pol I-like proteins are all more likely than POLIB to

replicate DNA with high processivity, a defining feature of a replicative polymerase. This suggests that out of these polymerases, POLIB is not likely to be a replicative enzyme.

Another model to account for the three essential Family A pols proposes that the enzymes are working cooperatively similar to nuclear DNA replication in other model systems. Pol α from yeast demonstrates increased nucleotidyl incorporation activity from an RNA primer in comparison to a DNA primer, likely due to a higher affinity to A-form helices typically formed by RNA/DNA duplexes [54]. In nuclear DNA replication, Pol α extends the RNA primers for completion of an RNA/DNA primer. POLIB extension activity is increased from an RNA primer, and this activity is less impacted by alkaline pH conditions than extension from a DNA primer (Figures 2.6, 2.7, 2.8). We hypothesize that POLIB may be playing a similar role to Pol a in trypanosomatid mitochondrial DNA replication to facilitate extension by a replicative pol.

When POLIB RNAi is induced, more linearized DNA minicircles are detected in Southern blots than in uninduced cells [9]. Based on the robust exonuclease activity of POLIB that we have demonstrated herein, one explanation for this observation is that POLIB is degrading these linearized molecules *in vivo.* Data from other systems support this hypothesis. In human mitochondria, the exonuclease activity of Pol γ participates along with other exonucleases in the degradation of linearized DNA molecules [55]. After the gene duplication event resulting in the four mitochondrial Pol I-like enzymes in *T. brucei*, POLIB may have retained this function of "cleaning up" molecules resulting from double-strand breaks. The preference for POLIB to degrade ssDNA may indicate that it is working

in tandem with another enzyme that is degrading linearized molecules in the 5'-3' direction, analogous to the function of MGME1 in humans [55].

Because of the divergent domain arrangement of POLIB and its conservation across kinetoplastid organisms, POLIB is an attractive potential drug target for the treatment of diseases caused by medically relevant trypanosomatids. Additional studies are needed to address the essential role of POLIB *in vivo.* For example, complementation experiments can be done with POLIB enzymatic mutants to determine if the exonuclease activity is the essential contribution of POLIB to kDNA replication.

This study represents a new framework for the catalytic activity of a divergent Family A DNA polymerase, adding to the body of knowledge of how these diverse and important proteins can be functionally specialized. Our work also provides new insight into possible mechanisms of the complex process of kDNA replication in trypanosomatid organisms.

2.6 References

1. Parent, J.-S.; Lepage, E.; Brisson, N. Divergent Roles for the Two PolI-Like Organelle DNA Polymerases of Arabidopsis1[W][OA]. *Plant Physiol* **2011**, *156* (1), 254–262. DOI:10.1104/pp.111.173849.
2. Seow, F.; Sato, S.; Janssen, C. S.; Riehle, M. O.; Mukhopadhyay, A.; Phillips, R. S.; Wilson, R. J. M. (Iain); Barrett, M. P. The Plastidic DNA Replication Enzyme Complex of Plasmodium Falciparum. *Molecular and Biochemical Parasitology* **2005**, *141* (2), 145–153. DOI:10.1016/j.molbiopara.2005.02.002.
3. Lukeš, J.; Lys Guilbride, D.; Votýpka, J.; Zíková, A.; Benne, R.; Englund, P. T. Kinetoplast DNA Network: Evolution of an Improbable Structure. *Eukaryot Cell* **2002**, *1* (4), 495–502. DOI:10.1128/EC.1.4.495-502.2002
4. Saxowsky, T. T.; Choudhary, G.; Klingbeil, M. M.; Englund, P. T. Trypanosoma Brucei Has Two Distinct Mitochondrial DNA Polymerase β Enzymes. *Journal of Biological Chemistry* **2003**, *278* (49), 49095–49101. DOI:10.1074/jbc.M308565200.
5. Klingbeil, M. M.; Motyka, S. A.; Englund, P. T. Multiple Mitochondrial DNA Polymerases in Trypanosoma Brucei. *Mol Cell* **2002**, *10* (1), 175–186. DOI:10.1016/s1097-2765(02)00571-3.
6. Poveda, A.; Méndez, M. Á.; Armijos-Jaramillo, V. Analysis of DNA Polymerases Reveals Specific Genes Expansion in Leishmania and Trypanosoma Spp. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 570493. DOI:10.3389/fcimb.2020.570493.
7. Harada, R.; Hirakawa, Y.; Yabuki, A.; Kashiyama, Y.; Maruyama, M.; Onuma, R.; Soukal, P.; Miyagishima, S.; Hampl, V.; Tanifuji, G.; Inagaki, Y. Inventory and Evolution of Mitochondrion-Localized Family A DNA Polymerases in Euglenozoa. *Pathogens* **2020**, *9* (4), 257.DOI:10.3390/pathogens9040257.
8. Chandler, J.; Vandoros, A. V.; Mozeleski, B.; Klingbeil, M. M. Stem-Loop Silencing Reveals That a Third Mitochondrial DNA Polymerase, POLID, Is Required for Kinetoplast DNA Replication in Trypanosomes. *Eukaryot Cell* **2008**, *7* (12), 2141–2146. DOI:10.1128/EC.00199-08.
9. Bruhn, D. F.; Mozeleski, B.; Falkin, L.; Klingbeil, M. M. Mitochondrial DNA Polymerase POLIB Is Essential for Minicircle DNA Replication in African Trypanosomes. *Molecular Microbiology* **2010**, *75* (6), 1414–1425. DOI:10.1111/j.1365-2958.2010.07061.x.
10. Miller, J. C.; Delzell, S. B.; Concepción-Acevedo, J.; Boucher, M. J.; Klingbeil, M. M. A DNA Polymerization-Independent Role for Mitochondrial DNA Polymerase I-like Protein C in African Trypanosomes. *J Cell Sci* **2020**, *133* (9), jcs233072. DOI:10.1242/jcs.233072.
11. Tabor, S.; Huber, H. E.; Richardson, C. C. Escherichia Coli Thioredoxin Confers Processivity on the DNA Polymerase Activity of the Gene 5 Protein of Bacteriophage T7. *Journal of Biological Chemistry* **1987**, *262* (33), 16212–16223. DOI:10.1016/S0021-9258(18)47718-6.
12. Baruch-Torres, N.; Brieba, L. G. Plant Organellar DNA Polymerases Are Replicative and Translesion DNA Synthesis Polymerases. *Nucleic Acids Research* **2017**, *45* (18), 10751. DOI:10.1093/nar/gkx744.

13. Lee, Y.-S.; Kennedy, W. D.; Yin, Y. W. Structural Insight into Processive Human Mitochondrial DNA Synthesis and Disease-Related Polymerase Mutations. *Cell* **2009**, *139* (2), 312–324. DOI:10.1016/j.cell.2009.07.050.
14. Aslett, M.; Aurrecoechea, C.; Berriman, M.; Brestelli, J.; Brunk, B. P.; Carrington, M.; Depledge, D. P.; Fischer, S.; Gajria, B.; Gao, X.; Gardner, M. J.; Gingle, A.; Grant, G.; Harb, O. S.; Heiges, M.; Hertz-Fowler, C.; Houston, R.; Innamorato, F.; Iodice, J.; Kissinger, J. C.; Kraemer, E.; Li, W.; Logan, F. J.; Miller, J. A.; Mitra, S.; Myler, P. J.; Nayak, V.; Pennington, C.; Phan, I.; Pinney, D. F.; Ramasamy, G.; Rogers, M. B.; Roos, D. S.; Ross, C.; Sivam, D.; Smith, D. F.; Srinivasamoorthy, G.; Stoeckert, C. J.; Subramanian, S.; Thibodeau, R.; Tivey, A.; Treatman, C.; Velarde, G.; Wang, H. TriTrypDB: A Functional Genomic Resource for the Trypanosomatidae. *Nucleic Acids Res* **2010**, *38* (Database issue), D457-462. DOI:10.1093/nar/gkp851.
15. Amos, B.; Aurrecoechea, C.; Barba, M.; Barreto, A.; Basenko, E. Y.; Bażant, W.; Belnap, R.; Blevins, A. S.; Böhme, U.; Brestelli, J.; Brunk, B. P.; Caddick, M.; Callan, D.; Campbell, L.; Christensen, M. B.; Christophides, G. K.; Crouch, K.; Davis, K.; DeBarry, J.; Doherty, R.; Duan, Y.; Dunn, M.; Falke, D.; Fisher, S.; Flicek, P.; Fox, B.; Gajria, B.; Giraldo-Calderón, G. I.; Harb, O. S.; Harper, E.; Hertz-Fowler, C.; Hickman, M. J.; Howington, C.; Hu, S.; Humphrey, J.; Iodice, J.; Jones, A.; Judkins, J.; Kelly, S. A.; Kissinger, J. C.; Kwon, D. K.; Lamoureux, K.; Lawson, D.; Li, W.; Lies, K.; Lodha, D.; Long, J.; MacCallum, R. M.; Maslen, G.; McDowell, M. A.; Nabrzyski, J.; Roos, D. S.; Rund, S. S. C.; Schulman, S. W.; Shanmugasundram, A.; Sitnik, V.; Spruill, D.; Starns, D.; Stoeckert, C. J., Jr; Tomko, S. S.; Wang, H.; Warrenfeltz, S.; Wieck, R.; Wilkinson, P. A.; Xu, L.; Zheng, J. VEuPathDB: The Eukaryotic Pathogen, Vector and Host Bioinformatics Resource Center. *Nucleic Acids Research* **2022**, *50* (D1), D898–D911. DOI:10.1093/nar/gkab929.
16. Papadopoulos, J. S.; Agarwala, R. COBALT: Constraint-Based Alignment Tool for Multiple Protein Sequences. *Bioinformatics* **2007**, *23* (9), 1073–1079. DOI:10.1093/bioinformatics/btm076.
17. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; Millán, C.; Park, H.; Adams, C.; Glassman, C. R.; DeGiovanni, A.; Pereira, J. H.; Rodrigues, A. V.; van Dijk, A. A.; Ebrecht, A. C.; Opperman, D. J.; Sagmeister, T.; Buhlheller, C.; Pavkov-Keller, T.; Rathinaswamy, M. K.; Dalwadi, U.; Yip, C. K.; Burke, J. E.; Garcia, K. C.; Grishin, N. V.; Adams, P. D.; Read, R. J.; Baker, D. Accurate Prediction of Protein Structures and Interactions Using a 3-Track Neural Network. *Science* **2021**, *373* (6557), 871–876. DOI:10.1126/science.abj8754.
18. Humphreys, I. R.; Pei, J.; Baek, M.; Krishnakumar, A.; Anishchenko, I.; Ovchinnikov, S.; Zhang, J.; Ness, T. J.; Banjade, S.; Bagde, S. R.; Stancheva, V. G.; Li, X.-H.; Liu, K.; Zheng, Z.; Barrero, D. J.; Roy, U.; Kuper, J.; Fernández, I. S.; Szakal, B.; Branzei, D.; Rizo, J.; Kisker, C.; Greene, E. C.; Biggins, S.; Keeney, S.; Miller, E. A.; Fromme, J. C.; Hendrickson, T. L.; Cong, Q.; Baker, D. Computed Structures of Core Eukaryotic Protein Complexes. *Science* **2021**, *374* (6573), eabm4805. DOI:10.1126/science.abm4805.
19. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat Methods* **2012**, *9* (7), 671–675. DOI:10.1038/nmeth.2089

20. Jensen, B. C.; Kifer, C. T.; Brekken, D. L.; Randall, A. C.; Wang, Q.; Drees, B. L.; Parsons, M. Characterization of Protein Kinase CK2 from Trypanosoma Brucei. *Mol Biochem Parasitol* **2007**, *151* (1), 28–40. DOI:10.1016/j.molbiopara.2006.10.002.
21. Wirtz, E.; Leal, S.; Ochatt, C.; Cross, GeorgeA. M. A Tightly Regulated Inducible Expression System for Conditional Gene Knock-Outs and Dominant-Negative Genetics in Trypanosoma Brucei. *Molecular and Biochemical Parasitology* **1999**, *99* (1), 89–101. 10.1016/S0166-6851(99)00002-X. DOI:10.1016/s0166-6851(99)00002-x.
22. Oates, M. E.; Stahlhacke, J.; Vavoulis, D. V.; Smithers, B.; Rackham, O. J. L.; Sardar, A. J.; Zaucha, J.; Thurlby, N.; Fang, H.; Gough, J. The SUPERFAMILY 1.75 Database in 2014: A Doubling of Data. *Nucleic Acids Research* **2015**, *43* (D1), D227–D233. DOI:10.1093/nar/gku1041.
23. Almagro Armenteros, J. J.; Salvatore, M.; Emanuelsson, O.; Winther, O.; von Heijne, G.; Elofsson, A.; Nielsen, H. Detecting Sequence Signals in Targeting Peptides Using Deep Learning. *Life Sci Alliance* **2019**, *2* (5), e201900429. DOI:10.26508/lsa.201900429.
24. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. DOI:10.1038/s41586-021-03819-2.
25. Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. LDDT: A Local Superposition-Free Score for Comparing Protein Structures and Models Using Distance Difference Tests. *Bioinformatics* **2013**, *29* (21), 2722–2728. DOI:10.1093/bioinformatics/btt473.
26. Hogg, M.; Seki, M.; Wood, R. D.; Doublié, S.; Wallace, S. S. Lesion Bypass Activity of DNA Polymerase θ (POLQ) Is an Intrinsic Property of the Pol Domain and Depends on Unique Sequence Inserts. *Journal of Molecular Biology* **2011**, *405* (3), 642–652. DOI:10.1016/j.jmb.2010.10.041.
27. Brautigam, C. A.; Sun, S.; Piccirilli, J. A.; Steitz, T. A. Structures of Normal Single-Stranded DNA and Deoxyribo-3‘-S-Phosphorothiolates Bound to the 3‘-5‘ Exonucleolytic Active Site of DNA Polymerase I from Escherichia Coli,. *Biochemistry* **1999**, *38* (2), 696–704. DOI:10.1021/bi981537g.
28. Sowers, M. L.; Anderson, A. P. P.; Wrabl, J. O.; Yin, Y. W. Networked Communication between Polymerase and Exonuclease Active Sites in Human Mitochondrial DNA Polymerase. *J. Am. Chem. Soc.* **2019**, *141* (27), 10821–10829. DOI:10.1021/jacs.9b04655.
29. Yang, W. Nucleases: Diversity of Structure, Function and Mechanism. *Q Rev Biophys* **2011**, *44* (1). DOI:10.1017/S0033583510000181.
30. Hemphill, W. O.; Perrino, F. W. Chapter Eight - Measuring TREX1 and TREX2 Exonuclease Activities. In *Methods in Enzymology*; Sohn, J., Ed.; DNA Sensors and Inflammasomes; Academic Press, 2019; Vol. 625, pp 109–133. DOI:10.1016/bs.mie.2019.05.004.

31. Johnson, A. A.; Johnson, K. A. Exonuclease Proofreading by Human Mitochondrial DNA Polymerase. *Journal of Biological Chemistry* **2001**, *276* (41), 38097–38107. DOI:10.1074/jbc.M106046200.
32. Singh, K.; Modak, M. J. Contribution of Polar Residues of the J-Helix in the 3‘−5‘ Exonuclease Activity of Escherichia Coli DNA Polymerase I (Klenow Fragment): Q677 Regulates the Removal of Terminal Mismatch. *Biochemistry* **2005**, *44* (22), 8101–8110. DOI:10.1021/bi050140r.
33. Teklemariam, T. A.; Rivera, O. D.; Nelson, S. W. Chapter Five - Kinetic Analysis of the Exonuclease Activity of the Bacteriophage T4 Mre11–Rad50 Complex. In *Methods in Enzymology*; Spies, M., Malkova, A., Eds.; Mechanisms of DNA Recombination and Genome Rearrangements: Methods to Study Homologous Recombination; Academic Press, 2018; Vol. 600, pp 135–156. DOI:10.1016/bs.mie.2017.12.007.
34. McGraw, N. J.; Bailey, J. N.; Cleaves, G. R.; Dembinski, D. R.; Gocke, C. R.; Joliffe, L. K.; MacWright, R. S.; McAllister, W. T. Sequence and Analysis of the Gene for Bacteriophage T3 RNA Polymerase. *Nucleic Acids Res* **1985**, *13* (18), 6753–6766. DOI:10.1093/nar/13.18.6753
35. Wijngaert, B. D.; Sultana, S.; Singh, A.; Dharia, C.; Vanbuel, H.; Shen, J.; Vasilchuk, D.; Martinez, S. E.; Kandiah, E.; Patel, S. S.; Das, K. Cryo-EM Structures Reveal Transcription Initiation Steps by Yeast Mitochondrial RNA Polymerase. *Mol Cell* **2021**, *81* (2), 268-280.e5. DOI:10.1016/j.molcel.2020.11.016
36. Porcelli, A. M.; Ghelli, A.; Zanna, C.; Pinton, P.; Rizzuto, R.; Rugolo, M. PH Difference across the Outer Mitochondrial Membrane Measured with a Green Fluorescent Protein Mutant. *Biochemical and Biophysical Research Communications* **2005**, *326* (4), 799–804. DOI:10.1016/j.bbrc.2004.11.105.
37. Orij, R.; Postmus, J.; Ter Beek, A.; Brul, S.; Smits, G. J. In Vivo Measurement of Cytosolic and Mitochondrial PH Using a PH-Sensitive GFP Derivative in Saccharomyces Cerevisiae Reveals a Relation between Intracellular PH and Growth. *Microbiology (Reading)* **2009**, *155* (Pt 1), 268–278. DOI:10.1099/mic.0.022038-0.
38. Llopis, J.; McCaffery, J. M.; Miyawaki, A.; Farquhar, M. G.; Tsien, R. Y. Measurement of Cytosolic, Mitochondrial, and Golgi PH in Single Living Cells with Green Fluorescent Proteins. *Proceedings of the National Academy of Sciences* **1998**, *95* (12), 6803–6808. DOI:10.1073/pnas.95.12.6803.
39. Abad, M. F. C.; Di Benedetto, G.; Magalhães, P. J.; Filippin, L.; Pozzan, T. Mitochondrial PH Monitored by a New Engineered Green Fluorescent Protein Mutant*. *Journal of Biological Chemistry* **2004**, *279* (12), 11521–11529. DOI:10.1074/jbc.M306766200.
40. Vanderheyden, N.; Wong, J.; Docampo, R. A Pyruvate-Proton Symport and an H+-ATPase Regulate the Intracellular PH of Trypanosoma Brucei at Different Stages of Its Life Cycle. *Biochem J* **2000**, *346* (Pt 1), 53–62. DOI:10.1042/bj3460053
41. Gluenz, E.; Shaw, M. K.; Gull, K. Structural Asymmetry and Discrete Nucleic Acid Subdomains in the Trypanosoma Brucei Kinetoplast. *Mol Microbiol* **2007**, *64* (6), 1529–1539. DOI:10.1111/j.1365-2958.2007.05749.x.

42. Melendy, T.; Sheline, C.; Ray, D. S. Localization of a Type II DNA Topoisomerase to Two Sites at the Periphery of the Kinetoplast DNA of Crithidia Fasciculata. *Cell* **1988**, *55* (6), 1083–1088. DOI:10.1016/0092-8674(88)90252-8.
43. Drew, M. E.; Englund, P. T. Intramitochondrial Location and Dynamics of Crithidia Fasciculata Kinetoplast Minicircle Replication Intermediates. *J Cell Biol* **2001**, *153* (4), 735–744. DOI:10.1083/jcb.153.4.735.
44. Concepción-Acevedo, J.; Luo, J.; Klingbeil, M. M. Dynamic Localization of Trypanosoma Brucei Mitochondrial DNA Polymerase ID. *Eukaryot Cell* **2012**, *11* (7), 844–855. DOI:10.1128/EC.05291-11.
45. Concepción-Acevedo, J.; Miller, J. C.; Boucher, M. J.; Klingbeil, M. M. Cell Cycle Localization Dynamics of Mitochondrial DNA Polymerase IC in African Trypanosomes. *Mol Biol Cell* **2018**, *29* (21), 2540–2552. DOI:10.1091/mbc.E18-02-0127.
46. Park, J.; Baruch-Torres, N.; Iwai, S.; Herrmann, G. K.; Brieba, L. G.; Yin, Y. W. Human Mitochondrial DNA Polymerase Metal Dependent UV Lesion Bypassing Ability. *Front Mol Biosci* **2022**, *9*, 808036. DOI:10.3389/fmolb.2022.808036.
47. Balint, E.; Unk, I. Selective Metal Ion Utilization Contributes to the Transformation of the Activity of Yeast Polymerase η from DNA Polymerization toward RNA Polymerization. *Int J Mol Sci* **2020**, *21* (21), 8248. DOI:10.3390/ijms21218248.
48. Vaisman, A.; Ling, H.; Woodgate, R.; Yang, W. Fidelity of Dpo4: Effect of Metal Ions, Nucleotide Selection and Pyrophosphorolysis. *EMBO J* **2005**, *24* (17), 2957–2967. DOI:10.1038/sj.emboj.7600786.
49. Konji, V.; Montag, A.; Sandri, G.; Nordenbrand, K.; Ernster, L. Transport of Ca2+ and Mn2+ by Mitochondria from Rat Liver, Heart and Brain. *Biochimie* **1985**, *67* (12), 1241–1250. DOI:10.1016/S0300-9084(85)80133-4.
50. Corkey, B. E.; Duszynski, J.; Rich, T. L.; Matschinsky, B.; Williamson, J. R. Regulation of Free and Bound Magnesium in Rat Hepatocytes and Isolated Mitochondria. *Journal of Biological Chemistry* **1986**, *261* (6), 2567–2574. DOI:10.1016/S0021-9258(17)35825-8.
51. Gunter, T. E.; Miller, L. M.; Gavin, C. E.; Eliseev, R.; Salter, J.; Buntinas, L.; Alexandrov, A.; Hammond, S.; Gunter, K. K. Determination of the Oxidation States of Manganese in Brain, Liver, and Heart Mitochondria. *Journal of Neurochemistry* **2004**, *88* (2), 266–280. DOI:10.1046/j.1471-4159.2003.02122.x.
52. Romani, A. M. P. Cellular Magnesium Homeostasis. *Archives of Biochemistry and Biophysics* **2011**, *512* (1), 1–23. DOI:10.1016/j.abb.2011.05.010.
53. Fisk, J. C.; Li, J.; Wang, H.; Aletta, J. M.; Qu, J.; Read, L. K. Proteomic Analysis Reveals Diverse Classes of Arginine Methylproteins in Mitochondria of Trypanosomes. *Mol Cell Proteomics* **2013**, *12* (2), 302–311. DOI:10.1074/mcp.M112.022533.
54. Perera, R. L.; Torella, R.; Klinge, S.; Kilkenny, M. L.; Maman, J. D.; Pellegrini, L. Mechanism for Priming DNA Synthesis by Yeast DNA Polymerase α. *eLife* **2013**, *2*, e00482. DOI:10.7554/eLife.00482.
55. Peeva, V.; Blei, D.; Trombly, G.; Corsi, S.; Szukszto, M. J.; Rebelo-Guiomar, P.; Gammage, P. A.; Kudin, A. P.; Becker, C.; Altmüller, J.; Minczuk, M.; Zsurka, G.; Kunz, W. S. Linear Mitochondrial DNA Is Rapidly Degraded by Components of the Replication Machinery. *Nat Commun* **2018**, *9* (1), 1727. DOI:10.1038/s41467-018-04131-w.

56. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Research* **2022**, *50* (D1), D439–D444. DOI:10.1093/nar/gkab1061.

Chapter 3: Future Study of POLIB and kDNA Replication

In the previous chapters, we outlined what is known about the kDNA replisome of *T. brucei* and generated new hypotheses for the mechanisms of kDNA replication based on the characteristics of POLIB. Here we propose future studies that are critical for further illumination of this enigmatic process and understanding the potential value of kDNA enzymes as drug targets.

3.1 Localization Studies

In 2012, Jeniffer Concepción-Acevedo used immunofluorescence techniques to track epitope-tagged POLID throughout the cell cycle of *T. brucei* [1]. This study revealed that POLID is distributed throughout the mitochondrial matrix until kDNA S-phase where it is redistributed to two foci at the antipodal sites, likely to participate in minicircle replication. Later, SIM microscopy allowed for high-resolution analysis of the localization of POLIC through the course of kDNA S-phase in relation to other characterized proteins and the kDNA disc. This study revealed that POLIC begins kDNA S-phase in the kinetoflagellar zone and migrates to antipodal sites in later stages of kDNA replication [2]. We hypothesize that other kDNA replication proteins have cell cycle-dependent localization patterns to the fine structures surrounding the kDNA network. High- resolution microscopy techniques such as SIM or expansion microscopy are excellent tools to facilitate the study of kDNA replication proteins.

POLIB localization throughout the cell cycle or through the course of kDNA S-phase has not been studied in detail, but should be to provide important evidence to support the

hypothesis that it is functioning as part of a multi-polymerase replisome in kDNA replication. Based on our hypothesis that POLIB extends an RNA primer, we would expect POLIB to localize to the antipodal sites for kDNA S-phase where both PRI1, thought to be the maxicircle primase, and PRI2, thought to be the minicircle primase, localize. POLID, which is the most likely candidate identified to be a replicative polymerase, also localizes to the antipodal sites, putting it in an ideal position to take over replication from POLIB after primer extension.

We also hypothesize that the relatively robust exonuclease activity of POLIB could be degrading linearized or otherwise damaged kDNA molecules. It is less clear where this activity would be taking place in relation to the kDNA disc based on what we know about *T. brucei* mitochondrial DNA repair. However, in *T. cruzi*, when DNA damage is caused by introduction of reactive oxygen species, the DNA repair enzyme Pol β re-localizes to the kinetoflagellar zone (KFZ). Based on this pattern, it is thought that the KFZ may be a DNA repair hub in the mitochondrion of *T. cruzi*. *T. brucei* could feasibly also use the KFZ as a DNA repair hub, and this is perhaps where POLIB would localize to degrade damaged DNA. Studying the localization of POLIB throughout the cell cycle, and also in conditions that induce DNA damage may provide evidence to support our current hypotheses for the *in vivo* functions of POLIB.

3.2 Protein-protein Interactions

Unlike mammalian mitochondrial DNA replication, which occurs continuously throughout the cell cycle, the kDNA network is replicated in a highly coordinated fashion just once per cell cycle in kDNA S-phase. Some critical proteins involved in kDNA

replication are only recruited to the disc for this phase. kDNA replication proteins that would have to physically interact with one another for successful replication may only being doing so during this relatively short time frame. These conditions present challenges for determining protein-protein interactions, as only a small percentage of cells of an unsynchronized population would be actively replicating kDNA, and have the replisome intact to capture. Therefore, trategies such as co-immunoprecipitations would result in the recovery of mostly protein that is not part of an active replisome. However, such studies remain an attractive approach to establish how kDNA replication machinery physically and functionally interact. In 2017, a study using protein correlation profiling mass spectrometry and machine learning to predict protein complexes in *T. brucei* generated a prediction that POLIB complexes with two other proteins: kinetoplastid-associated protein 3, KAP3, and a putative RNA binding protein DRBD10 [3]. Experiments that directly test whether this complex indeed forms *in vivo* would determine whether the function of POLIB is tied to these proteins, perhaps implicating it in a role in RNA metabolism.

An alternative approach to understanding the protein-protein interactions of the kDNA replisome would be to use a model trypanosomatid, *Crithidia fasciculata,* that can be grown at high cell densities in economically friendly media and can be reliably synchronized. Using this system, cells could be harvested for analysis in kDNA S-phase to enrich the number of replisome-related complexes formed. However, in order to draw comparisons of the mitochondrial DNA replisome of *T. brucei* and *C. fasciculata*, studies will also need to be done to establish whether the kDNA replication proteins that have been established as essential for kDNA maintenance in *T. brucei* have analogous roles in

C. fasciculata. These studies would contribute understanding to the common mechanisms of kDNA replication of trypanosomatids and could lead to pan-trypanosomal therapies.

3.3 RNAi Complementation

In a recent study, complementing POLIC knockdown with exogenously expressed POLIC mutants provided useful data into how two domains of this pol are contributing to its *in vivo* role [4]. Similar studies with POLIB and POLID should be pursued to determine how conserved structural features of these enzymes contribute to kDNA replication. Initial attempts to complement POLIB knockdown did not result in successful rescue of the knockdown phenotype, likely due to insufficient levels of exogenous protein being expressed [5]. However, the design of these complementation cell lines should be redesigned to improve protein expression to investigate the role of POLIB *in vivo.* A complementation study will be particularly interesting for POLIB catalytic mutants to determine whether its highly divergent exonuclease domain is the sole essential feature for kDNA maintenance. Based on the dominant negative phenotype caused by overexpression of the exonuclease-dead variant of POLIB, we can assume that an exonuclease-ablated variant will not rescue the knockdown kDNA replication defect phenotype. However, complementation with a polymerase-ablated variant will provide valuable information about the importance of this domain.

These studies could also be used to determine if the N-terminal uncharacterized region present in all of the Pol I like kDNA proteins are critical for their respective localization patterns, as is the case for POLIC. One more open question that can be addressed using

deletion mutants in a complementation system is the role of the thumb insertions of POLIC and POLID that do not have homology to other characterized enzymes. Dissecting each of the Pol-I like polymerases in this way will likely lead to not only a better understanding of kDNA replication but a better understanding of the mutability and evolution of these divergent DNA pols.

3.4 Drug Targeting Potential of kDNA Replication Proteins

Proteins and processes surrounding mitochondrial DNA replication in *T. brucei* are significantly divergent from those processes in humans and other animal hosts of the parasites. Furthermore, successful maintenance of the kDNA network is essential for the survival of these parasites. These features have long made kDNA replication an enticing potential drug targets for treatment of diseases caused by trypanosomes. Recently, a study revealed that some commonly used trypanocidal drugs indeed impact the kDNA with some specificity, solidifying the value of this strategy [6]. It is well established that the targeting of some kDNA replication enzymes by RNAi results in a loss of fitness of trypanosomes, following the loss or disruption of the kDNA network. These enzymes that are essential for kDNA replication and maintenance could be studied further in inhibition screens, particularly those with enzymatic activity that could be measured easily in a high-throughput format. Identifying drugs that target kDNA proteins with high specificity has the potential to provide new therapies with fewer toxic side effects, addressing a long-term goal of the trypanosome biology field.

References

1. Concepción-Acevedo, J.; Luo, J.; Klingbeil, M. M. Dynamic Localization of Trypanosoma Brucei Mitochondrial DNA Polymerase ID. *Eukaryot Cell* **2012**, *11* (7), 844–855. https://doi.org/10.1128/EC.05291-11.
2. Concepción-Acevedo, J.; Miller, J. C.; Boucher, M. J.; Klingbeil, M. M. Cell Cycle Localization Dynamics of Mitochondrial DNA Polymerase IC in African Trypanosomes. *Mol Biol Cell* **2018**, *29* (21), 2540–2552. https://doi.org/10.1091/mbc.E18-02-0127.
3. Crozier, T. W. M.; Tinti, M.; Larance, M.; Lamond, A. I.; Ferguson, M. A. J. Prediction of Protein Complexes in Trypanosoma Brucei by Protein Correlation Profiling Mass Spectrometry and Machine Learning. *Mol Cell Proteomics* **2017**, *16* (12), 2254–2267. https://doi.org/10.1074/mcp.O117.068122.
4. Miller, J. C.; Delzell, S. B.; Concepción-Acevedo, J.; Boucher, M. J.; Klingbeil, M. M. A DNA Polymerization-Independent Role for Mitochondrial DNA Polymerase I-like Protein C in African Trypanosomes. *J Cell Sci* **2020**, *133* (9), jcs233072. https://doi.org/10.1242/jcs.233072.
5. Armstrong, R. Structure-Function Studies of the Trypanosome Mitochondrial Replication Protein POLIB. UMass Amherst, 2021.
6. Miskinyte, M.; Dawson, J. C.; Makda, A.; Doughty-Shenton, D.; Carragher, N. O.; Schnaufer, A. A Novel High-Content Phenotypic Screen To Identify Inhibitors of Mitochondrial DNA Maintenance in Trypanosomes. *Antimicrob Agents Chemother* **2022**, *66* (2), e01980-21. https://doi.org/10.1128/AAC.01980-21.

Appendix I Supplementary Data for Chapter 2

```
[task]
task = fit
data = progress
model = 2step?
[mechanism]
E*DDDDDDDDDDDDDDDDDDDDDD <===> E*  +          DDDDDDDDDDDDDDDDDDDDDD  : k-2 k1
E*DDDDDDDDDDDDDDDDDDDDDD -->                 EDDDDDDDDDDDDDDDDDDDDDD : kconf
EDDDDDDDDDDDDDDDDDDDDDD -->                  EDDDDDDDDDDDDDDDDDDDDDD + R   : kexo
EDDDDDDDDDDDDDDDDDDDDDD <==> E  +            DDDDDDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDDDDD -->                   EDDDDDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDDDDDDD <==> E  +             DDDDDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDDDDD -->                   EDDDDDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDDDDDD <==> E  +              DDDDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDDDD -->                    EDDDDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDDDDD <==> E  +               DDDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDDD -->                     EDDDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDDDD <==> E  +                DDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDD -->                      EDDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDDD <==> E  +                 DDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDD -->                       EDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDD <==> E  +                  DDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDD -->                        EDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDD <==> E  +                   DDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDD -->                         EDDDDDDDDDDDDDDD + Q   : kexo
EDDDDDDDDDDDDDD <==> E  +                    DDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDD -->                          EDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDD <==> E  +                     DDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDD -->                           EDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDD <==> E  +                      DDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDD -->                            EDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDD <==> E  +                       DDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDD  -->                            EDDDDDDDDDD + P   : kexo
EDDDDDDDDDD <==> E  +                        DDDDDDDDDD  : k-2 k1
EDDDDDDDDDD  -->                             EDDDDDDDDDD + P   : kexo
EDDDDDDDDD <==> E  +                         DDDDDDDDD  : k-2 k1
EDDDDDDDDD  -->                              EDDDDDDDDD + P   : kexo
EDDDDDDDD <==> E  +                          DDDDDDDD  : k-2 k1
EDDDDDDDD  -->                               EDDDDDDDD + P   : kexo
EDDDDDDD <==> E  +                           DDDDDDD  : k-2 k1
EDDDDDDD  -->                                EDDDDDDD + P   : kexo
EDDDDDD <==> E  +                            DDDDDD  : k-2 k1
EDDDDDD  -->                                 EDDDDDD + P   : kexo
EDDDDD <==> E  +                             DDDDD  : k-2 k1
EDDDDD  -->                                  EDDDDD + P   : kexo
EDDDD <==> E  +                              DDDD  : k-2 k1
EDDDD  -->                                   EDDDD + P   : kexo
EDDD <==> E  +                               DDD : k-2 k1
EDDD  -->                                    EDDD + P   : kexo
EDD <==> E  +                                DD : k-2 k1
EDD  -->                                     E  +  D  +  P   : kexo
[constants]
k1 = 50
kexo = 1??
k-2 = 34??
kconf = 1??
```

Figure AI.1: Dynafit script for global fitting of 1AP and 2AP progress curves for the ssDNA and blunt dsDNA substrates.

```
[task]
task = fit
data = progress
model = 2step?
[mechanism]
EDDDDDDDDDDDDDDDDDDDDD <===> E  +        DDDDDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDDDDD -->               EDDDDDDDDDDDDDDDDDDDDD + R   : kexo
EDDDDDDDDDDDDDDDDDDDD <===> E  +         DDDDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDDDD -->                EDDDDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDDDD <==> E  +            DDDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDDD -->                  EDDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDDD <==> E  +             DDDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDDD -->                   EDDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDDD <==> E  +              DDDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDDD -->                    EDDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDDD <==> E  +               DDDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDDD -->                     EDDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDDD <==> E  +                DDDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDDD -->                      EDDDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDDDD <==> E  +                 DDDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDDD -->                       EDDDDDDDDDDDDD + Q   : kexo
EDDDDDDDDDDDD <==> E  +                  DDDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDDD -->                        EDDDDDDDDDDDD + P   : kexo
EDDDDDDDDDDD <==> E  +                   DDDDDDDDDDD  : k-2 k1
EDDDDDDDDDDD -->                         EDDDDDDDDDDD + P   : kexo
EDDDDDDDDDD <==> E  +                    DDDDDDDDDD  : k-2 k1
EDDDDDDDDDD  -->                         EDDDDDDDDDD + P   : kexo
EDDDDDDDD <==> E  +                      DDDDDDDDD  : k-2 k1
EDDDDDDDD  -->                           EDDDDDDDD + P   : kexo
EDDDDDDD <==> E  +                       DDDDDDDD  : k-2 k1
EDDDDDDD  -->                            EDDDDDDD + P   : kexo
EDDDDDD <==> E  +                        DDDDDDD  : k-2 k1
EDDDDDD  -->                             EDDDDDD + P   : kexo
EDDDDD <==> E  +                         DDDDDD  : k-2 k1
EDDDDD  -->                              EDDDDD + P   : kexo
EDDDD <==> E  +                          DDDDD  : k-2 k1
EDDDD  -->                               EDDDD + P   : kexo
EDDD <==> E  +                           DDDD  : k-2 k1
EDDD  -->                                EDDD + P   : kexo
EDD <==> E  +                            DDD : k-2 k1
EDD  -->                                 EDD + P  : kexo
E + P <==>                               DD : k-2 k1
[constants]                              E + D + P : kexo
k1 = 50                                  EP : k1 k-3;
kexo = 1??
k-2 = 34??
k-3 =5??
```

Figure AI.2: Dynafit script for global fitting of 1AP and 2AP progress curves for the recessed dsDNA substrate.

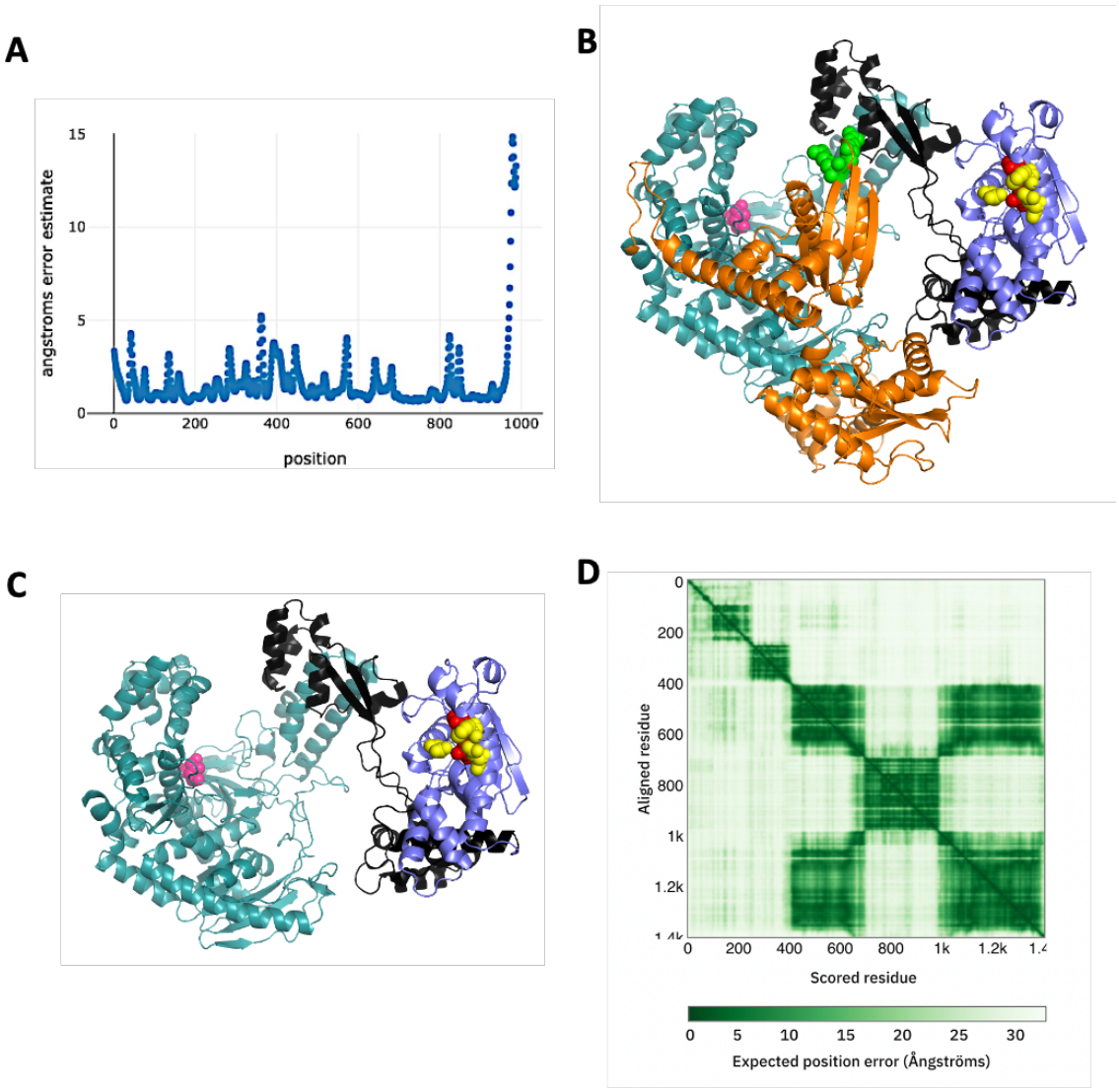
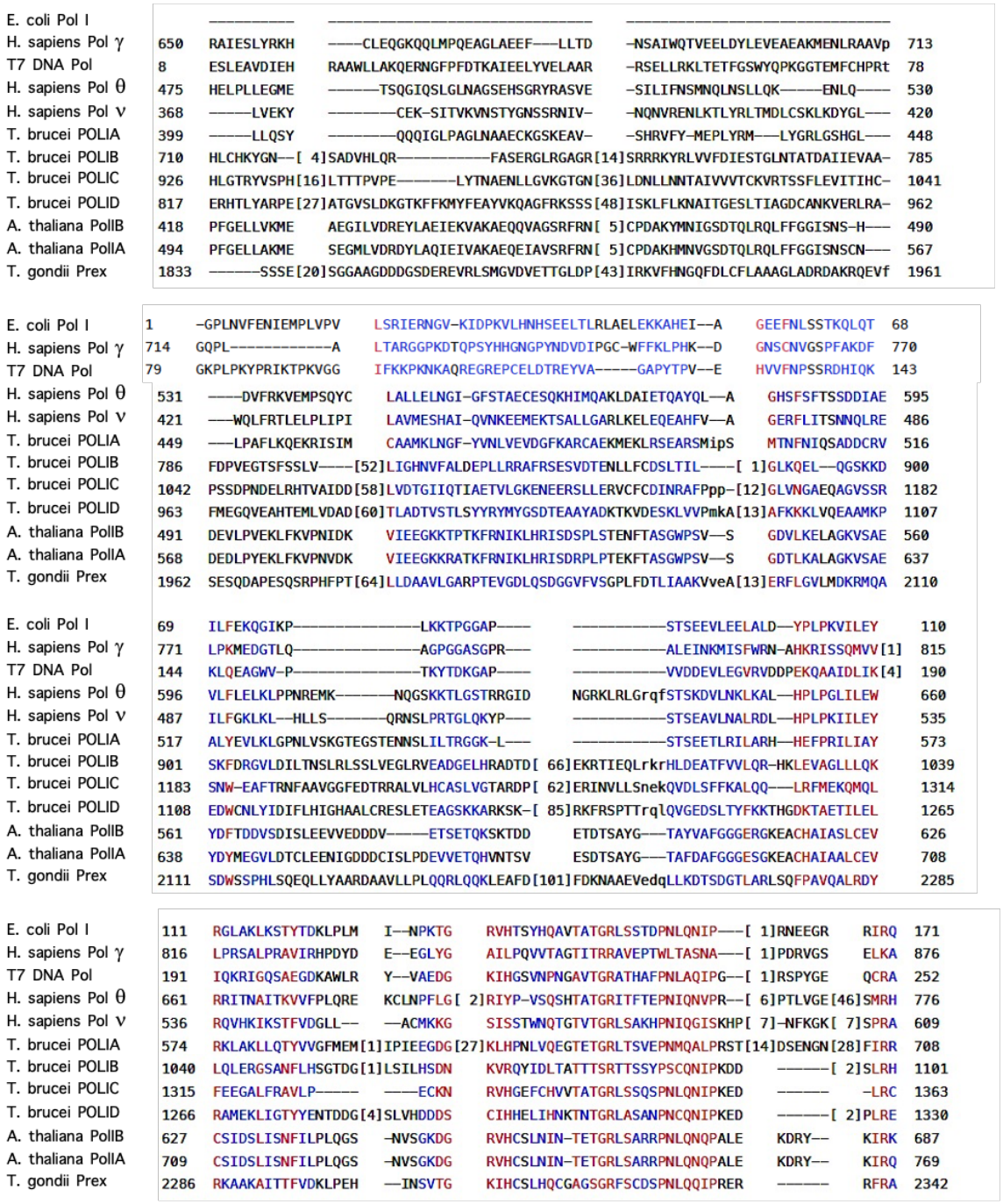


Figure AI.3: Quality of Rosettafold model and comparison of domain structure with Alphafold l model (A) Per Residue Error Estimate of TbPOLIB RoseTTAfold Model displayed in Figure 1. (B) Alphafold Model of Full-length POLIB. Model retrieved from Alphafold database (Entry: Mitochondrial DNA pol I protein B, Tb11.02.2300, *Trypanosoma brucei brucei* (strain 927/4 GUTat10.1).[56] Teal, pol domain; purple, portion of insertion with homology to an exonuclease domain; black, thumb insertion linker regions; Magenta, pol domain active sites mutated for IBpol- variant, red, exo catalytic residues for IBexo- variant; yellow, additional exonuclease residues of the DEDDh motif. (C) Alphafold model of POLIB (residues 414-1400). Same color scheme as in B. Orange, uncharacterized C-terminal domain (UCR); bright green, arginine methylation sites (R266, R270).[53] (D) Predicted aligned error for full length Alphafold model.


E. coli Pol I
H. sapiens Pol γ
T7 DNA Pol
H. sapiens Pol θ
H. sapiens Pol ν
T. brucei POLIA
T. brucei POLIB
T. brucei POLIC
T. brucei POLID
A. thaliana PolIB
A. thaliana PolIA
T. gondii Prex

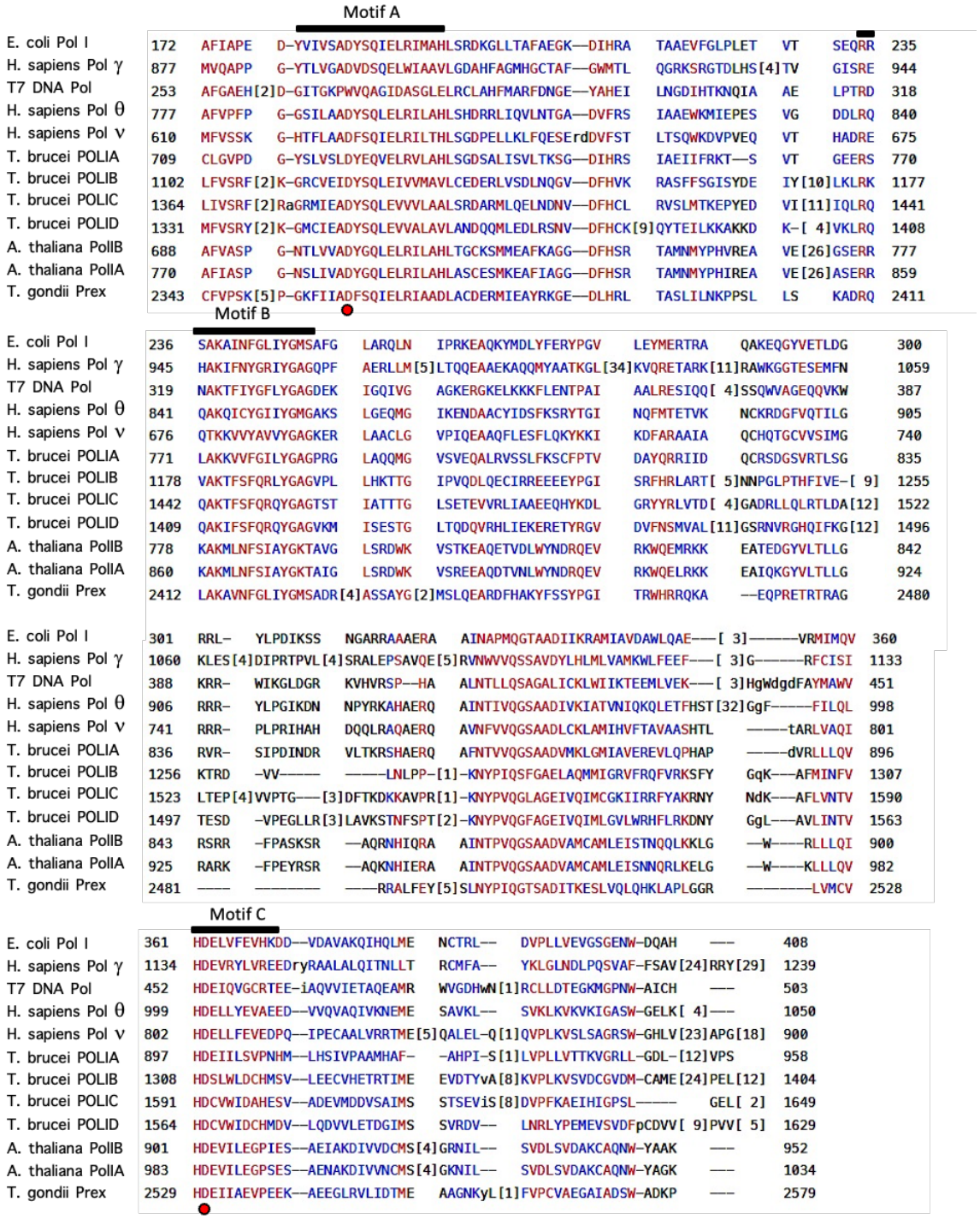
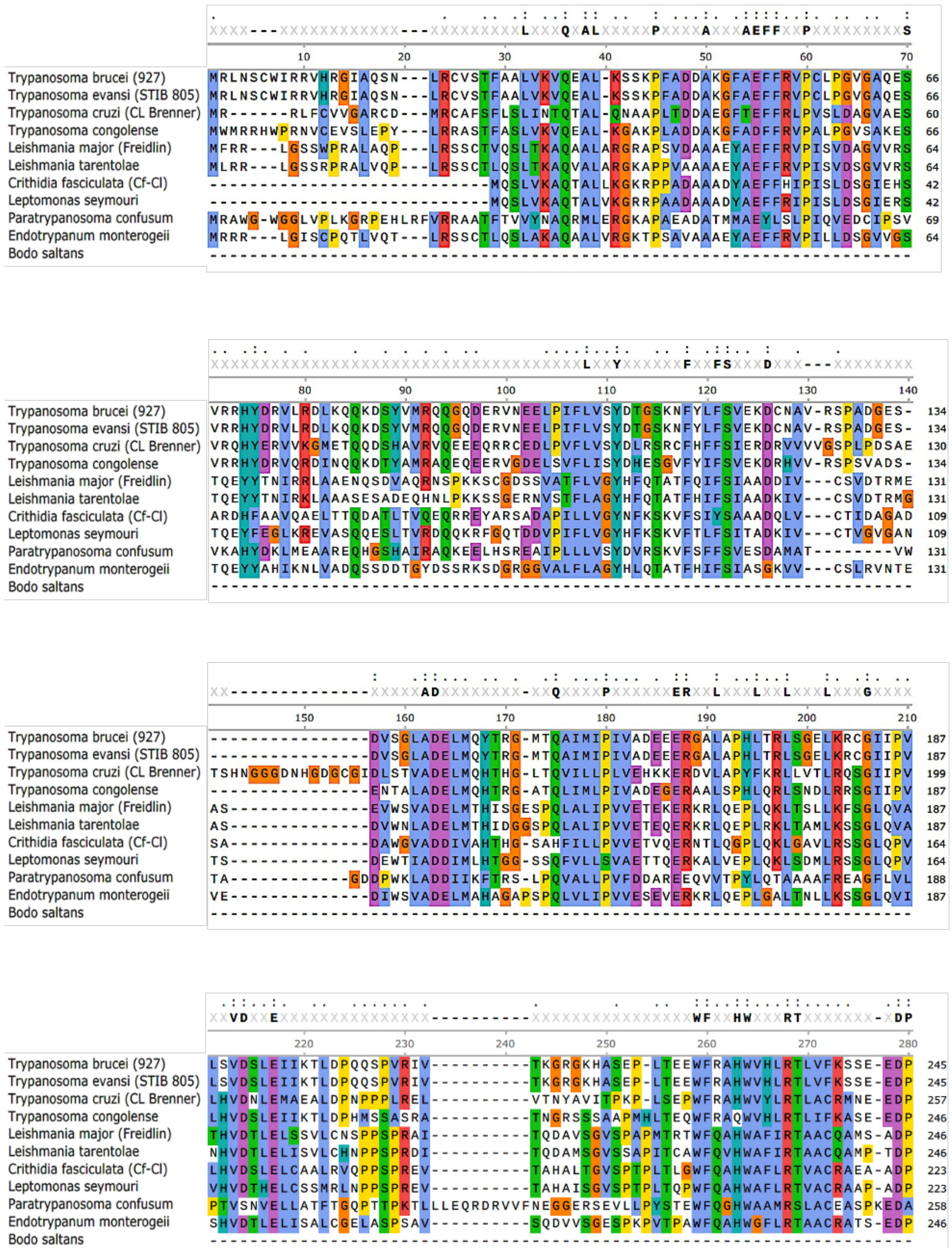
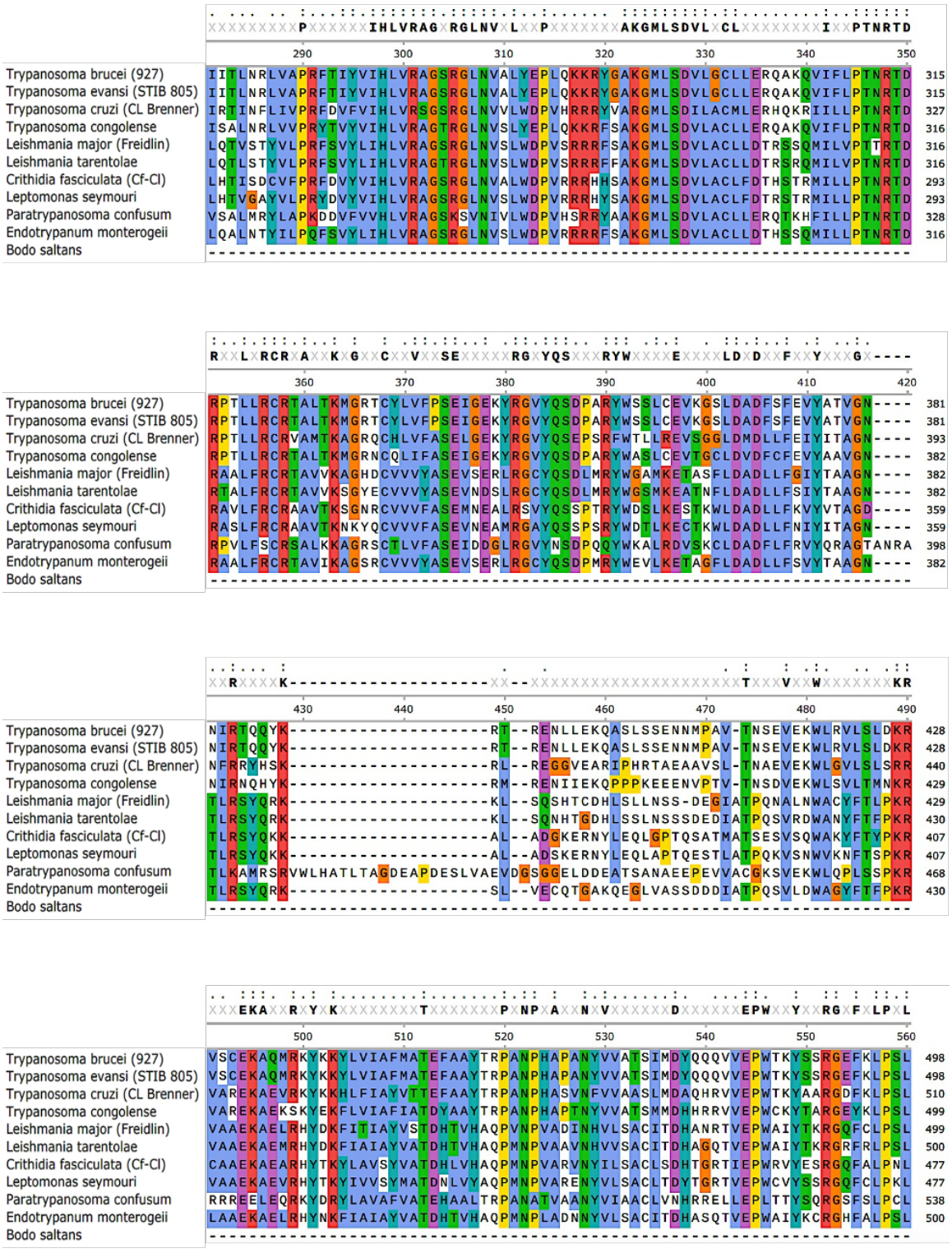
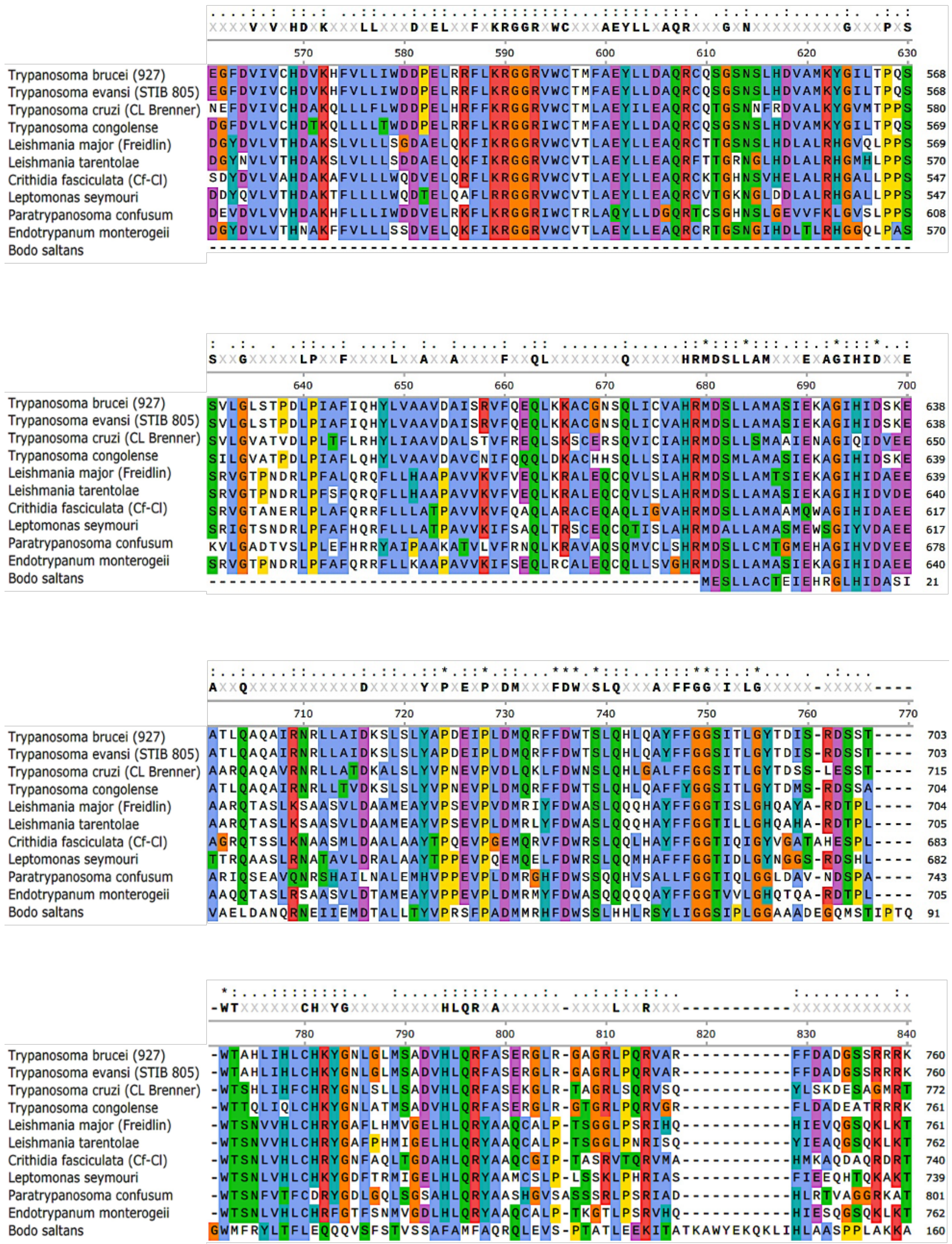
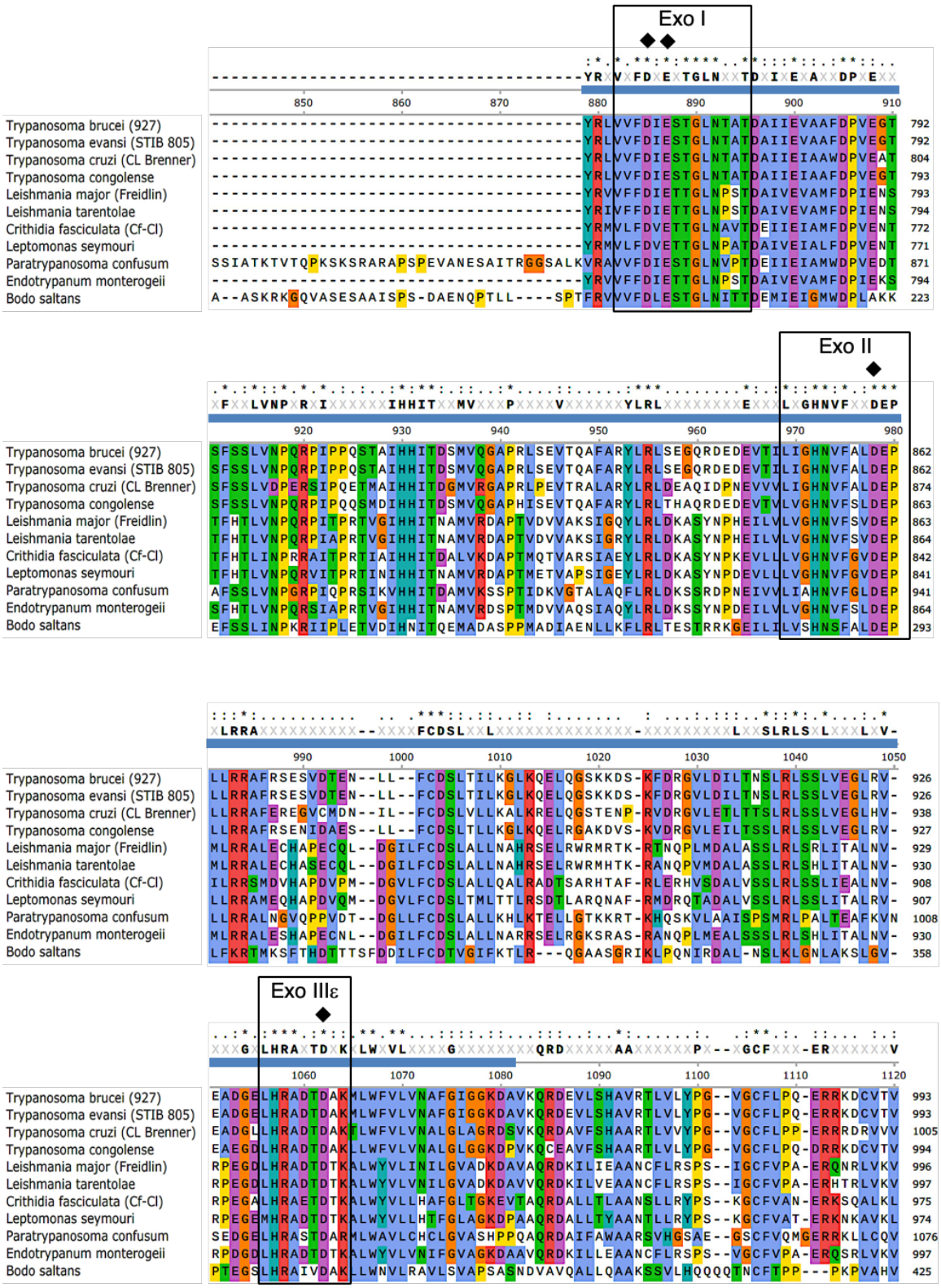


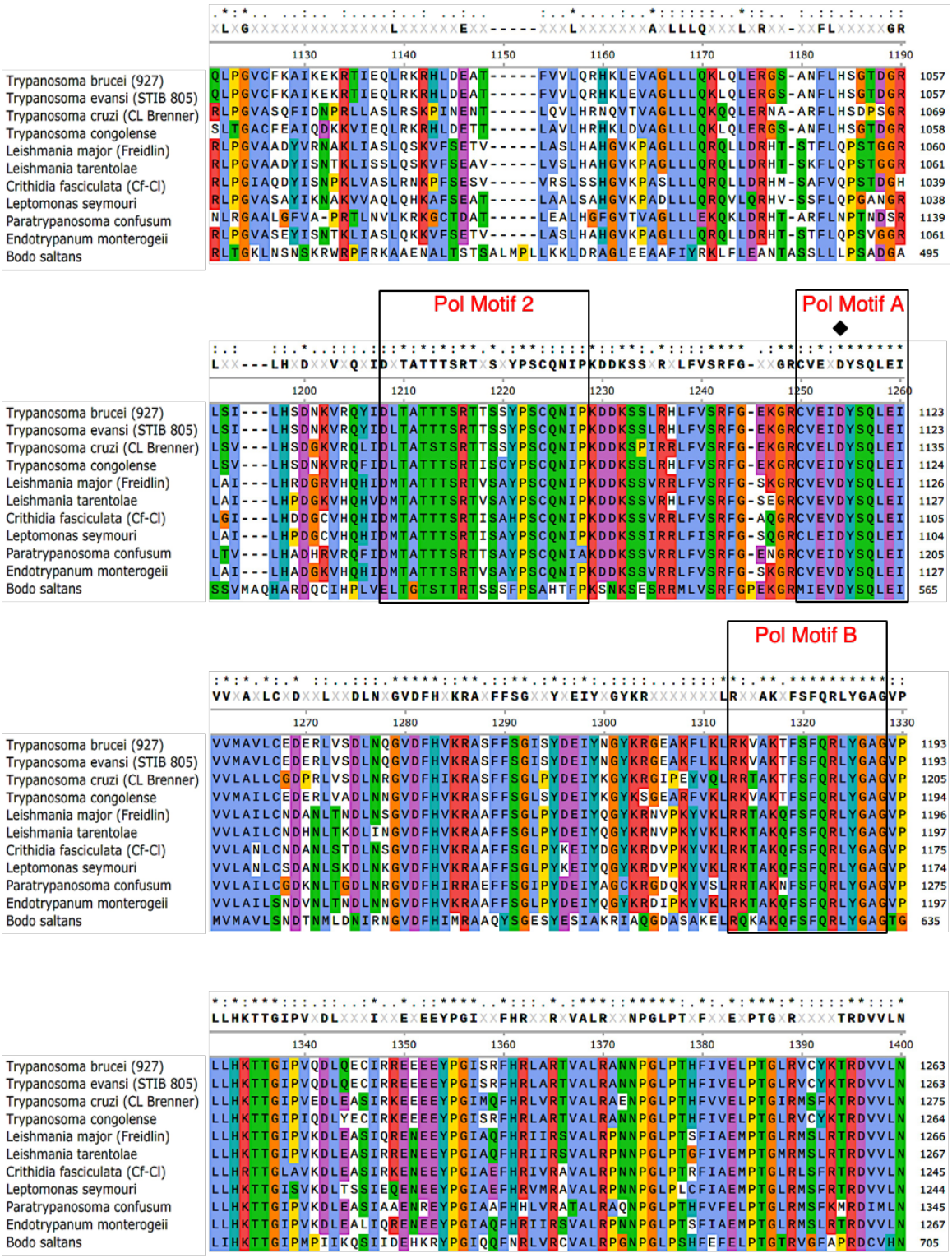
Figure AI.4: Sequence alignment of Family A Polymerase Domain. COBALT Multiple sequence alignment of the POLA domain from selected Family A DNA polymerases. Essential catalytic residues are indicated in motifs I and III (red circles). Motifs A, B and C are indicated with a black line.


Trypanosoma brucei (927)
Trypanosoma evansi (STIB 805)
Trypanosoma cruzi (CL Brenner)
Trypanosoma congolense
Leishmania major (Freidlin)
Leishmania tarentolae
Crithidia fasciculata (Cf-Cl)
Leptomonas seymouri
Paratrypanosoma confusum
Endotrypanum monterogeii
Bodo saltans


Trypanosoma brucei (927)
Trypanosoma evansi (STIB 805)
Trypanosoma cruzi (CL Brenner)
Trypanosoma congolense
Leishmania major (Freidlin)
Leishmania tarentolae
Crithidia fasciculata (Cf-Cl)
Leptomonas seymouri
Paratrypanosoma confusum
Endotrypanum monterogeii
Bodo saltans


Trypanosoma brucei (927)
Trypanosoma evansi (STIB 805)
Trypanosoma cruzi (CL Brenner)
Trypanosoma congolense
Leishmania major (Freidlin)
Leishmania tarentolae
Crithidia fasciculata (Cf-Cl)
Leptomonas seymouri
Paratrypanosoma confusum
Endotrypanum monterogeii
Bodo saltans


Exo I
Exo II
Exo IIIε
Trypanosoma brucei (927)
Trypanosoma evansi (STIB 805)
Trypanosoma cruzi (CL Brenner)
Trypanosoma congolense
Leishmania major (Freidlin)
Leishmania tarentolae
Crithidia fasciculata (Cf-Cl)
Leptomonas seymouri
Paratrypanosoma confusum
Endotrypanum monterogeii
Bodo saltans


Pol Motif 2
Pol Motif A
Pol Motif B
Trypanosoma brucei (927)
Trypanosoma evansi (STIB 805)
Trypanosoma cruzi (CL Brenner)
Trypanosoma congolense
Leishmania major (Freidlin)
Leishmania tarentolae
Crithidia fasciculata (Cf-Cl)
Leptomonas seymouri
Paratrypanosoma confusum
Endotrypanum monterogeii
Bodo saltans

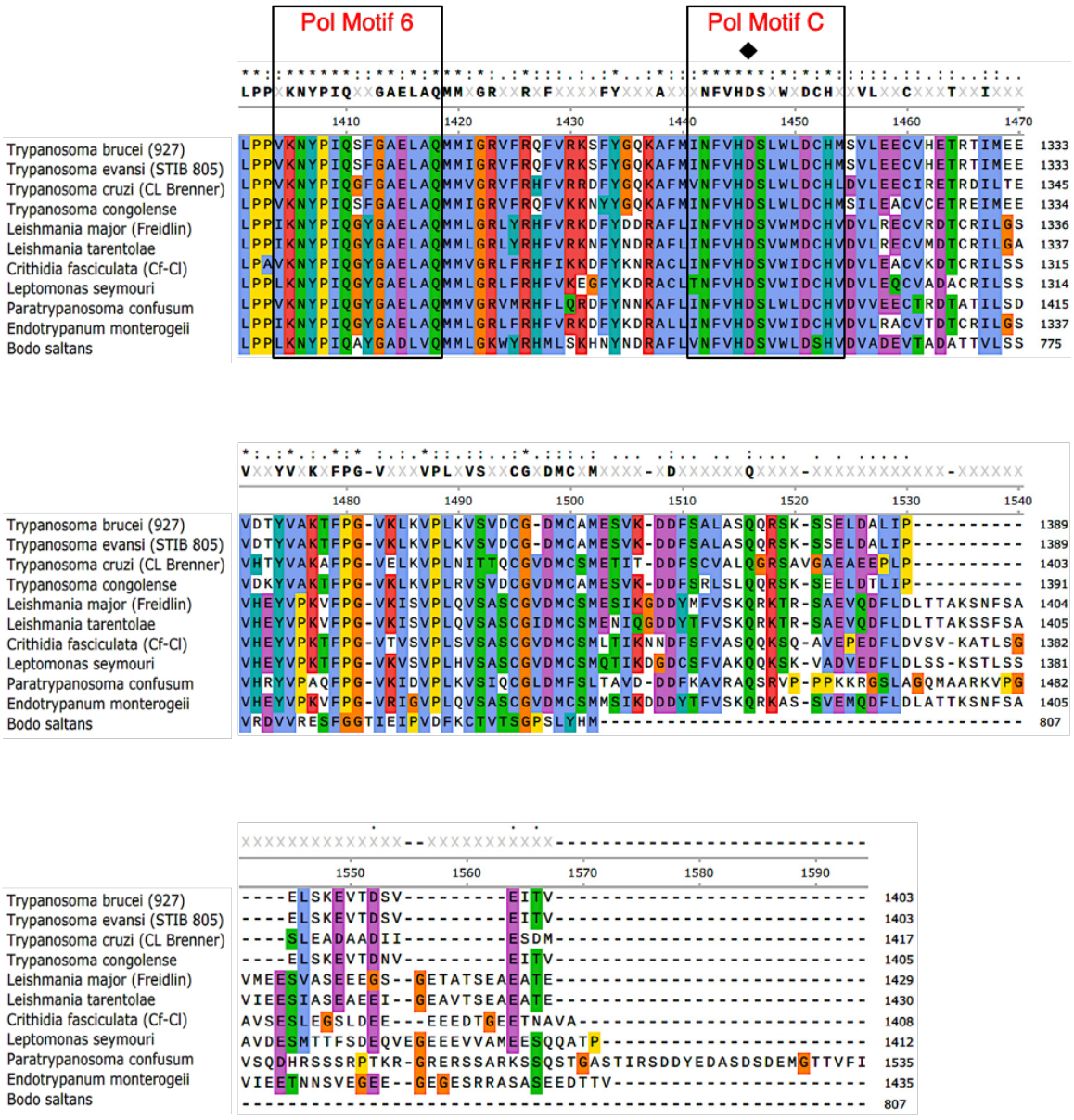


Figure AI.5: Alignment of Kinetoplastid POLIB sequences. The aligned POLIB sequences are: *Trypanosoma brucei brucei, Trypanosoma evansi*, *Trypanosoma cruzi*, *Trypanosoma congolense*, *Leishmania major, Leishmania tarentolae*, *Crithidia fasciculata*, *Leptomonas seymouri*, *Paratrypanosoma confusum*, *Endotrypanum monterogeli* and *Bodo saltans*. Coloring is based on Clustal X formatting. Concensus sequence was generated with a threshold >75%. Highly conserved and invariant residues are small circles and asterisks, respectively. Diamonds indicate the catalytic aspartic or glutamic acid residue. Black squares indicate five of the six most conserved regions in family A DNA polymerases. Blue bar indicates the thumb insertion with sequence similarity to DEDDh 3'-5' exonucleases.

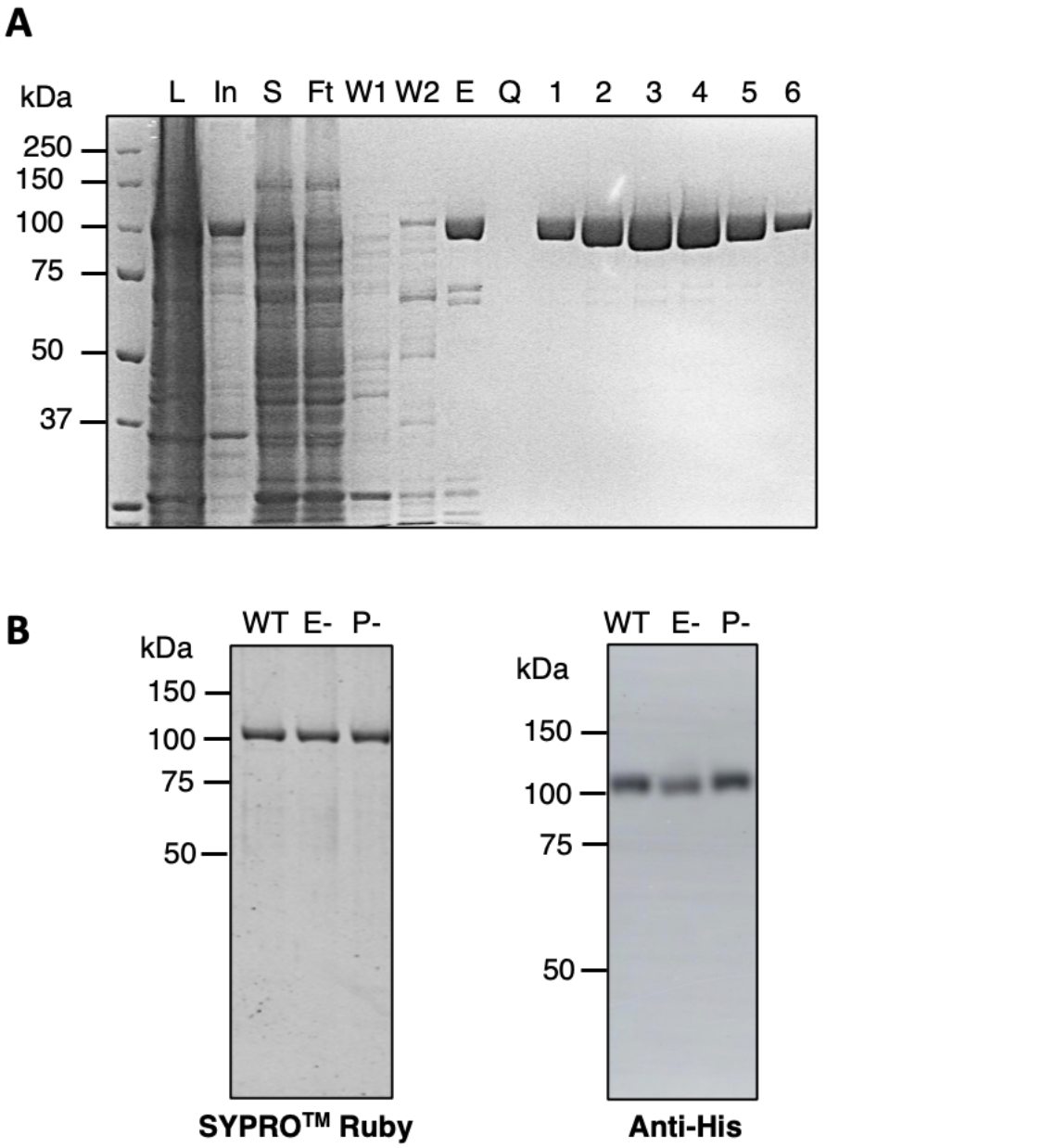


Figure AI.6: Preparation of recombinant TbPOLIB variants. (A) Coomasie stained 8% SDS-PAGE gel of IBWT purification progression. L; cell lysate, In; insoluble fraction, S; soluble fraction, Ft; flowthrough of the Ni column, W1; wash with lysis buffer, W2; wash with wash buffer, E; eluate from Nickel column, Q; flowthrough from the Q column, 1-6; selected fractions from Q column NaCl gradient elution. (B) 500 ng each POLIB variant was analyzed on 8% SDS-PAGE gel and either stained with Sypro ruby (left) or transferred to membrane and detected with anti-His antibody (right). WT, wild type variant; E-, exonuclease deficient variant; P-, polymerase deficient variant.

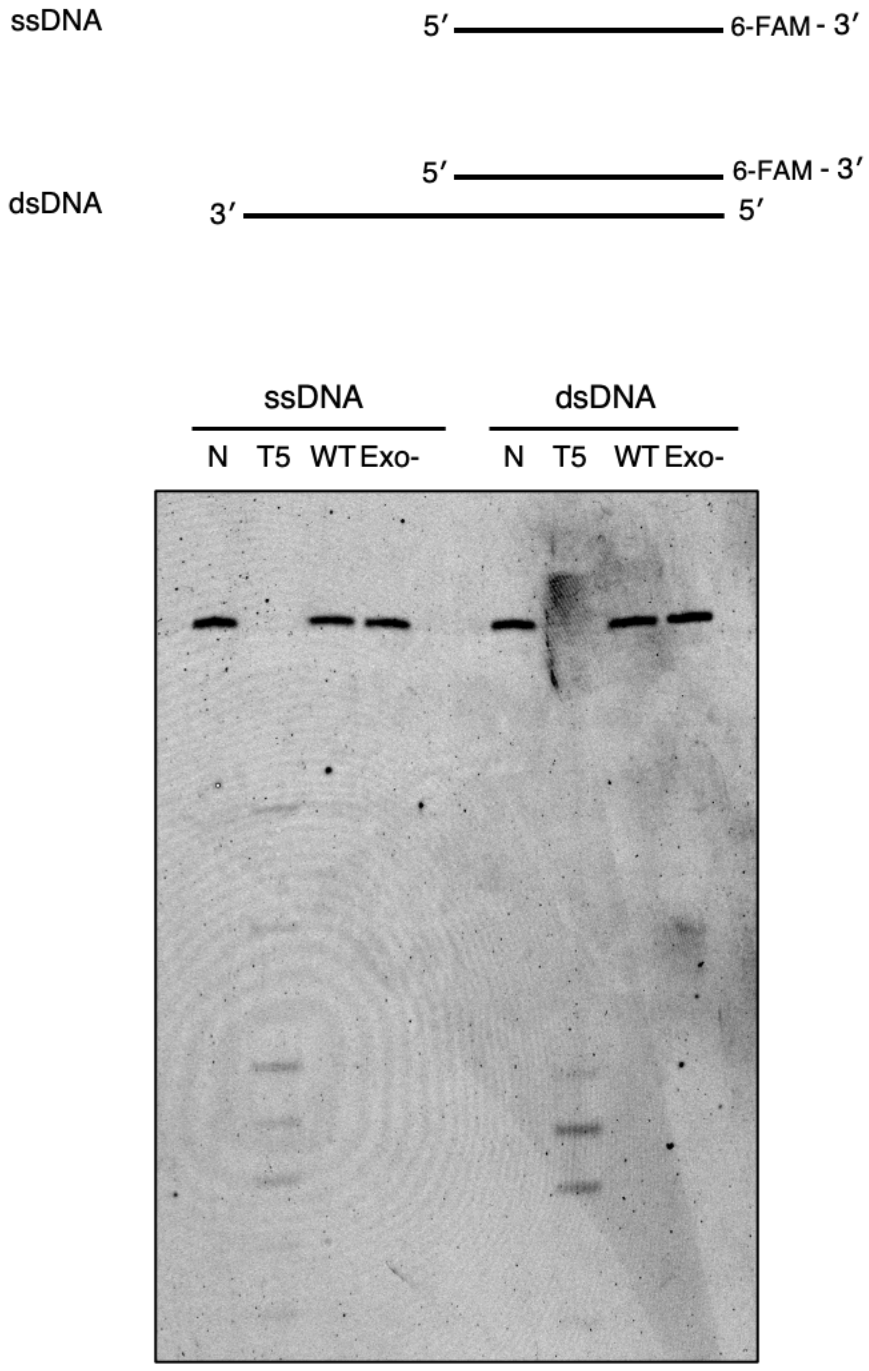


Figure AI.7: POLIB Lacks 5′-3′ Activity on DNA substrates. 200 nM POLIB variant was used to initiate each reaction using standard conditions (pH 7.0, 5 mM $MgCl_2$). T5 exonuclease (10U), positive control. All reactions were quenched after 30 minutes. Negative controls contain no enzyme and were quenched at time 0. dsDNA, 3′ FAM labeled 22mer annealed to the unlabeled 42mer template; ssDNA, 3′ FAM labeled 22mer. WT, wild type variant; Exo-, exonuclease deficient variant.

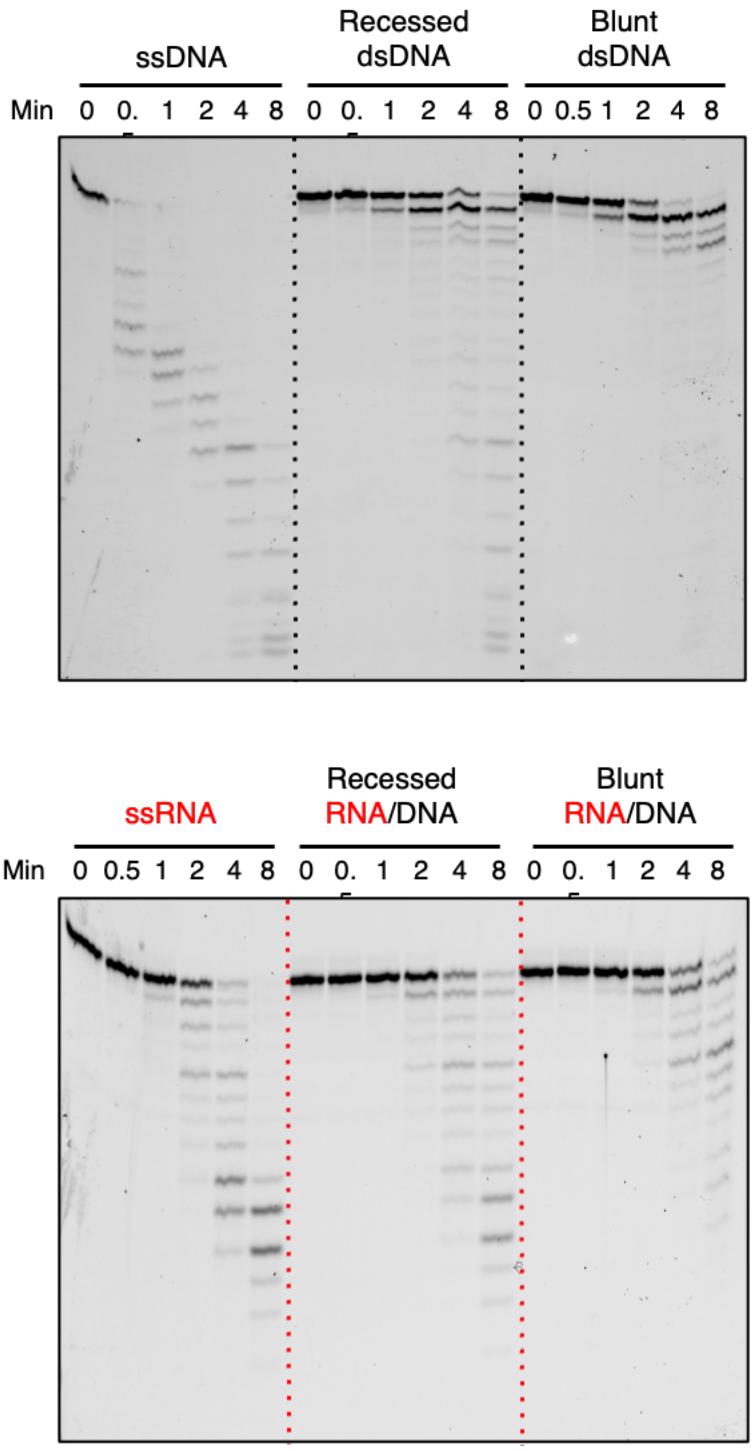


Figure AI.8: Representative images used for quantification of data in Figure 3.

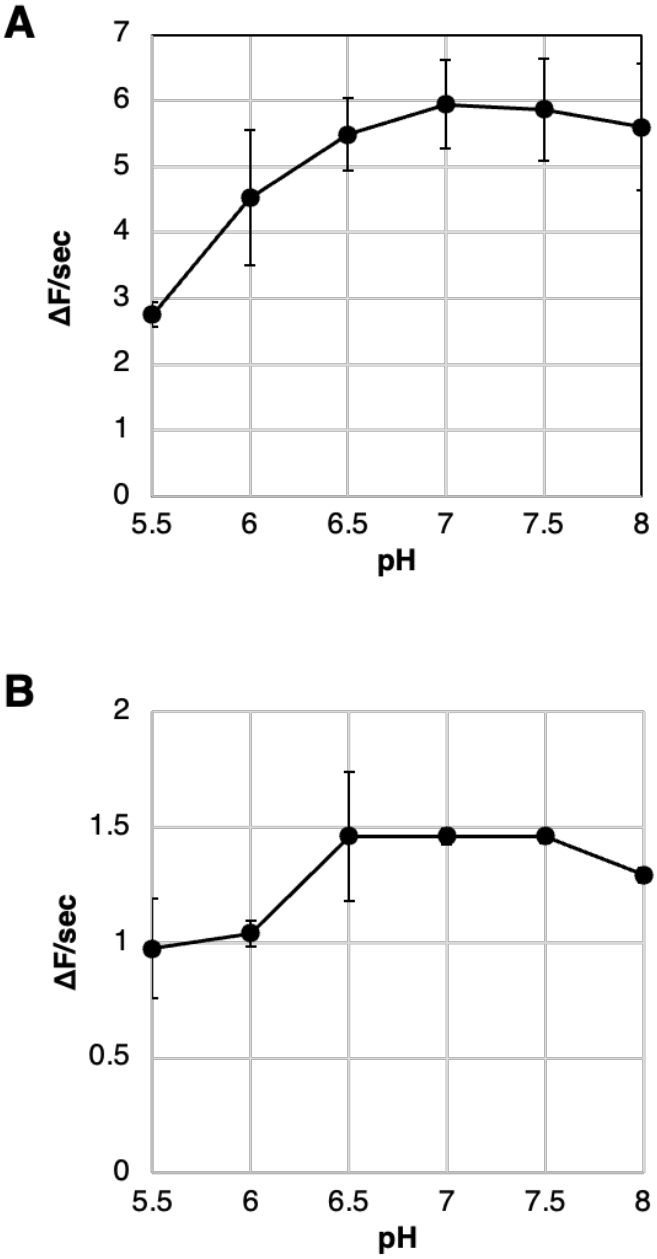


Figure AI.9: Rate of POLIB exonuclease activity on A) ssDNA and B) blunt dsDNA. Each 2AP excision assay contained standard buffer conditions, 200 nM IBWT and 5 mM 2AP position 1. Only the pH varied as indicated. Error bars represent standard deviation of three replicates, some error bars are too small to be visible on this chart.

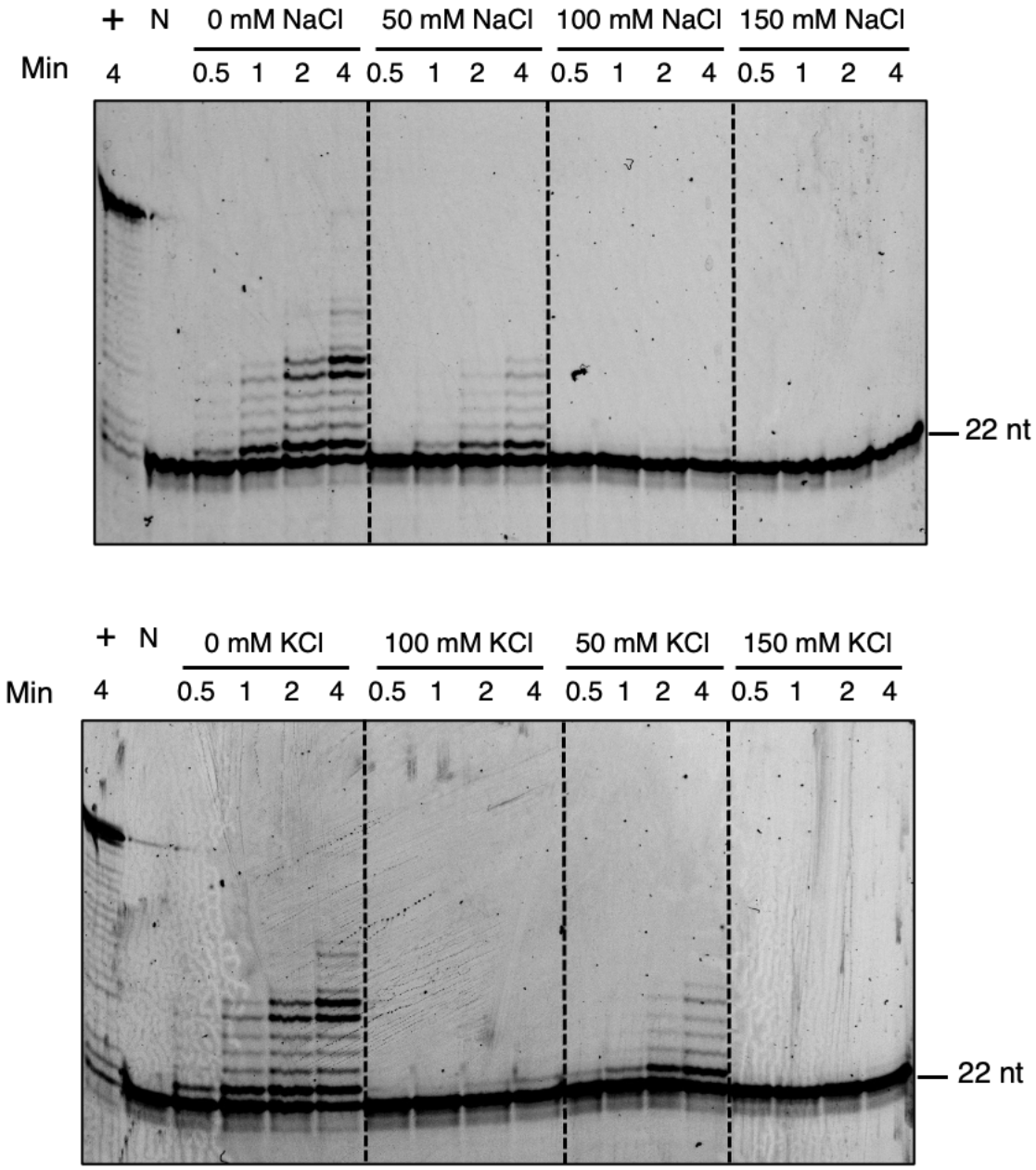


Figure AI.10: Salt optimum of primer extension activity for IBexo-. Reactions contained standard rection buffer (pH 7.0), varying salt concentrations as indicated and was initiated with 200 nM IBExo-. Positive control (+), 1U Klenow pH 8.0, no salt in buffer. Negative control (N) quenched at time 0.

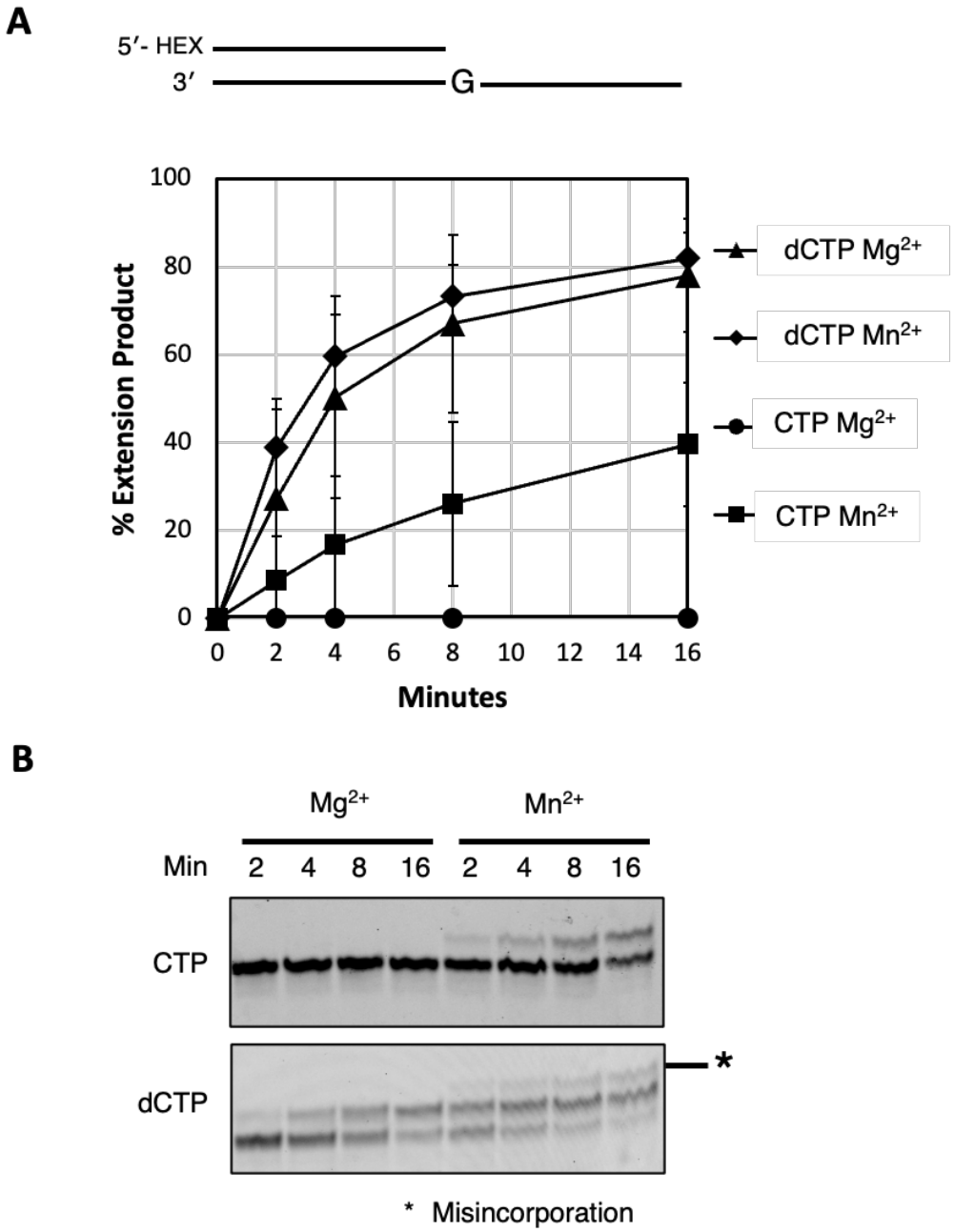


Figure AI.11: POLIB incorporates nucleotides more rapidly with Mn^{2+} than Mg^{2+}. Reactions include 200 nM IBExo- 5 mM CTP or dCTP, 50 mM Tris pH 7.0, 1 mM DTT, 0.1% BSA, 8.4 μM DNA, reactions started with 500 M $MgCl_2$ or 5 mM $MnCl_2$ A) quantification of % primer extended across three replicates with different POLIB preparations. B) Representative of primer extension reactions.

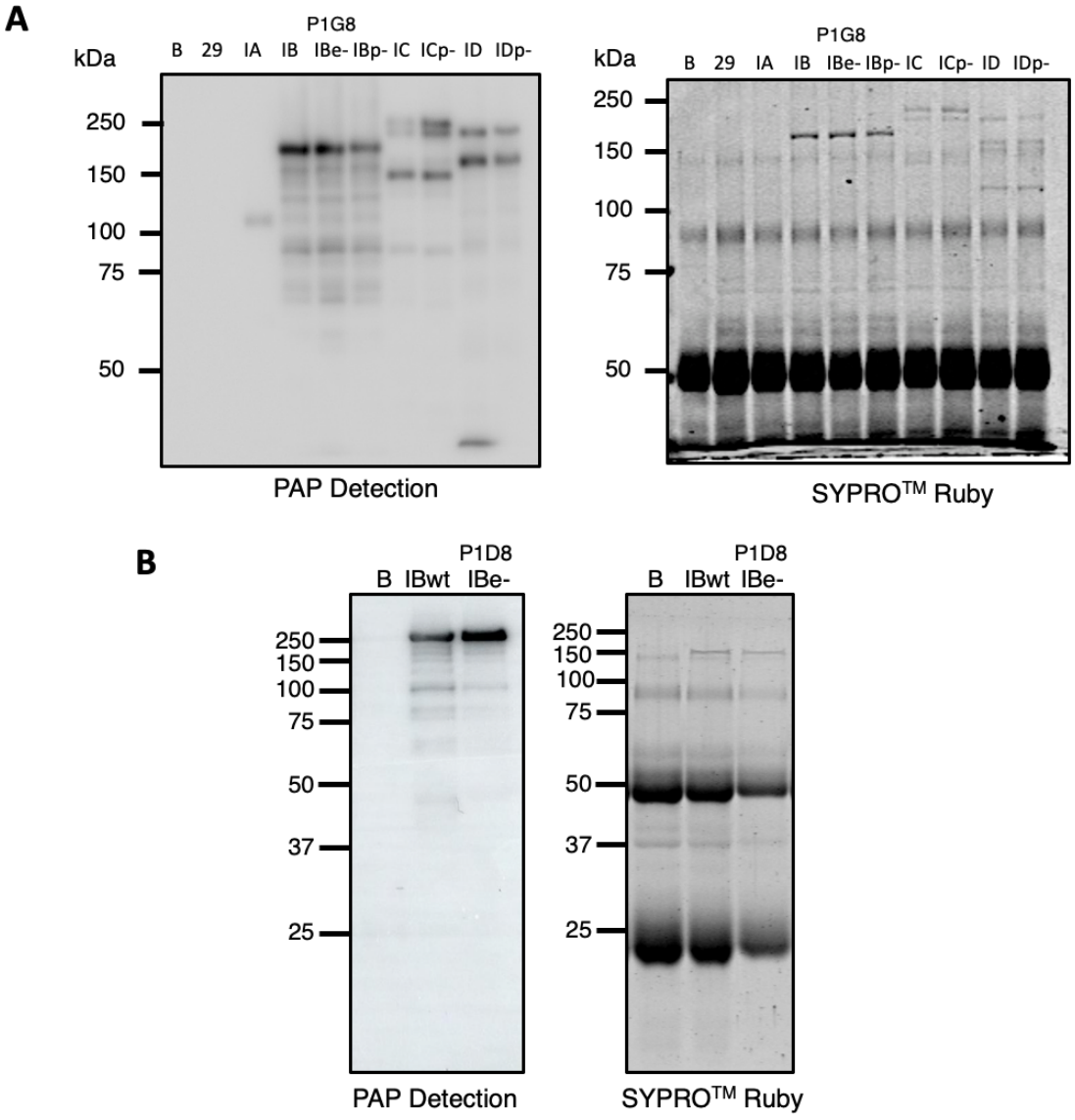


Figure AI.12: SDS-PAGE and Western Blots of PTP-tagged Pol I-like paralog variants immunoprecipitated from *T. brucei* procyclic cells. (A) Detection of Pol I-like variants on IgG Sepharose beads used in primer extension assay Figure 7A. PAP detection or SYPRO™ Ruby for total protein. B; Beads-only, 29; IP with 29-13 procyclic cells, IA; IA^{MHTAP}, IB; IBwt^{PTP}, IBe-; IBexo-^{PTP} (clone P1G8), IBp-; IBpol-^{PTP}, IC; ICwt^{PTP}, ICp-; ICpol-^{PTP}, ID; IDwt^{PTP}, IDp-; IDpol^{PTP}. (B) Representative PAP and SYPRO™ Ruby detection for IP samples of IBwt^{PTP} and IBexo-^{PTP} (clone P1D8) used in Figure 7B. Labels as in A.

A. POLIB Codon-Optimized Nucleotide Sequence

AGCGAAGTTGAAAAATGGCTGCGTGTTCTGAGCCTGGATAAACGTGTTAGCTGTGAAAAAGCACAGATGCGCAAATACAAA
AAGTATCTGGTGATTGCATTTATGGCCACCGAATTTGCAGCATATACCCGTCCGGCAAATCCGCATGCACCGGCAAATTAT
GTTGTTGCAACCAGCATTATGGATTACCAGCAGCAGGTTGTTGAACCGTGGACCAAATATAGCAGCCGTGGTGAATTTAAA
CTGCCGAGCCTGGAAGGTTTTGATGTTATTGTTTGCCATGACGTGAAACACTTTGTTCTGCTGATTTGGGATGATCCGGAA
CTGCGTCGTTTTCTGAAACGTGGTGGTCGTGTTTGGTGTACCATGTTTGCAGAATATCTGCTGGATGCACAGCGTTGTCAG
AGCGGTAGCAATAGCCTGCATGATGTTGCAATGAAATATGGTATTCTGACACCGCAGAGCAGCGTTCTGGGTCTGAGCACA
CCGGATCTGCCGATTGCCTTTATTCAGCATTATCTGGTTGCAGCAGTTGATGCAATTAGCCGTGTTTTTCAAGAGCAGCTG
AAAAAAGCATGTGGTAATAGCCAGCTGATTTGTGTTGCACATCGTATGGATAGCCTGCTGGCAATGGCAAGCATTGAAAAA
GCCGGTATTCATATCGATAGCAAAGAAGCAACCCTGCAGGCACAGGCAATTCGTAATCGTCTGCTGGCCATTGATAAATCA
CTGAGCCTGTATGCACCGGATGAAATTCCGCTGGATATGCAGCGTTTTTTTGATTGGACCAGCCTGCAGCATCTGCAAGCA
TATTTCTTTGGTGGTAGCATTACCCTGGGTTATACCGATATTAGTCGTGATAGCAGCACCTGGACCGCACATCTGATTCAT
CTGTGTCATAAATATGGCAATCTGGGCCTGATGAGCGCAGATGTTCATCTGCAGCGCTTTGCAAGCGAACGTGGTCTGCGT
GGTGCAGGTCGTCTGCCGCAGCGTGTTGCACGCTTTTTTGATGCAGATGGTAGCAGTCGTCGTCGTAAATATCGTCTGGTT
GTGTTTGATATTGAAAGCACCGGTCTGAATACCGCAACCGATGCCATTATTGAAGTTGCAGCCTTTGATCCGGTTGAAGGC
ACCAGCTTTAGCAGCCTGGTTAATCCGCAGCGTCCGATTCCGCCTCAGAGCACCGCAATTCATCATATTACCGATAGCATG
GTTCAGGGTGCACCGCGTCTGAGCGAAGTTACCCAGGCATTTGCACGTTATCTGCGCCTGAGTGAAGGTCAGCGTGATGAA
GATGAAGTTACCATTCTGATTGGCCATAATGTTTTTGCACTGGATGAACCGCTGCTGCGTCGTGCATTTCGTAGCGAAAGC
GTTGATACCGAAAATCTGCTGTTTTGTGATAGCCTGACAATTCTGAAAGGCCTGAAACAAGAACTGCAGGGTAGCAAAAAA
GATAGCAAATTTGATCGTGGCGTGCTGGATATTCTGACCAATAGTCTGCGTCTGAGTAGCCTGGTTGAAGGTCTGCGCGTT
GAAGCCGATGGTGAACTGCATCGTGCAGATACCGATGCAAAAATGCTGTGGTTTGTTCTGGTTAATGCCTTTGGTATTGGT
GGTAAAGATGCAGTTAAACAGCGTGACGAAGTGCTGAGCCATGCAGTTCGTACCCTGGTTCTGTATCCTGGTGTTGGTTGT
TTTCTGCCGCAAGAACGTCGTAAAGATTGTGTTACCGTTCAGCTGCCTGGTGTTTGTTTTAAAGCCATTAAAGAAAAACGC
ACCATTGAGCAGCTGCGTAAACGTCATCTGGATGAAGCAACCTTTGTTGTTCTGCAGCGTCATAAACTGGAAGTTGCCGGT
CTGCTGCTGCAGAAACTGCAGCTGGAACGTGGTAGCGCAAATTTTCTGCATAGCGGCACCGATGGTCGTCTGAGCATTCTG
CATTCAGATAATAAAGTGCGCCAGTATATTGATCTGACCGCAACCACCACCAGTCGTACCACCAGTAGCTATCCGAGCTGC
CAGAATATTCCGAAAGATGATAAAAGCAGCCTGCGTCACCTGTTTGTTAGCCGTTTTGGTGAAAAAGGTCGTTGCGTGGAA
ATTGATTATTCCCAGCTGGAAATTGTTGTTATGGCAGTTCTGTGCGAAGATGAACGTCTGGTTAGCGATCTGAATCAGGGT
GTTGATTTTCATGTTAAACGTGCCAGCTTTTTTAGCGGCATTAGCTATGATGAAATCTATAACGGTTATAAACGCGGTGAA
GCCAAATTCCTGAAACTGCGTAAAGTTGCAAAAACCTTTAGCTTTCAGCGTCTGTATGGTGCCGGTGTTCCTCTGCTGCAT
AAAACCACCGGTATTCCGGTTCAGGATCTGCAAGAATGTATTCGTCGTGAAGAAGAAGAATATCCGGGTATTTCACGTTTT
CATCGTCTGGCACGTACCGTTGCACTGCGTGCAAATAATCCGGGTCTGCCGACACATTTTATTGTTGAACTGCCGACAGGC
CTGCGTGTTTGCTATAAAACCCGTGATGTTGTGCTGAATCTGCCTCCGGTTAAAAACTATCCGATTCAGAGCTTTGGTGCA
GAACTGGCACAGATGATGATTGGTCGTGTGTTTCGTCAGTTTGTGCGTAAAAGTTTCTATGGCCAGAAAGCCTTTATGATC
AACTTTGTGCATGATAGTCTGTGGCTGGATTGTCATATGAGTGTTCTGGAAGAATGTGTTCATGAAACCCGTACCATTATG
GAAGAGGTTGATACCTATGTGGCCAAAACCTTTCCGGGTGTTAAACTGAAAGTGCCGCTGAAAGTTAGCGTTGATTGTGGT
GTTGATATGTGTGCAATGGAAAGCGTGAAAGATGACTTTAGCGCACTGGCAAGCCAGCAGCGTAGCAAATCAAGCGAACTG
GATGCCCTGATTCCTGAACTGTCAAAAGAAGTGACCGATAGCGTTTAA

B. Oligonucleotides used for Site Directed Mutagenesis

Mutagenic site	Sequence (5'-3')
Pol Domain D1117A	ATTTCCAGCTGGGAATAA**G**CAATTTCCACGCAACGAC
Pol Domain D1309A	CAATCCAGCCACAGACTA**G**CATGCACAAAGTTGATCA
Exo Domain D767N, E769Q	CAGACCGGTGCTTT**G**AATAT**T**AAACACAACCAGACGATATTTACGA

Table AI.1: Nucleotide reagents used for POLIB recombinant constructs. A. recoded region corresponding to POLIB nucleotide optimized for E. coli expression. B. Primers for site-directed mutagenesis. Mutagenic sites are in red bold.

A. Epitope tagged trypanosome cell lines for this study

Cell Line		Mutation
POLIA^{MHTAP}	(IAwt)	none
POLIB^{PTP}	(IBwt)	none
POLIBExo-^{PTP}	(IBe-)	D767N, E769Q
POLIBPol-^{PTP}	(IBp-)	D1117A, D1309A
POLIC^{PTP}	(ICwt)	none
POLICPol-^{PTP}	(ICp-)	D1380A, D1592A
POLID^{PTP}	(IDwt)	none
POLIDPol-^{PTP}	(IDp-)	D1346A, D1565A

B. Oligonucleotides used for generating trypanosomes cell lines

Name	Sequence (5'-3')	Purpose
POLIA F	TTATTAGGATCCATGTCGTTACATGGTCATC	Subcloning
POLIA R	TTATTAGGATCCACTCGGCACCCGCAGTC	
POLIB F	GTAACTCGAGATGCGGCTAAATAGCTGC	
POLIB R	ACACTCTAGACACCGTAATTTCTACACTGTCAG	
POLID F	TTATTACAATTGATGCTGCGGCGGCTCTTC	
POLID R	TTATTAGGATCCCAGTGTCTCCTCAATGACAACGG	
POLIB Rc F	gcttcaattgctcgaGAGATGCGGCTAAATAGCTGC	Gibson
POLIB Rc R	gtaccgggccggatcCGACTGATATACCCCACGATAC	
POLIB M1	GGCCGTTGCGTGGAAATTGC**C**TACTCGCAGCTG	Mutagenesis
POLIB M2	GATAAACTTCGTACACG**C**TTCGCTTTGGCTCGAC	
POLIB M3	GCGCAAGTACAGACTAGTTGTATTT**A**ATATT**C**AATCCACGGGATTG	
POLID M1	GGGAATGTGTATTGAGGCAG**C**TTATTCACAGCTTGAAGTCG	
POLID M2	ACTAATCAACACCGTGCACG**C**CTGCGTATGGATTGACTG	
POLID M3	ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGG	
POLID M4	CCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGG	

Table AI.2: Generation of epitope tagged *T. brucei* Pol I-like paralogs for overexpression and immunoprecipitation. Underline, restriction enzyme sites, small case, vector backbone sequence; Red bold, mutagenic sites.

APPENDIX II - Overexpression of POLIB Variants in Procyclic Form *T. brucei*

Contributions to cloning and cell line development made by Matthew P. Frost

When POLIB is knocked down in procyclic or bloodstream form *T. brucei,* the cells lose fitness, their kDNA networks reduce in size, and there is a loss of minicircle replication intermediates [1]. This phenotype indicated a role of POLIB in kDNA replication, most likely minicircle replication [1]. We hypothesized that the highly divergent predicted structure of POLIB is essential for its specialized role in kDNA replication. In chapter 2, we demonstrated that in basic conditions wildtype POLIB exo activity outcompetes pol activity on DNA substrates. We expect that the mitochondrion of *T. brucei* is a basic environment, as in other eukaryotic systems [1]. These data suggested the possibility that the exonuclease activity of POLIB may be the essential contribution of POLIB to kDNA replication, and it may be possible that a variant of POLIB lacking pol activity could fully rescue POLIB knockdown.

In order to test this, we designed a complementation system to determine if the loss of fitness (LOF) and minicircle replication defects seen during POLIB depletion via RNAi could be rescued by a variant of POLIB with pol and/or exo activity ablated These experiments would tell us whether both enzymatic activities of POLIB are critical for its function in kDNA replication, or if perhaps part of the protein is vestigial. The complementation system in which epitope-tagged POLIB variants (described in chapter 2) would be overexpressed in the POLIB RNAi parental cell line.

The complementation system is a dual-function TET-ON system that can be used to simultaneously induce dsRNA to deplete mRNA of a specific gene and co-express a copy of the gene, that is resistance to the knowckdown, ectopically. This dual-function TET-ON was recently used to successfully determine a catalytic and noncatalytic role for POLIC, another kDNA polymerase required for kDNA maintenance [2]. We aimed to embark on a similar study to better understand the role of POLIB. Two critical controls for complementation include resuce with a copy of wildtype protein to observe any off-target effects and overexpression controls in 29-13 cells. Unexpectedly, the overexpression of POLIB wildtype simultaneous with POLIB RNAi did not fully rescue the loss of fitness phenotype characteristic of SLIB induction, likely due to not enough exogenous protein being expressed to replace endogenous levels [3].

One of the critical controls for RNAi complementation is to overexpress the POLIB variants without simultaneous POLIB knockdown in order to detect any dominant-negative phenotypes. These POLIB overexpression cell lines developed are a valuable tool for studying these variants. We confirmed that POLIBwt retained both 3'-5' exo activity and nucleotidyl incorporation activity by using immunoprecipitated protein in gel-based activity assays, discussed in Chapter 2 (Fig 2.8). POLIBexodead was shown to lack exo activity but retained the ability to perform nucleotidyl incorporation activity, especially from an RNA primer. Observing enzymatic properties of immunoprecipated protein variants gives us an idea of the potential roles of POLIB, but it is critical to follow up with *in vivo* studies to fully understand the role of POLIB. loss of exo activity of the mutant

Methods

Cells lines were transfected as described in chapter 2 (page 42) Eight clones of POLIBwt-PTP OE were selected for screening and seven clones of POLIBexodead-PTP OE were selected for screening. Clone were chosen for subsequent experiments based on protein expression after a 2-day induction with 1 μg/mL tetracycline. For POLIBwt-PTP OE clone P2B5 was chosen for further experiments based on protein expression, and for POLIBexodead-PTP OE P1G8 was chosen. Cells were cultured in SDM-79 medium supplemented with 15% FBS, G418 (15 μg/ml), hygromycin, and puromycin (1 μg/ml). Cultures were maintained at $5x10^5$-$1x10^6$ cells/mL, and induced with 1 μg/mL tetracycline that was added each time the cells were diluted. On days when the cells were not diluted, cultures were supplemented with an additional 0.5 μg/mL tetracycline. Counts were taken approximately every 24 hours for growth curves via Beckman coulter counter. Western blot detection was performed as previously described using Sigma peroxidase anti peroxidase (PAP) and anti-EF1α as a loading control

Results

Fitness of POLIB Overexpression Cell lines

Overexpression of POLIBwt appeared to have no impact on fitness of procyclic cells, and POLIBwt protein levels appeared consistent throughout the induction as measured by Western blot. This provided confirmation that the recoded region cloned into this POLIB construct (described in chapter 2 methods) to evade RNAi did not cause a toxic effect on the cells (Figure AII.1). Overexpression of POLIBexo- has a dominant negative

phenotype on the fitness of these cells and had POLIB-PTP expression levels ~15-fold higher than the single expressor cell line (Figure AII.2). This result demonstrates that the presence of endogenous wildtype POLIB does not compensate for the overexpression of this catalytically dead variant.

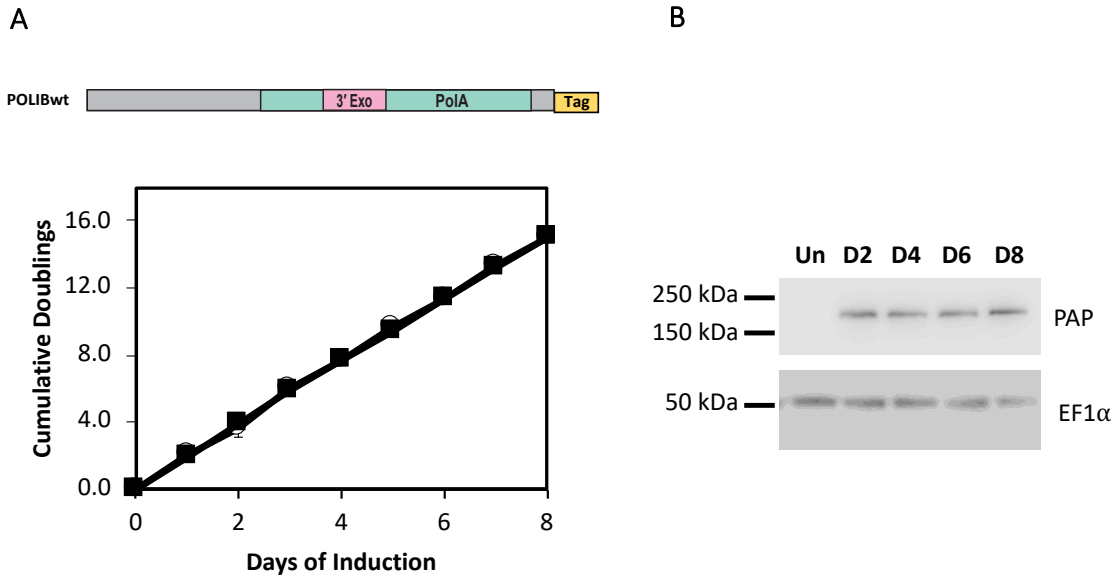


Figure AII.1: Induction of Overexpression of POLIBwt-PTP. A)Growth curves of POLIBwtcell lines induced (black squares) or uninduced (open circles) for POLIBexodead overexpression. N=3 B) Western blot showing POLIBwt overexpression, $1x10^6$cell equivalence loaded for each time point. PAP was used to detect PTP-tagged protein, EF1α detected as a loading control.

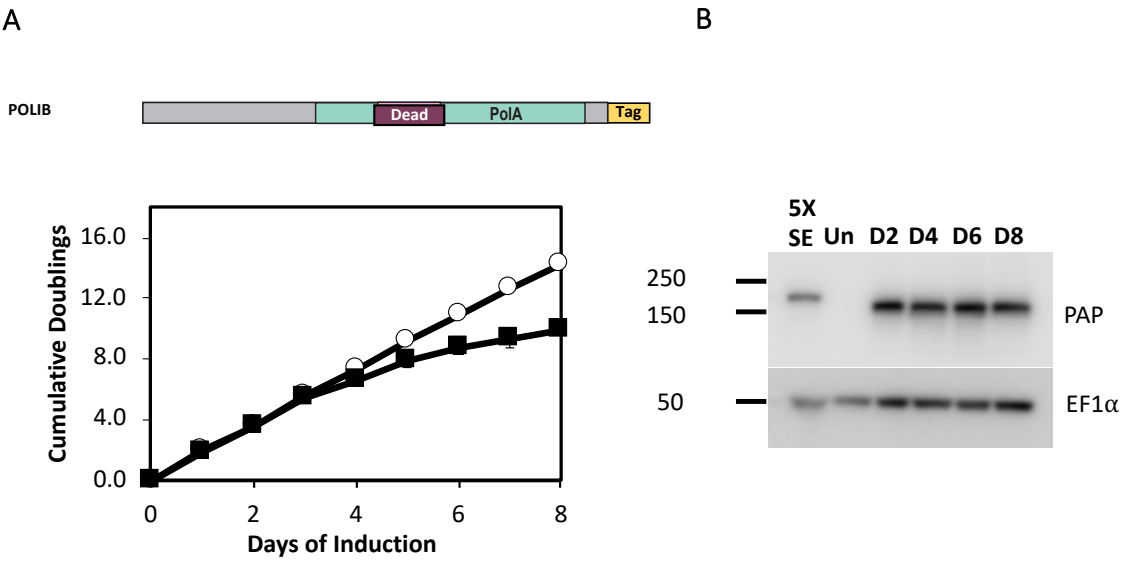


Figure AII.2: Induction of Overexpression of POLIBexodead-PTP. A) Growth curves of POLIBexodead cell lines induced (black squares) or uninduced (open circles) for POLIBexodead overexpression N=3. B) Western blot showing POLIBexodead overexpression, $1x10^6$cell equivalence loaded for each time point. PAP was used to detect PTP-tagged protein, EF1α detected as a loading control. 5X SE= $5x10^6$ cell equivalence of POLIB-PTP single expressor as a control for protein expression levels

Impact of POLIBexodead Overexpression on Minicircle Replication Intermediates

In order to understand how the POLIBexodead variant may be conferring the dominant-negative fitness phenotype, we tested the impact of this variant on free minicircle species (Fig AII.3A). These molecules can be detected using a minicircle-specific probe after separation via agarose gel electrophoresis and transfer to a membrane. At day two of the induction, covalently closed (CC) minicircles transiently increase in numbers, followed by a gradual decrease through the rest of the 8-day induction. Meanwhile, nicked-gapped (N/G) species increase at days four and six, before dramatically decreasing on day 8 (Fig AII.3B). These data provide evidence that the dominant-negative phenotype of

POLIBexodead overexpression is a result of this mutant is somehow impacting minicircle replication.

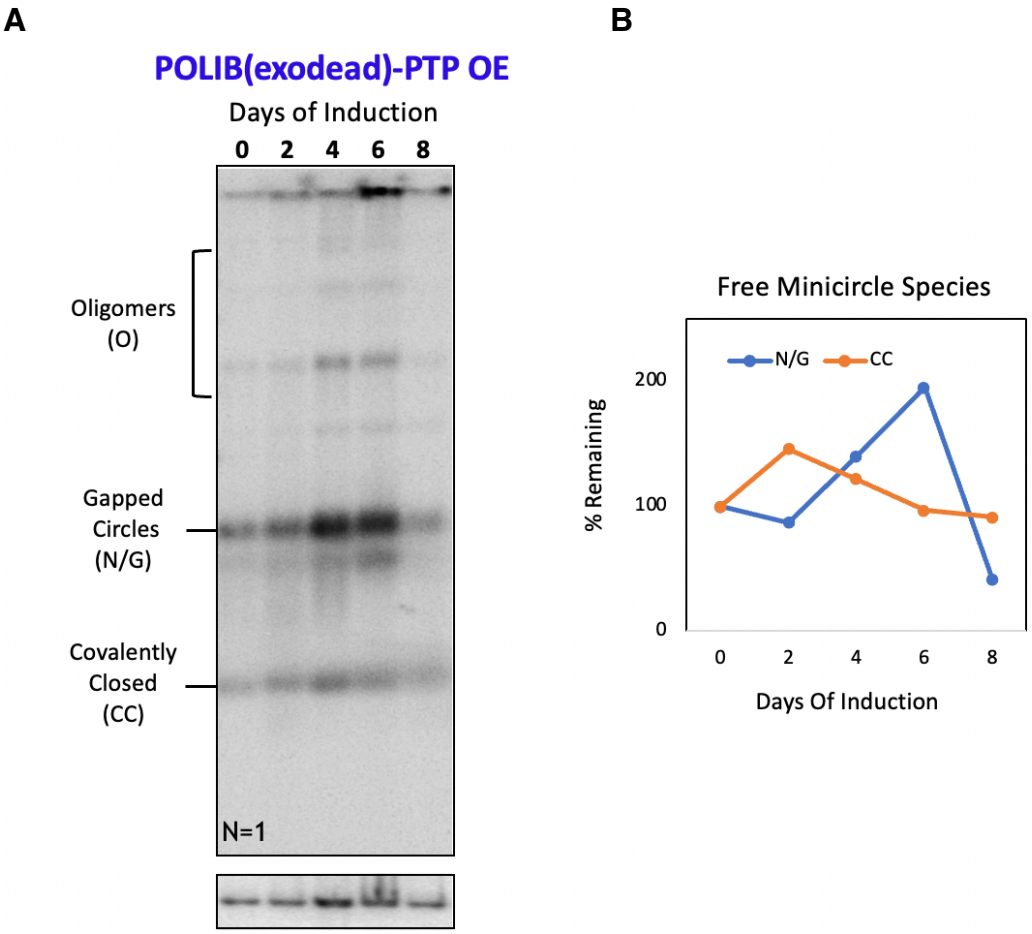


Figure AII.3) Southern Blot Analysis of Free Minicircle species through Induction of Overexpression of POLIBexodead-PTP. A) Southern blot detection of minicircle species using DNA isolated from trypanosomes at days 0, 2, 4, 6, and 8 days of induction. B) Quantitation of nick-gapped and covalently closed free minicircle species from Southern blot. N=1.

Discussion

The dominant negative fitness phenotype associated with the POLIBexodead variant could be an indicator that the exonuclease activity of POLIB is critical for its role in kDNA replication. The presence of much more POLIBexodead than the endogenous wildtype protein could lead to the variant diluting the endogenous protein so it cannot effectively perform its essential role. Alternatively, POLIBexodead may localize or bind DNA substrates inappropriately and block other essential kDNA replication enzymes from carrying out their roles. The Southern blot analysis supports this hypothesis (Fig 3). In knockdown of POLIB, both nick/gapped and covalently closed minicircles reduced in quantity over the course of an 8-day induction (Bruhn 2010). However, in POLIBexodead overexpression, while ultimately both CC and N/G species are reduced in number, we see a transient increase in N/G minicircle species that does not phenocopy SLIB induction. The transient increase of N/G species more closely resembles the phenotype seen when another kDNA pol, POLID, is knocked down (Klingbeil 2002). Together these data support the hypothesis that POLIBexodead may be interfering with another component of the kDNA replisome, such as POLID, although more direct evidence is needed.

In future studies, microscopy experiments to determine POLIBexodead localization compared to that of POLIBwt may also reveal more about this mechanism, for example, whether the POLIBexodead variant is incapable of localizing as the wildtype enzyme. Repeated Southern blot experiments should be performed to provide more robust data on how POLIBexodead impacts minicircle species. If POLIB exo activity is important for

the degradation of linearized kDNA molecules, Southern blot analysis may reveal a buildup of these molecules, similar to what was observed in POLIB knockdown experiments (Bruhn 2010). Development of complementation cell lines will be crucial to obtaining a higher resolution understanding of the roles of the two enzymatic functions of POLIB.

APPENDIX III - Additional Variants of Recombinant POLIB

In order to test the enzymatic properties of some of the divergent structural features of POLIB, we have designed and completed cloning on a suite of mutations in POLIB. As previously described in chapter 2, POLIB has a divergent amino acid sequence in the pol domain motif C (amino acid 1310) that is most commonly a conserved aspartic acid in Family A pols, but in POLIB is a serine. In POLIC and POLID, the residue in this motif is a cystine. Some RNA polymerases have a serine at this site, so to investigate if this residue confers the ability of POLIB to extend more rapidly from an RNA primer, we used site-directed mutagenesis to make variants of POLIB to make either a Ser→Asp mutation modeled after other DNA pols, or a Ser→Cys mutation modeled after POLIC and POLID pol domains. In order to test the activity of these mutated pol domains without competition from the exonuclease domain, we made two additional variants that contained these described pol domain mutations along with the IBexo- active site mutations that ablated exonuclease activity.

The exonuclease domain of POLIB is a fascinating structural feature of a polymerase. We hypothesized based on biochemical characterization of IBwt that the position of the exo

domain may influence activity of both the pol and exo domains. Because the exo domain projects out of the pol thumb subdomain in the predicted protein structure, one mechanism could be structural interference with the thumb domain's participation in extension activity. Another mechanism could be the competition between the two domains to bind to similar DNA substrates. To understand these potential influences, we designed truncated POLIB variants that delete one domain or the other. In the pol domain construct, the two linker regions (Figure 2.1A) were included while the exo domain was deleted. To have the best chance of getting an exo domain construct that would properly fold, we designed two variants, one that included the linker regions, and one that included exclusively the exonuclease-homologous region of the protein.

Methods

Primer Name	Primer Sequence (5'-3')	POLIB Construct
ForwardExoOnly/ UM107	GCGCGCTAGCGGTGGTTCTGGTGGTGGTTCTGGTCGTCTGGTTGTGTTTGATATTGAAAGCACC	IB_Exodomain
ReverseExoOnly/ UM103	GCGCGAGCTCTTATTTACCACCAATACCAAAGGCATTAACCAG	IB_Exodomain
ForwardInsertOnly/ UM105	GCGCGCTAGCGGTGGTTCTGGTGGTGGTTCTGGTGCACCGGATGAAATTCCGCT	IB_Insert
ReverseInsertOnly/ UM104	CGCGGAGCTCCGCGTTACAGTTTATGACGCTGCAGAACAAAG	IB_Insert
HDS→HDE/ UM090	GAAAGCCTTTATGATCAACTTTGTGCATGATGAACTGTGGCTGGATTGTCA	IB_HDE
HDS→HDC/ UM089	CTTTATGATCAACTTTGTGCATGATTGTCTGTGGCTGGAT	IB_HDC
Exodead mutations/ UM039	CAGACCGGTGCTTTGAATATTAAACACAACCAGACGATATTTACGA	POLIB_HDEexo-POLIB_HDCexo-IB_doubledead
ForwardExoDeletion/ UM101	ATATTTACGACGACGACTG	POLIB_PolDomainOnly
ReverseExoDeletion/ UM102	GATGCAGTTAAACAGCGTG	POLIB_PolDomainOnly

Table AIII.1: Primers used for generation of additional variants of Recombinant POLIB

Generation of IB_exo and IB_insert

The DNA sequences corresponding to the exo domain (amino acids 762-955) and the thumb insertion (amino acids 661-1030) of recombinant POLIBtrunc were amplified from pET/dTOPO_His_IBwt$_{trunc}$ using the respective primers in Table 1. The resulting fragments were digested with Nhe1 and Sac1 (NEB). For the vector backbone pET/dTOPO_His_IBwt$_{trunc}$ was digested with nhe1 and sac1 and the vector was extracted from agarose gel after separation. Resulting vector was ligated with either the exo domain fragment or the insertion fragment using T4 ligase, generating pET/dTOPO_His_IBexodomain and pET/dTOPO_His_IBinsert. Final constructs were verified by Sanger sequencing.

Generation of IB_HDE, IB_HDEexo-, IB_HDC, IB_HDCexo-, and IBdoubledead

These polymerase domain catalytic site mutations were introduced to pET/dTOPO_His_IBwt$_{trunc}$ using the Quikchange Lighting Site-Directed Mutagenesis Kit (Agilent) according to manufacturer's instructions. For constructs with both pol and exo catalytic site mutations (IB_HDEexo-, IB_HDCexo-, and IBdoubledead) the construct with pol domain mutations was used as the template for site directed mutagenesis to add exo mutations. Mutations were introduced into these constructs using the respective primers in Table AIII.1. Final constructs were verified by Sanger sequencing.

Generation of IB_poldomain

The exonuclease domain (amino acids) was deleted from the recombinant POLIB open reading frame in pET/dTOPO_His_IBwt$_{trunc,}$ using the Q5 Site-Directed Mutagenesis Kit (NEB) according to manufacturer's instructions and primers in Table AIII.1. The final construct was verified by Sanger sequencing.

Pilot Inductions

Constructs were each transformed into BL21(DE3) chemically competent cells and plated onto LB agar containing 100 μg/mL ampicillin. Multiple clones were selected from each transformation to screen for best protein expression. To do this, 5 mL cultures were inoculated to grow for 16 hours at 37 degrees C. This culture was used to inoculate two more 5 mL cultures that were grown to an OD of 0.4-1 before inducing protein expression in one of the two cultures with 0.04 mM IPTG. Once protein expression was induced, cultures were incubated at room temperature (~20 degrees C). After 16 hours, 1 mL of culture was pelleted, and the pellet was resuspended in 60 μL Laemmli Buffer with 5% βME. 10 μL of each resuspension was loaded per well of an 8% polyacrylamide SDS-PAGE gel. Protein was detected using anti-His antibody as previously described.

HDC Exo- Purification

Purification of HDCexo- was performed as described for other recombinant POLIB variants in chapter 2 (page 37).

Results

Pilot Inductions

Inducible expression of each of these variants has been confirmed by Western blot detection using anti-His antibodies (Figure AIII.1). For clones of IB_HDCexo-, IB_HDEexo-, and IB_poldomain, some expression was seen in uninduced cultures. These constructs were re-transformed into BL21(DE3) cells to test clones for better controlled expression. At least one clone with successfully controlled induction was selected for each POLIB variant to be used for larger scale protein inductions and purifications. These constructs will add valuable insight into the enzymatic mechanisms of POLIB and could inform exciting *in vivo* experiments to observe how specific changes in the activity of POLIB can influence kDNA replication.

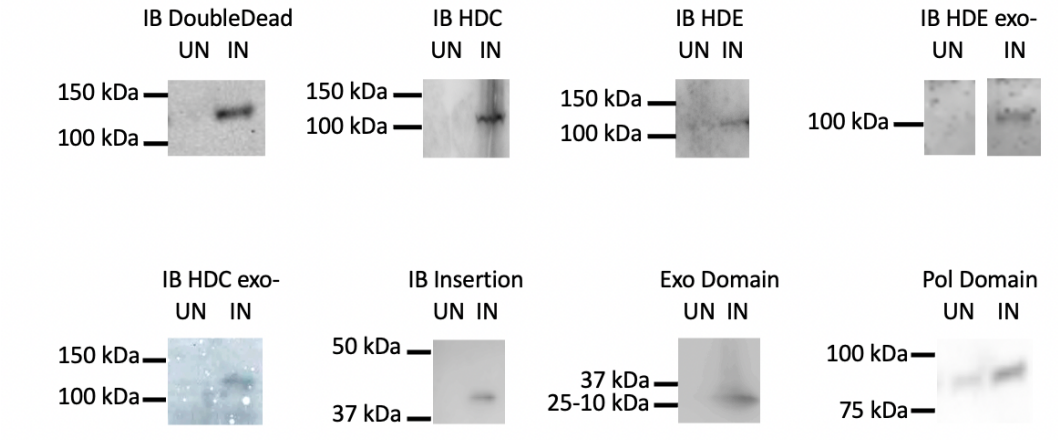


Figure AIII.1: Anti-His detection of recombinant POLIB variants from whole cell extracts of pilot protein inductions. Expected molecular weight of IB_insertion: 44 kDa, IB_exodomain: 24 kDa, IB_poldomain: 94 kDa. All other constructs expected to be 116 kDa.

HDC_exo- Purification and Activity

IBHDC_exo- was purified using the protein purification protocol established for other recombinant POLIB variants. In this purification, some IBHDC_exo- was seen in the flow through on from the Q-column, perhaps due to residual salt carried over from the Nickel column elution competitively binding to the Q column. Resulting Q column elutions were highly pure, and appear at the expected size on and SDS-PAGE gel (115.6 kDa) (Figure AIII.2). Fractions of Q column elution containing pure protein were pooled, concentrated and flash frozen in liquid nitrogen for future assays.

To determine whether the IBHDC_exo- S1310C mutation in the POLIB pol domain motif C would impact the substrate preference for RNA primed DNA that we described in chapter 2, IBHDC_exo- was used in extension assay with both a DNA primed substrate and an RNA primed substrate (Figure AIII.3A,B). More efficient extension of the RNA primer which closely resembles the activity of other characterized POLIB variants. Also similar to other characterized constructs, IBHDC_exo- minimally incorporates ribonucleotides from an RNA primer, and no incorporation of deoxyribonucleotides can be observed from a DNA primer.

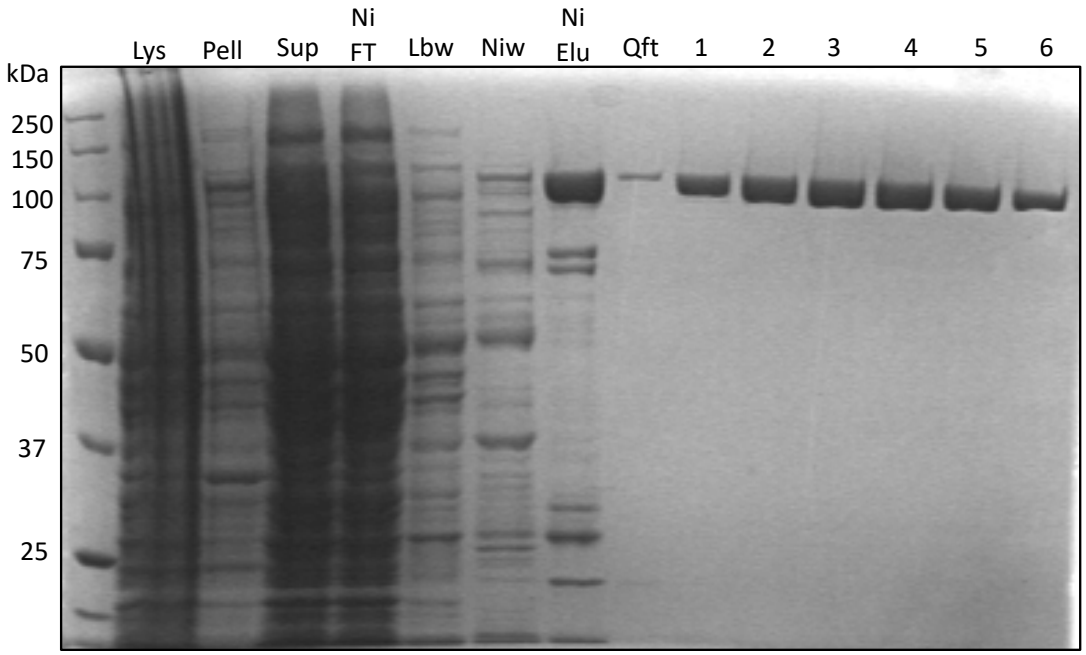


Figure AIII.2: Coomassie stained SDS-PAGE of Purification Fractions of HDC_Exo-. Each well loaded with 10 µL sample (7.5 µL protein sample with 2.5 µL Laemmli Buffer. Lys: whole cell lysate, Pell: insoluble pellet, Sup: soluble fraction, Ni FT: flow through from sample loading of nickel column, Lbw: loading buffer wash, niw: Nickel was buffer wash, Ni Elu: eluted protein from nickel column, Qft: flow through from loading of sample on Q column, 1-6: Q column elution fractions selected based on A280 absorbance N=1.

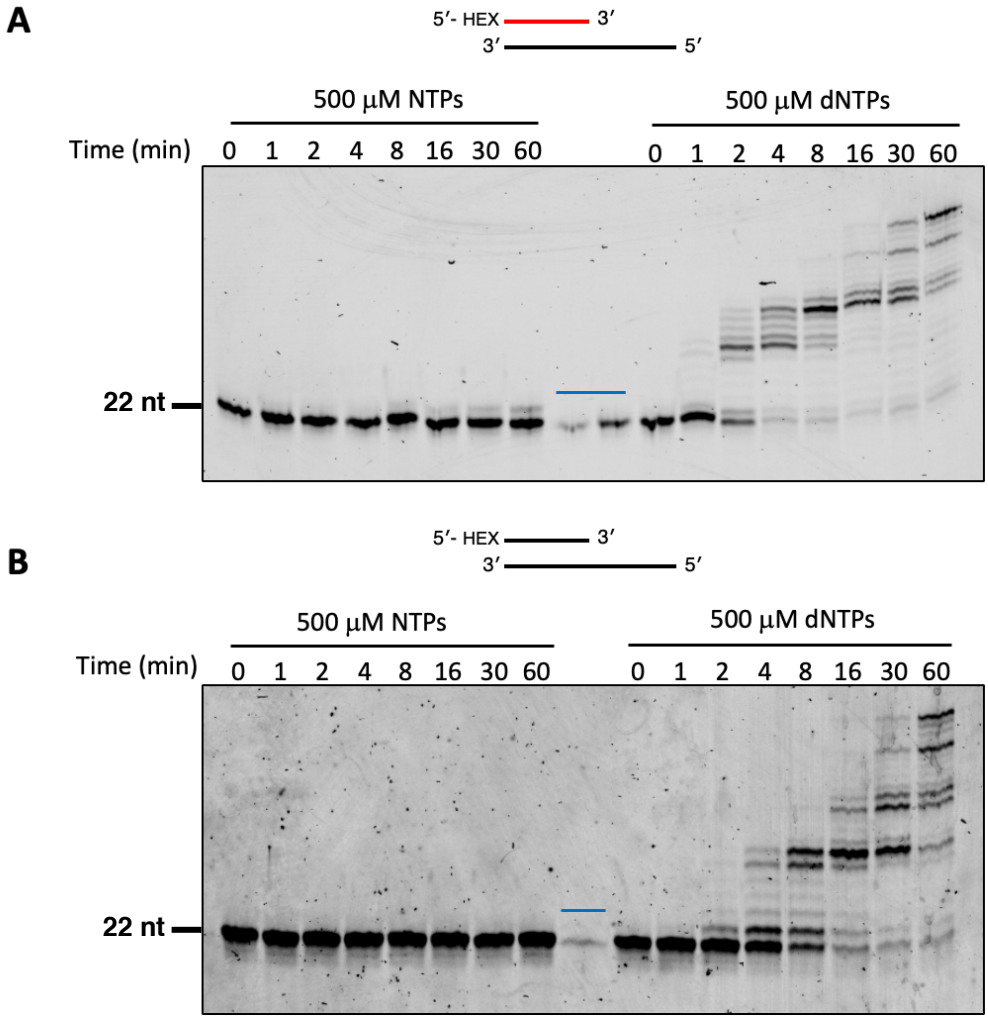


Figure AIII.3: Extension Activity of HDC Exo-. Quenched reaction time points visualized on denaturing polyacrylamide gel. Lanes marked with blue lines should be disregarded (Loading error) A) Progression of extension reaction including 840 nM RNA primed DNA substrate, 200 nM HDC exo-, 0.1% BSA, 1 mM DTT, 500 µM NTPs or dNTPs, 50 mM Tris pH 7.0. B) Progression of extension reaction including 840 nM DNA primed DNA substrate, 200 nM HDC exo-, 0.1% BSA, 1 mM DTT, 500 µM NTPs or dNTPs, 50 mM Tris pH 7.0. N=1.

Discussion

High-quality purifications need to be performed on most of these constructs, with the exception of HDC_exo-, to determine the impact of mutations on the activity of each

variant. It is likely that the dramatic truncations of the IB_insert and IB_exo POLIB variants will require a redesigned purifications scheme. These smaller proteins have higher predicted isoelectric points than the original POLIBwt_{trunc} and other recombinant POLIB variants with the same molecular weight. So, the pH of buffers used for the Quaternary ammonium column purification scheme should be adjusted to ensure that these variants are negatively charged in the buffer bind to the positively charged column.

We predict that enzymatic characterization of each of these POLIB variants will reveal valuable information about the functional impact of the unique structure of POLIB. Using the IB_insert and IB_exo constructs in exonuclease assays and DNA binding assays will allow us to understand how the structure of the pol domain impacts exo activity or the protein's DNA affinity.

The IB_HDC and IB_HDE variants will allow us to test if the HDS motif in the polymerase domain of wildtype POLIB is involved in the substrate preference of POLB for an RNA primer, or if this mutated S1310C residue impacts other aspects of POLIB extension activity such as rate or processivity. The inclusion of the exo- mutations is critical for evaluation of these variants, allowing us to quantify extension without competing exonuclease activity. In a pilot extension assay (N=1) of purified HDC exo- (Figure 3) the activity closely resembles IBexo- activity. More specifically, both variants exhibit more efficient extension from an RNA primer when compared to extension from an RNA primer. Both variants also only incorporate a single ribonucleotide from an RNA primer, and do not incorporate ribonucleotides from a DNA primer. This initial data

suggests that the serine variation in the Motif C (amino acid 1310) of POLIB may not be directly contributing to the protein's substrate preference for RNA primed DNA, or that other residues are also contributing to this aspect of the activity.

Assays of IBpoldomain will test whether the exonuclease domain embedded in the polymerase domain contributes to the relatively low polymerase activity of POLIB. These constructs will add valuable insight into the enzymatic mechanisms of POLIB and could inform novel *in vivo* experiments to observe how specific changes in the activity of POLIB can influence kDNA replication.

References

1. Porcelli, A. M.; Ghelli, A.; Zanna, C.; Pinton, P.; Rizzuto, R.; Rugolo, M. PH Difference across the Outer Mitochondrial Membrane Measured with a Green Fluorescent Protein Mutant. *Biochemical and Biophysical Research Communications* **2005**, *326* (4), 799–804. DOI:10.1016/j.bbrc.2004.11.105.
2. Bruhn, D. F.; Mozeleski, B.; Falkin, L.; Klingbeil, M. M. Mitochondrial DNA Polymerase POLIB Is Essential for Minicircle DNA Replication in African Trypanosomes. *Mol Microbiol* **2010**, *75* (6), 1414–1425. https://doi.org/10.1111/j.1365-2958.2010.07061.x.
3. Miller, J. C.; Delzell, S. B.; Concepción-Acevedo, J.; Boucher, M. J.; Klingbeil, M. M. A DNA Polymerization-Independent Role for Mitochondrial DNA Polymerase I-like Protein C in African Trypanosomes. *J Cell Sci* **2020**, *133* (9), jcs233072. https://doi.org/10.1242/jcs.233072.
4. Armstrong, R. Structure-Function Studies of the Trypanosome Mitochondrial Replication Protein POLIB. UMass Amherst, 2021.

Bibliography

1. Abad, M. F. C.; Di Benedetto, G.; Magalhães, P. J.; Filippin, L.; Pozzan, T. Mitochondrial PH Monitored by a New Engineered Green Fluorescent Protein Mutant*. *Journal of Biological Chemistry* **2004**, *279* (12), 11521–11529. DOI:10.1074/jbc.M306766200.
2. Abu-Elneel, K.; Robinson, D. R.; Drew, M. E.; Englund, P. T.; Shlomai, J. Intramitochondrial Localization of Universal Minicircle Sequence-Binding Protein, a Trypanosomatid Protein That Binds Kinetoplast Minicircle Replication Origins. *J Cell Biol* **2001**, *153* (4), 725–734. https://doi.org/10.1083/jcb.153.4.725.
3. Almagro Armenteros, J. J.; Salvatore, M.; Emanuelsson, O.; Winther, O.; von Heijne, G.; Elofsson, A.; Nielsen, H. Detecting Sequence Signals in Targeting Peptides Using Deep Learning. *Life Sci Alliance* **2019**, *2* (5), e201900429. DOI:10.26508/lsa.201900429.
4. Álvarez-Rodríguez, A.; Jin, B.-K.; Radwanska, M.; Magez, S. Recent Progress in Diagnosis and Treatment of Human African Trypanosomiasis Has Made the Elimination of This Disease a Realistic Target by 2030. *Front Med (Lausanne)* **2022**, *9*, 1037094. https://doi.org/10.3389/fmed.2022.1037094.
5. Amodeo, S.; Bregy, I.; Ochsenreiter, T. Mitochondrial Genome Maintenance - the Kinetoplast Story. *FEMS Microbiol Rev* **2022**, fuac047. https://doi.org/10.1093/femsre/fuac047.
6. Amos, B.; Aurrecoechea, C.; Barba, M.; Barreto, A.; Basenko, E. Y.; Bażant, W.; Belnap, R.; Blevins, A. S.; Böhme, U.; Brestelli, J.; Brunk, B. P.; Caddick, M.; Callan, D.; Campbell, L.; Christensen, M. B.; Christophides, G. K.; Crouch, K.; Davis, K.; DeBarry, J.; Doherty, R.; Duan, Y.; Dunn, M.; Falke, D.; Fisher, S.; Flicek, P.; Fox, B.; Gajria, B.; Giraldo-Calderón, G. I.; Harb, O. S.; Harper, E.; Hertz-Fowler, C.; Hickman, M. J.; Howington, C.; Hu, S.; Humphrey, J.; Iodice, J.; Jones, A.; Judkins, J.; Kelly, S. A.; Kissinger, J. C.; Kwon, D. K.; Lamoureux, K.; Lawson, D.; Li, W.; Lies, K.; Lodha, D.; Long, J.; MacCallum, R. M.; Maslen, G.; McDowell, M. A.; Nabrzyski, J.; Roos, D. S.; Rund, S. S. C.; Schulman, S. W.; Shanmugasundram, A.; Sitnik, V.; Spruill, D.; Starns, D.; Stoeckert, C. J., Jr; Tomko, S. S.; Wang, H.; Warrenfeltz, S.; Wieck, R.; Wilkinson, P. A.; Xu, L.; Zheng, J. VEuPathDB: The Eukaryotic Pathogen, Vector and Host Bioinformatics Resource Center. *Nucleic Acids Research* **2022**, *50* (D1), D898–D911. DOI:10.1093/nar/gkab929.
7. Armstrong, R. Structure-Function Studies of the Trypanosome Mitochondrial ReplicationProtein POLIB. UMass Amherst, 2021.
8. Aslett, M.; Aurrecoechea, C.; Berriman, M.; Brestelli, J.; Brunk, B. P.; Carrington, M.; Depledge, D. P.; Fischer, S.; Gajria, B.; Gao, X.; Gardner, M. J.; Gingle, A.; Grant, G.; Harb, O. S.; Heiges, M.; Hertz-Fowler, C.; Houston, R.; Innamorato, F.; Iodice, J.; Kissinger, J. C.; Kraemer, E.; Li, W.; Logan, F. J.; Miller, J. A.; Mitra, S.; Myler, P. J.; Nayak, V.; Pennington, C.; Phan, I.; Pinney, D. F.; Ramasamy, G.; Rogers, M. B.; Roos, D. S.; Ross, C.; Sivam, D.; Smith, D. F.; Srinivasamoorthy, G.; Stoeckert, C. J.; Subramanian, S.; Thibodeau, R.; Tivey, A.; Treatman, C.; Velarde, G.; Wang, H. TriTrypDB: A Functional Genomic Resource for the Trypanosomatidae. *Nucleic Acids Res* **2010**, *38* (Database issue), D457-462. DOI:10.1093/nar/gkp851.

9. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; Millán, C.; Park, H.; Adams, C.; Glassman, C. R.; DeGiovanni, A.; Pereira, J. H.; Rodrigues, A. V.; van Dijk, A. A.; Ebrecht, A. C.; Opperman, D. J.; Sagmeister, T.; Buhlheller, C.; Pavkov-Keller, T.; Rathinaswamy, M. K.; Dalwadi, U.; Yip, C. K.; Burke, J. E.; Garcia, K. C.; Grishin, N. V.; Adams, P. D.; Read, R. J.; Baker, D. Accurate Prediction of Protein Structures and Interactions Using a 3-Track Neural Network. *Science* **2021**, *373* (6557), 871–876. DOI:10.1126/science.abj8754.
10. Balint, E.; Unk, I. Selective Metal Ion Utilization Contributes to the Transformation of the Activity of Yeast Polymerase η from DNA Polymerization toward RNA Polymerization. *Int J Mol Sci* **2020**, *21* (21), 8248. DOI:10.3390/ijms21218248.
11. Balmer, O.; Beadell, J. S.; Gibson, W.; Caccone, A. Phylogeography and Taxonomy of Trypanosoma Brucei. *PLoS Negl Trop Dis* **2011**, *5* (2), e961. https://doi.org/10.1371/journal.pntd.0000961.
12. Baruch-Torres, N.; Brieba, L. G. Plant Organellar DNA Polymerases Are Replicative and Translesion DNA Synthesis Polymerases. *Nucleic Acids Research* **2017**, *45* (18), 10751. DOI:10.1093/nar/gkx744.
13. Baudouin, H. C. M.; Pfeiffer, L.; Ochsenreiter, T. A Comparison of Three Approaches for the Discovery of Novel Tripartite Attachment Complex Proteins in Trypanosoma Brucei. *PLoS Negl Trop Dis* **2020**, *14* (9), e0008568. https://doi.org/10.1371/journal.pntd.0008568.
14. Belmonte-Reche, E.; Martínez-García, M.; Guédin, A.; Zuffo, M.; Arévalo-Ruiz, M.; Doria, F.; Campos-Salinas, J.; Maynadier, M.; López-Rubio, J. J.; Freccero, M.; Mergny, J.-L.; Pérez-Victoria, J. M.; Morales, J. C. G-Quadruplex Identification in the Genome of Protozoan Parasites Points to Naphthalene Diimide Ligands as New Antiparasitic Agents. *J Med Chem* **2018**, *61* (3), 1231–1240. https://doi.org/10.1021/acs.jmedchem.7b01672.
15. Bochman, M. L.; Sabouri, N.; Zakian, V. A. Unwinding the Functions of the Pif1 Family Helicases. *DNA Repair (Amst)* **2010**, *9* (3), 237–249. https://doi.org/10.1016/j.dnarep.2010.01.008.
16. Brautigam, C. A.; Sun, S.; Piccirilli, J. A.; Steitz, T. A. Structures of Normal Single-Stranded DNA and Deoxyribo-3‘-S-Phosphorothiolates Bound to the 3‘-5‘ Exonucleolytic Active Site of DNA Polymerase I from Escherichia Coli,. *Biochemistry* **1999**, *38* (2), 696–704. DOI:10.1021/bi981537g.
17. Brieba, L. G. Structure-Function Analysis Reveals the Singularity of Plant Mitochondrial DNA Replication Components: A Mosaic and Redundant System. *Plants (Basel)* **2019**, *8* (12), 533. https://doi.org/10.3390/plants8120533.
18. Bruhn, D. F.; Mozeleski, B.; Falkin, L.; Klingbeil, M. M. Mitochondrial DNA Polymerase POLIB Is Essential for Minicircle DNA Replication in African Trypanosomes. *Mol Microbiol* **2010**, *75* (6), 1414–1425. https://doi.org/10.1111/j.1365-2958.2010.07061.x.
19. Bruhn, D. F.; Sammartino, M. P.; Klingbeil, M. M. Three Mitochondrial DNA Polymerases Are Essential for Kinetoplast DNA Replication and Survival of Bloodstream Form Trypanosoma Brucei. *Eukaryot Cell* **2011**, *10* (6), 734–743. https://doi.org/10.1128/EC.05008-11.

20. Capewell, P.; Cren-Travaillé, C.; Marchesi, F.; Johnston, P.; Clucas, C.; Benson, R. A.; Gorman, T.-A.; Calvo-Alvarez, E.; Crouzols, A.; Jouvion, G.; Jamonneau, V.; Weir, W.; Stevenson, M. L.; O'Neill, K.; Cooper, A.; Swar, N.-R. K.; Bucheton, B.; Ngoyi, D. M.; Garside, P.; Rotureau, B.; MacLeod, A. The Skin Is a Significant but Overlooked Anatomical Reservoir for Vector-Borne African Trypanosomes. *Elife* **2016**, *5*, e17716. https://doi.org/10.7554/eLife.17716.
21. Carter, N. S.; Berger, B. J.; Fairlamb, A. H. Uptake of Diamidine Drugs by the P2 Nucleoside Transporter in Melarsen-Sensitive and -Resistant Trypanosoma Brucei Brucei. *Journal of Biological Chemistry* **1995**, *270* (47), 28153–28157. https://doi.org/10.1074/jbc.270.47.28153.
22. Chandler, J.; Vandoros, A. V.; Mozeleski, B.; Klingbeil, M. M. Stem-Loop Silencing Reveals That a Third Mitochondrial DNA Polymerase, POLID, Is Required for Kinetoplast DNA Replication in Trypanosomes. *Eukaryot Cell* **2008**, *7* (12), 2141–2146. https://doi.org/10.1128/EC.00199-08.
23. Chowdhury, Roy.; Bakshi, R.; Wang, J.; Yildirir, G.; Liu, B.; Pappas-Brown, V.; Tolun, G.; Griffith, J. D.; Shapiro, T. A.; Jensen, R. E.; Englund, P. T. The Killing of African Trypanosomes by Ethidium Bromide. *PLoS Pathog* **2010**, *6* (12), e1001226. https://doi.org/10.1371/journal.ppat.1001226.
24. Concepción-Acevedo, J.; Luo, J.; Klingbeil, M. M. Dynamic Localization of Trypanosoma Brucei Mitochondrial DNA Polymerase ID. *Eukaryot Cell* **2012**, *11* (7), 844–855. DOI:10.1128/EC.05291-11.
25. Concepción-Acevedo, J.; Miller, J. C.; Boucher, M. J.; Klingbeil, M. M. Cell Cycle Localization Dynamics of Mitochondrial DNA Polymerase IC in African Trypanosomes. *Mol Biol Cell* **2018**, *29* (21), 2540–2552. DOI:10.1091/mbc.E18-02-0127.
26. Corkey, B. E.; Duszynski, J.; Rich, T. L.; Matschinsky, B.; Williamson, J. R. Regulation of Free and Bound Magnesium in Rat Hepatocytes and Isolated Mitochondria. *Journal of Biological Chemistry* **1986**, *261* (6), 2567–2574. DOI:10.1016/S0021-9258(17)35825-8.
27. Crozier, T. W. M.; Tinti, M.; Larance, M.; Lamond, A. I.; Ferguson, M. A. J. Prediction of Protein Complexes in Trypanosoma Brucei by Protein Correlation Profiling Mass Spectrometry and Machine Learning. *Mol Cell Proteomics* **2017**, *16* (12), 2254–2267. https://doi.org/10.1074/mcp.O117.068122.
28. Dahan, D.; Tsirkas, I.; Dovrat, D.; Sparks, M. A.; Singh, S. P.; Galletto, R.; Aharoni, A. Pif1 Is Essential for Efficient Replisome Progression through Lagging Strand G-Quadruplex DNA Secondary Structures. *Nucleic Acids Res* **2018**, *46* (22), 11847–11857. https://doi.org/10.1093/nar/gky1065.
29. Dewar, C. E.; MacGregor, P.; Cooper, S.; Gould, M. K.; Matthews, K. R.; Savill, N. J.; Schnaufer, A. Mitochondrial DNA Is Critical for Longevity and Metabolism of Transmission Stage Trypanosoma Brucei. *PLoS Pathog* **2018**, *14* (7), e1007195. https://doi.org/10.1371/journal.ppat.1007195.
30. Drew, M. E.; Englund, P. T. Intramitochondrial Location and Dynamics of Crithidia Fasciculata Kinetoplast Minicircle Replication Intermediates. *J Cell Biol* **2001**, *153* (4), 735–744. https://doi.org/10.1083/jcb.153.4.735

31. Downey, N.; Hines, J. C.; Sinha, K. M.; Ray, D. S. Mitochondrial DNA Ligases of Trypanosoma Brucei. *Eukaryot Cell* **2005**, *4* (4), 765–774. https://doi.org/10.1128/EC.4.4.765- 774.2005.
32. Eshetu, E.; Begejo, B. The Current Situation and Diagnostic Approach of Nagana in Africa: A Review. **2015**.
33. Fisk, J. C.; Li, J.; Wang, H.; Aletta, J. M.; Qu, J.; Read, L. K. Proteomic Analysis Reveals Diverse Classes of Arginine Methylproteins in Mitochondria of Trypanosomes. *Mol Cell Proteomics* **2013**, *12* (2), 302–311. DOI:10.1074/mcp.M112.022533.
34. García-Medel, P. L.; Baruch-Torres, N.; Peralta-Castro, A.; Trasviña-Arenas, C. H.; Torres-Larios, A.; Brieba, L. G. Plant Organellar DNA Polymerases Repair Double-Stranded Breaks by Microhomology-Mediated End-Joining. *Nucleic Acids Res* **2019**, *47* (6), 3028–3044. https://doi.org/10.1093/nar/gkz039.
35. Gluenz, E.; Shaw, M. K.; Gull, K. Structural Asymmetry and Discrete Nucleic Acid Subdomains in the Trypanosoma Brucei Kinetoplast. *Mol Microbiol* **2007**, *64* (6), 1529–1539. DOI:10.1111/j.1365-2958.2007.05749.x.
36. Graham SV, Barry JD. Transcriptional regulation of metacyclic variant surface glycoprotein gene expression during the life cycle of Trypanosoma brucei. *Mol Cell Biol*. 1995;15: 5945–5956. doi: 10.1128/MCB.15.11.5945
37. Gunter, T. E.; Miller, L. M.; Gavin, C. E.; Eliseev, R.; Salter, J.; Buntinas, L.; Alexandrov, A.; Hammond, S.; Gunter, K. K. Determination of the Oxidation States of Manganese in Brain, Liver, and Heart Mitochondria. *Journal of Neurochemistry* **2004**, *88* (2), 266–280. DOI:10.1046/j.1471-4159.2003.02122.x.
38. Harada, R.; Hirakawa, Y.; Yabuki, A.; Kashiyama, Y.; Maruyama, M.; Onuma, R.; Soukal, P.; Miyagishima, S.; Hampl, V.; Tanifuji, G.; Inagaki, Y. Inventory and Evolution of Mitochondrion-Localized Family A DNA Polymerases in Euglenozoa. *Pathogens* **2020**, *9* (4), 257.DOI:10.3390/pathogens9040257.
39. Hemphill, W. O.; Perrino, F. W. Chapter Eight - Measuring TREX1 and TREX2 Exonuclease Activities. In *Methods in Enzymology*; Sohn, J., Ed.; DNA Sensors and Inflammasomes; Academic Press, 2019; Vol. 625, pp 109–133. DOI:10.1016/bs.mie.2019.05.004.
40. Hines, J. C.; Ray, D. S. A Mitochondrial DNA Primase Is Essential for Cell Growth and Kinetoplast DNA Replication in Trypanosoma Brucei. *Mol Cell Biol* **2010**, *30* (6), 1319–1328. https://doi.org/10.1128/MCB.01231-09.
41. Hines, J. C.; Ray, D. S. A Second Mitochondrial DNA Primase Is Essential for Cell Growth and Kinetoplast Minicircle DNA Replication in Trypanosoma Brucei. *Eukaryot Cell* **2011**, *10* (3), 445–454. https://doi.org/10.1128/EC.00308-10.
42. Hogg, M.; Seki, M.; Wood, R. D.; Doublié, S.; Wallace, S. S. Lesion Bypass Activity of DNA Polymerase θ (POLQ) Is an Intrinsic Property of the Pol Domain and Depends on Unique Sequence Inserts. *Journal of Molecular Biology* **2011**, *405* (3), 642–652. DOI:10.1016/j.jmb.2010.10.041.
43. Humphreys, I. R.; Pei, J.; Baek, M.; Krishnakumar, A.; Anishchenko, I.; Ovchinnikov, S.; Zhang, J.; Ness, T. J.; Banjade, S.; Bagde, S. R.; Stancheva, V. G.; Li, X.-H.; Liu, K.; Zheng, Z.; Barrero, D. J.; Roy, U.; Kuper, J.; Fernández, I. S.; Szakal, B.; Branzei, D.; Rizo, J.; Kisker, C.; Greene, E. C.; Biggins, S.; Keeney, S.; Miller, E. A.; Fromme, J. C.;

Hendrickson, T. L.; Cong, Q.; Baker, D. Computed Structures of Core Eukaryotic Protein Complexes. *Science* **2021**, *374* (6573), eabm4805. DOI:10.1126/science.abm4805.
44. Jensen, R. E.; Englund, P. T. Network News: The Replication of Kinetoplast DNA. *Annu Rev Microbiol* **2012**, *66*, 473–491. https://doi.org/10.1146/annurev-micro-092611-15005
45. Jensen, B. C.; Kifer, C. T.; Brekken, D. L.; Randall, A. C.; Wang, Q.; Drees, B. L.; Parsons, M. Characterization of Protein Kinase CK2 from Trypanosoma Brucei. *Mol Biochem Parasitol* **2007**, *151* (1), 28–40. DOI:10.1016/j.molbiopara.2006.10.002.
46. Johnson, A. A.; Johnson, K. A. Exonuclease Proofreading by Human Mitochondrial DNA Polymerase. *Journal of Biological Chemistry* **2001**, *276* (41), 38097–38107. DOI:10.1074/jbc.M106046200.
47. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.
48. Kasiviswanathan, R.; Collins, T. R. L.; Copeland, W. C. The Interface of Transcription and DNA Replication in the Mitochondria. *Biochim Biophys Acta* **2012**, *1819* (9–10), 970–978. https://doi.org/10.1016/j.bbagrm.2011.12.005.
49. Kennedy, P. G. Clinical Features, Diagnosis, and Treatment of Human African Trypanosomiasis (Sleeping Sickness). *Lancet Neurol* **2013**, *12* (2), 186–194. https://doi.org/10.1016/S1474-4422(12)70296-X.
50. Klingbeil, M. M.; Motyka, S. A.; Englund, P. T. Multiple Mitochondrial DNA Polymerases in Trypanosoma Brucei. *Mol Cell* **2002**, *10* (1), 175–186. https://doi.org/10.1016/s1097-2765(02)00571-3.
51. Konji, V.; Montag, A.; Sandri, G.; Nordenbrand, K.; Ernster, L. Transport of Ca2+ and Mn2+ by Mitochondria from Rat Liver, Heart and Brain. *Biochimie* **1985**, *67* (12), 1241–1250. DOI:10.1016/S0300-9084(85)80133-4.
52. Lamour, N.; Rivière, L.; Coustou, V.; Coombs, G. H.; Barrett, M. P.; Bringaud, F. Proline Metabolism in Procyclic Trypanosoma Brucei Is Down-Regulated in the Presence of Glucose. *J Biol Chem* **2005**, *280* (12), 11902–11910. https://doi.org/10.1074/jbc.M414274200.
53. Lee, Y.-S.; Kennedy, W. D.; Yin, Y. W. Structural Insight into Processive Human Mitochondrial DNA Synthesis and Disease-Related Polymerase Mutations. *Cell* **2009**, *139* (2), 312–324. https://doi.org/10.1016/j.cell.2009.07.050.
54. Liu, B.; Wang, J.; Yaffe, N.; Lindsay, M. E.; Zhao, Z.; Zick, A.; Shlomai, J.; Englund, P. T. Trypanosomes Have Six Mitochondrial DNA Helicases with One Controlling Kinetoplast Maxicircle Replication. *Mol Cell* **2009**, *35* (4), 490–501. https://doi.org/10.1016/j.molcel.2009.07.004.
55. Llopis, J.; McCaffery, J. M.; Miyawaki, A.; Farquhar, M. G.; Tsien, R. Y. Measurement of Cytosolic, Mitochondrial, and Golgi PH in Single Living Cells with Green Fluorescent

Proteins. *Proceedings of the National Academy of Sciences* **1998**, *95* (12), 6803–6808. DOI:10.1073/pnas.95.12.6803.
56. Longbottom, J.; Wamboga, C.; Bessell, P. R.; Torr, S. J.; Stanton, M. C. Optimising Passive Surveillance of a Neglected Tropical Disease in the Era of Elimination: A Modelling Study. *PLoS Negl Trop Dis* **2021**, *15* (3), e0008599. https://doi.org/10.1371/journal.pntd.0008599.
57. Lopes, D. de O.; Schamber-Reis, B. L. F.; Regis-da-Silva, C. G.; Rajão, M. A.; DaRocha, W. D.; Macedo, A. M.; Franco, G. R.; Nardelli, S. C.; Schenkman, S.; Hoffmann, J.-S.; Cazaux, C.; Pena, S. D. J.; Teixeira, S. M. R.; Machado, C. R. Biochemical Studies with DNA Polymerase β and DNA Polymerase β-PAK of Trypanosoma Cruzi Suggest the Involvement of These Proteins in Mitochondrial DNA Maintenance. *DNA Repair* **2008**, *7* (11), 1882–1892.
58. Lukeš, J.; Lys Guilbride, D.; Votýpka, J.; Zíková, A.; Benne, R.; Englund, P. T. Kinetoplast DNA Network: Evolution of an Improbable Structure. *Eukaryot Cell* **2002**, *1* (4), 495–502. DOI:10.1128/EC.1.4.495-502.2002.
59. MacGregor P, Savill NJ, Hall D, Matthews KR. Transmission stages dominate trypanosome within-host dynamics during chronic infections. Cell Host Microbe. 2011 Apr 21;9(4):310-8. doi: 10.1016/j.chom.2011.03.013. PMID: 21501830; PMCID: PMC3094754.
60. Maldonado, E.; Rojas, D. A.; Urbina, F.; Solari, A. T. Cruzi DNA Polymerase Beta (Tcpolβ) Is Phosphorylated in Vitro by CK1, CK2 and TcAUK1 Leading to the Potentiation of Its DNA Synthesis Activity. *PLoS Negl Trop Dis* **2021**, *15* (7), e0009588. https://doi.org/10.1371/journal.pntd.0009588.
61. Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. LDDT: A Local Superposition-Free Score for Comparing Protein Structures and Models Using Distance Difference Tests. *Bioinformatics* **2013**, *29* (21), 2722–2728. DOI:10.1093/bioinformatics/btt473.
62. Melendy, T.; Sheline, C.; Ray, D. S. Localization of a Type II DNA Topoisomerase to Two Sites at the Periphery of the Kinetoplast DNA of Crithidia Fasciculata. *Cell* **1988**, *55* (6), 1083–1088. DOI:10.1016/0092-8674(88)90252-8.
63. McGraw, N. J.; Bailey, J. N.; Cleaves, G. R.; Dembinski, D. R.; Gocke, C. R.; Joliffe, L. K.; MacWright, R. S.; McAllister, W. T. Sequence and Analysis of the Gene for Bacteriophage T3 RNA Polymerase. *Nucleic Acids Res* **1985**, *13* (18), 6753–6766. DOI:10.1093/nar/13.18.6753
64. Miller, J. C.; Delzell, S. B.; Concepción-Acevedo, J.; Boucher, M. J.; Klingbeil, M. M. A DNA Polymerization-Independent Role for Mitochondrial DNA Polymerase I-like Protein C in African Trypanosomes. *J Cell Sci* **2020**, *133* (9), jcs233072. https://doi.org/10.1242/jcs.233072.
65. Milman, N.; Motyka, S. A.; Englund, P. T.; Robinson, D.; Shlomai, J. Mitochondrial Origin-Binding Protein UMSBP Mediates DNA Replication and Segregation in Trypanosomes. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (49), 19250–19255. https://doi.org/10.1073/pnas.0706858104.
66. Miskinyte, M.; Dawson, J. C.; Makda, A.; Doughty-Shenton, D.; Carragher, N. O.; Schnaufer, A. A Novel High-Content Phenotypic Screen To Identify Inhibitors of

Mitochondrial DNA Maintenance in Trypanosomes. *Antimicrob Agents Chemother* **2022**, *66* (2), e01980-21. https://doi.org/10.1128/AAC.01980-21.
67. Muellner, J.; Schmidt, K. H. Yeast Genome Maintenance by the Multifunctional PIF1 DNA Helicase Family. *Genes (Basel)* **2020**, *11* (2), 224. https://doi.org/10.3390/genes11020
68. Oates, M. E.; Stahlhacke, J.; Vavoulis, D. V.; Smithers, B.; Rackham, O. J. L.; Sardar, A. J.; Zaucha, J.; Thurlby, N.; Fang, H.; Gough, J. The SUPERFAMILY 1.75 Database in 2014: A Doubling of Data. *Nucleic Acids Research* **2015**, *43* (D1), D227–D233. DOI:10.1093/nar/gku1041.
69. Orij, R.; Postmus, J.; Ter Beek, A.; Brul, S.; Smits, G. J. In Vivo Measurement of Cytosolic and Mitochondrial PH Using a PH-Sensitive GFP Derivative in Saccharomyces Cerevisiae Reveals a Relation between Intracellular PH and Growth. *Microbiology (Reading)* **2009**, *155* (Pt 1), 268–278. DOI:10.1099/mic.0.022038-0.
70. Paeschke, K.; Bochman, M. L.; Garcia, P. D.; Cejka, P.; Friedman, K. L.; Kowalczykowski, S. C.; Zakian, V. A. Pif1 Family Helicases Suppress Genome Instability at G-Quadruplex Motifs. *Nature* **2013**, *497* (7450), 458–462. https://doi.org/10.1038/nature12149.
71. Papadopoulos, J. S.; Agarwala, R. COBALT: Constraint-Based Alignment Tool for Multiple Protein Sequences. *Bioinformatics* **2007**, *23* (9), 1073–1079. DOI:10.1093/bioinformatics/btm076.
72. Parent, J.-S.; Lepage, E.; Brisson, N. Divergent Roles for the Two PolI-Like Organelle DNA Polymerases of Arabidopsis1[W][OA]. *Plant Physiol* **2011**, *156* (1), 254–262. DOI:10.1104/pp.111.173849.
73. Park, J.; Baruch-Torres, N.; Iwai, S.; Herrmann, G. K.; Brieba, L. G.; Yin, Y. W. Human Mitochondrial DNA Polymerase Metal Dependent UV Lesion Bypassing Ability. *Front Mol Biosci* **2022**, *9*, 808036. DOI:10.3389/fmolb.2022.808036.
74. Pays, E.; Vanhollebeke, B.; Uzureau, P.; Lecordier, L.; Pérez-Morga, D. The Molecular Arms Race between African Trypanosomes and Humans. *Nat Rev Microbiol* **2014**, *12* (8), 575–584. https://doi.org/10.1038/nrmicro3298.
75. Perera, R. L.; Torella, R.; Klinge, S.; Kilkenny, M. L.; Maman, J. D.; Pellegrini, L. Mechanism for Priming DNA Synthesis by Yeast DNA Polymerase α. *eLife* **2013**, *2*, e00482. DOI:10.7554/eLife.00482.
76. Peeva, V.; Blei, D.; Trombly, G.; Corsi, S.; Szukszto, M. J.; Rebelo-Guiomar, P.; Gammage, P. A.; Kudin, A. P.; Becker, C.; Altmüller, J.; Minczuk, M.; Zsurka, G.; Kunz, W. S. Linear Mitochondrial DNA Is Rapidly Degraded by Components of the Replication Machinery. *Nat Commun* **2018**, *9* (1), 1727. DOI:10.1038/s41467-018-04131-w.
77. Peter, B.; Falkenberg, M. TWINKLE and Other Human Mitochondrial DNA Helicases: Structure, Function and Disease. *Genes (Basel)* **2020**, *11* (4), 408. https://doi.org/10.3390/genes11040408.
78. Pike, J. E.; Henry, R. A.; Burgers, P. M. J.; Campbell, J. L.; Bambara, R. A. An Alternative Pathway for Okazaki Fragment Processing: Resolution of Fold-Back Flaps by Pif1 Helicase. *J Biol Chem* **2010**, *285* (53), 41712–41723. https://doi.org/10.1074/jbc.M110.146894.

79. Porcelli, A. M.; Ghelli, A.; Zanna, C.; Pinton, P.; Rizzuto, R.; Rugolo, M. PH Difference across the Outer Mitochondrial Membrane Measured with a Green Fluorescent Protein Mutant. *Biochemical and Biophysical Research Communications* **2005**, *326* (4), 799–804. DOI:10.1016/j.bbrc.2004.11.105.
80. Poveda, A.; Méndez, M. Á.; Armijos-Jaramillo, V. Analysis of DNA Polymerases Reveals Specific Genes Expansion in Leishmania and Trypanosoma Spp. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 570493. DOI:10.3389/fcimb.2020.570493.
81. Romani, A. M. P. Cellular Magnesium Homeostasis. *Archives of Biochemistry and Biophysics* **2011**, *512* (1), 1–23. DOI:10.1016/j.abb.2011.05.010.
82. Rotureau, B.; Van Den Abbeele, J. Through the Dark Continent: African Trypanosome Development in the Tsetse Fly. *Front Cell Infect Microbiol* **2013**, *3*, 53. https://doi.org/10.3389/fcimb.2013.00053.
83. Saxowsky, T. T.; Choudhary, G.; Klingbeil, M. M.; Englund, P. T. Trypanosoma Brucei Has Two Distinct Mitochondrial DNA Polymerase β Enzymes. *Journal of Biological Chemistry* **2003**, *278* (49), 49095–49101. DOI:10.1074/jbc.M308565200.
84. Seow, F.; Sato, S.; Janssen, C. S.; Riehle, M. O.; Mukhopadhyay, A.; Phillips, R. S.; Wilson, R. J. M. (Iain); Barrett, M. P. The Plastidic DNA Replication Enzyme Complex of Plasmodium Falciparum. *Molecular and Biochemical Parasitology* **2005**, *141* (2), 145–153. DOI:10.1016/j.molbiopara.2005.02.002.
85. Schamber-Reis, B. L. F.; Nardelli, S.; Régis-Silva, C. G.; Campos, P. C.; Cerqueira, P. G.; Lima, S. A.; Franco, G. R.; Macedo, A. M.; Pena, S. D. J.; Cazaux, C.; Hoffmann, J.-S.; Motta, M. C. M.; Schenkman, S.; Teixeira, S. M. R.; Machado, C. R. DNA Polymerase Beta from Trypanosoma Cruzi Is Involved in Kinetoplast DNA Replication and Repair of Oxidative Lesions. *Mol Biochem Parasitol* **2012**, *183* (2), 122–131.
86. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat Methods* **2012**, *9* (7), 671–675. DOI:10.1038/nmeth.2089
87. Sinha, K. M.; Hines, J. C.; Downey, N.; Ray, D. S. Mitochondrial DNA Ligase in Crithidia Fasciculata. *Proc Natl Acad Sci U S A* **2004**, *101* (13), 4361–4366. https://doi.org/10.1073/pnas.0305705101.
88. Singh, K.; Modak, M. J. Contribution of Polar Residues of the J-Helix in the 3‘−5‘ Exonuclease Activity of Escherichia Coli DNA Polymerase I (Klenow Fragment): Q677 Regulates the Removal of Terminal Mismatch. *Biochemistry* **2005**, *44* (22), 8101–8110. DOI:10.1021/bi050140r.
89. Sowers, M. L.; Anderson, A. P. P.; Wrabl, J. O.; Yin, Y. W. Networked Communication between Polymerase and Exonuclease Active Sites in Human Mitochondrial DNA Polymerase. *J. Am. Chem. Soc.* **2019**, *141* (27), 10821–10829. DOI:10.1021/jacs.9b04655.
90. Sykora, P.; Kanno, S.; Akbari, M.; Kulikowicz, T.; Baptiste, B. A.; Leandro, G. S.; Lu, H.; Tian, J.; May, A.; Becker, K. A.; Croteau, D. L.; Wilson, D. M.; Sobol, R. W.; Yasui, A.; Bohr, V. A. DNA Polymerase Beta Participates in Mitochondrial DNA Repair. *Mol Cell Biol* **2017**, *37* (16), e00237-17. https://doi.org/10.1128/MCB.00237-17.
91. Tabor, S.; Huber, H. E.; Richardson, C. C. Escherichia Coli Thioredoxin Confers Processivity on the DNA Polymerase Activity of the Gene 5 Protein of Bacteriophage T7.

Journal of Biological Chemistry **1987**, *262* (33), 16212–16223. DOI:10.1016/S0021-9258(18)47718-6.

92. Taladriz, S.; Hanke, T.; Ramiro, M. J.; García-Díaz, M.; García De Lacoba, M.; Blanco, L.; Larraga, V. Nuclear DNA Polymerase Beta from Leishmania Infantum. Cloning, Molecular Analysis and Developmental Regulation. *Nucleic Acids Res* **2001**, *29* (18), 3822–3834. https://doi.org/10.1093/nar/29.18.3822.
93. Teklemariam, T. A.; Rivera, O. D.; Nelson, S. W. Chapter Five - Kinetic Analysis of the Exonuclease Activity of the Bacteriophage T4 Mre11–Rad50 Complex. In *Methods in Enzymology*; Spies, M., Malkova, A., Eds.; Mechanisms of DNA Recombination and Genome Rearrangements: Methods to Study Homologous Recombination; Academic Press, 2018; Vol. 600, pp 135–156. DOI:10.1016/bs.mie.2017.12.007.
94. Thomas, J. A.; Baker, N.; Hutchinson, S.; Dominicus, C.; Trenaman, A.; Glover, L.; Alsford, S.; Horn, D. Insights into Antitrypanosomal Drug Mode-of-Action from Cytology-Based Profiling. *PLoS Negl Trop Dis* **2018**, *12* (11), e0006980.
95. Torri, A. F.; Englund, P. T. Purification of a Mitochondrial DNA Polymerase from Crithidia Fasciculata. *J Biol Chem* **1992**, *267* (7), 4786–4792.
96. Torri, A. F.; Kunkel, T. A.; Englund, P. T. A Beta-like DNA Polymerase from the Mitochondrion of the Trypanosomatid Crithidia Fasciculata. *J Biol Chem* **1994**, *269* (11), 8165–8171.
97. Trindade, S.; Rijo-Ferreira, F.; Carvalho, T.; Pinto-Neves, D.; Guegan, F.; Aresta-Branco, F.; Bento, F.; Young, S. A.; Pinto, A.; Van Den Abbeele, J.; Ribeiro, R. M.; Dias, S.; Smith, T. K.; Figueiredo, L. M. Trypanosoma Brucei Parasites Occupy and Functionally Adapt to the Adipose Tissue in Mice. *Cell Host Microbe* **2016**, *19* (6), 837–848. https://doi.org/10.1016/j.chom.2016.05.002.
98. Schneider, A.; Ochsenreiter, T. Failure Is Not an Option – Mitochondrial Genome Segregation in Trypanosomes. *Journal of Cell Science* **2018**, *131* (18), jcs221820. https://doi.org/10.1242/jcs.221820.
99. Shapiro, T. A.; Englund, P. T. Selective Cleavage of Kinetoplast DNA Minicircles Promoted by Antitrypanosomal Drugs. *Proc Natl Acad Sci U S A* **1990**, *87* (3), 950–954. https://doi.org/10.1073/pnas.87.3.950.
100. Sima, N.; McLaughlin, E. J.; Hutchinson, S.; Glover, L. Escaping the Immune System by DNA Repair and Recombination in African Trypanosomes. *Open Biol.* **2019**, *9* (11), 190182. https://doi.org/10.1098/rsob.190182.
101. Swallow, B. M. Impacts of Trypanosomiasis on African Agriculture. **1999.**
102. Vaisman, A.; Ling, H.; Woodgate, R.; Yang, W. Fidelity of Dpo4: Effect of Metal Ions, Nucleotide Selection and Pyrophosphorolysis. *EMBO J* **2005**, *24* (17), 2957–2967. DOI:10.1038/sj.emboj.7600786.
103. Vanderheyden, N.; Wong, J.; Docampo, R. A Pyruvate-Proton Symport and an H+-ATPase Regulate the Intracellular PH of Trypanosoma Brucei at Different Stages of Its Life Cycle. *Biochem J* **2000**, *346* (Pt 1), 53–62. DOI:10.1042/bj3460053
104. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold

Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Research* **2022**, *50* (D1), D439–D444. DOI:10.1093/nar/gkab1061.

105. Venturelli, A.; Tagliazucchi, L.; Lima, C.; Venuti, F.; Malpezzi, G.; Magoulas, G. E.; Santarem, N.; Calogeropoulou, T.; Cordeiro-da-Silva, A.; Costi, M. P. Current Treatments to Control African Trypanosomiasis and One Health Perspective. *Microorganisms* **2022**, *10* (7), 1298. https://doi.org/10.3390/microorganisms10071298.
106. Wang, J.; Englund, P. T.; Jensen, R. E. TbPIF8, a Trypanosoma Brucei Protein Related to the Yeast Pif1 Helicase, Is Essential for Cell Viability and Mitochondrial Genome Maintenance. *Mol Microbiol* **2012**, *83* (3), 471–485. https://doi.org/10.1111/j.1365-2958.2011.07938.x.
107. Watkins, T. I.; Woolfe, G. Effect of Changing the Quaternizing Group on the Trypanocidal Activity of Dimidium Bromide. *Nature* **1952**, *169* (4299), 506–506. https://doi.org/10.1038/169506a0.
108. Wheeler, R. J. A Resource for Improved Predictions of Trypanosoma and Leishmania Protein Three-Dimensional Structure. *PLoS One* **2021**, *16* (11), e0259871. https://doi.org/10.1371/journal.pone.0259871.
109. *WHO Interim Guidelines for the Treatment of Gambiense Human African Trypanosomiasis*; WHO Guidelines Approved by the Guidelines Review Committee; World Health Organization: Geneva, 2019.
110. Wijngaert, B. D.; Sultana, S.; Singh, A.; Dharia, C.; Vanbuel, H.; Shen, J.; Vasilchuk, D.; Martinez, S. E.; Kandiah, E.; Patel, S. S.; Das, K. Cryo-EM Structures Reveal Transcription Initiation Steps by Yeast Mitochondrial RNA Polymerase. *Mol Cell* **2021**, *81* (2), 268-280.e5. DOI:10.1016/j.molcel.2020.11.016
111. Wirtz, E.; Leal, S.; Ochatt, C.; Cross, GeorgeA. M. A Tightly Regulated Inducible Expression System for Conditional Gene Knock-Outs and Dominant-Negative Genetics in Trypanosoma Brucei. *Molecular and Biochemical Parasitology* **1999**, *99* (1), 89–101. 10.1016/S0166-6851(99)00002-X. DOI:10.1016/s0166-6851(99)00002-x.
112. Yang, W. Nucleases: Diversity of Structure, Function and Mechanism. *Q Rev Biophys* **2011**, *44* (1). DOI:10.1017/S0033583510000181.
113. Yu, M.; Muteti, C.; Ogugo, M.; Ritchie, W. A.; Raper, J.; Kemp, S. Cloning of the African Indigenous Cattle Breed Kenyan Boran. *Anim Genet* **2016**, *47* (4), 510–511. https://doi.org/10.1111/age.12441.