Study of DNA double strand break repair proteins of *Toxoplasma gondii*: an implication for inefficient gene targeting

A Thesis submitted to the University of Hyderabad for the award of a Ph.D degree in Department of Biochemistry, School of Life Sciences

By

G. SITA SWATI

(07LBPH05)

Under the Supervision of

Dr. Mrinal Kanti Bhattacharyya

Department of Biochemistry
School of Life Sciences
University of Hyderabad
(P.O.) Central University, Gachibowli,
Hyderabad – 500 046
Andhra Pradesh (India)

October 2012

University of Hyderabad
School of Life Sciences
Department of Biochemistry

DECLARATION

I, G. Sita Swati, hereby declare that this thesis entitled, **"Study of DNA double strand break repair proteins of *Toxoplasma gondii*: an implication for inefficient gene targeting"** submitted by me under the guidance and supervision of Dr. Mrinal Kanti Bhattacharyya, is an original and independent research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

G. Sita Swati
(Research Scholar)
Reg. No. 07LBPH05

Dr. Mrinal Kanti Bhattacharyya
(Research Supervisor)

University of Hyderabad
School of Life Sciences
Department of Biochemistry

CERTIFICATE

This is to certify that this thesis entitled, **"Study of DNA double strand break repair proteins of *Toxoplasma gondii*: an implication for inefficient gene targeting"** is a record of bonafide work done by G. Sita Swati, a research scholar for Ph.D. programme in Department of Biochemistry, School of Life Sciences, University of Hyderabad under my guidance and supervision. The thesis has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Dr. Mrinal Kanti Bhattacharyya
(Research Supervisor)

Head of the Department

Dean of the School

Dedicated to my beloved parents and husband

Acknowledgements

I specially thank my research supervisor, Dr. Mrinal Kanti Bhattacharyya for giving me an opportunity to pursue research in his lab. It was a privilege to work in his lab. My heartfelt thanks for his excellent scientific guidance, valuable suggestions and cooperation given through out my work. I am very much grateful for all his help and patience throughout my Ph.D. He has been very patient, encouraging and helpful throughout my research work.

I thank my doctoral committee members Prof. K.V.A. Ramaiah and Dr. Krishnaveni Mishra for all their suggestions and guidance in my work.

I am very much thankful to Dr. Sunanda Bhattacharyya, Department of Biotechnology, for her guidance in my research work through useful discussions and data interpretation and for extending her lab facilities during my work.

I thank the Head of Biochemistry department, Prof. O.H. Setty and the previous heads, Prof. K.V.A. Ramaiah and Prof. M. Ramanadham for giving all the research facilities.

I thank the Dean of life sciences, Prof. M. Ramanadham and the previous Dean Prof. A. S. Raghavendra for giving all the research facilities which facilitated my research.

I thank CSIR for all its financial support throughout my research.

I thank DBT and CSIR funding agencies for the financial assistance.

I thank DST-FIST, DBT-CREBB and UGC funding to school and departmental facilities that aided in my research work.

I thank Ms. Monika of proteomic facility who helped in the MALDI analysis.

I thank my lab mates, Ms. Nabamita, Mr. Sugith, Ms. Shalu, Ms. Shyamasree, Mr. Suresh, Ms. Nidhi and Ms. Tanvi for all their support and cooperation during my research work.

I thank my previous lab member Mr. Rakesh for helping in my protein work.

I thank our technical assistant Mr. Joseph Thambi and attender Mr. Ramesh for their help in lab.

I thank all the teaching and nonteaching staff members of School of Life Sciences for all their help and support given to me during my research work.

I thank my colleagues of School of life sciences who have helped me in completion of the research work.

I thank my classmates of Advanced PG Diploma in Bioinformatics for all their support.

I thank my friends Mr. V. Srinivas, Ms. Sumera, Dr. Usha and Mrs. Shobha for all their emotional and moral support given at the frustating moments of my Ph.D. work.

I specially thank my father, Mr. G. V. J. Raju and mother Mrs. G. Lakshmi who are the pillars behind me for this completion. Their constant encouragement and

timely advice greatly helped me.

I specially thank my mother-in-law, A. Padmaja whose goodwill and support immensely helped me to complete my research.

I specially thank my sister, Mrs.Y. Sunita and brother-in-law, Mr. Y. Giridhar and Baby Sai Snehita whose goodwill and understanding greatly supported my research work.

Last but not the least, I specially thank my husband, Mr. Eshwar Sharma for all his emotional support, understanding and patience till my research work completion without which this success was impossible.

Above all, I sincerely thank the almighty whose blessings and goodwill made me complete this research work. He alone knows all the hard work and anxious moments that I have undergone during this period.

G. Sita Swati

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

aa	: Amino acid
ADH4	: Alcohol dehydrogenase 4
ADP	: Adenosine di-phosphate
A-NHEJ	: Alternate NHEJ
APN1	: Apurinic/apyrimidinic endonuclease-1
ATP	: Adenosine tri phosphate
BER	: Base excision repair
BIR	: Break induced replication
B-NHEJ	: Backup NHEJ
bp	: Base pair
BSA	: Bovine serum albumin
cDNA	: complementary DNA
C-NHEJ	: Classical NHEJ
Da	: Dalton
DDR	: DNA damage response
DEPC	: Diethyl pyrocarbonate
DNA	: Deoxyribonucleic acid
DNA-PKcs	: DNA protein kinase catalytic subunit
DSB	: Double strand break
dsDNA	: Double stranded DNA
DTT	: Dithiothreitol
EDTA	: Ethylene diamine tetraacetic acid

FEN-1 : Flap endonuclease-1

GC : Gene conversion

GPD : Glyceraldehyde 3-phosphate dehydrogenase

HDR : Homology- mediated DNA repair

HR : Homologous recombination

HXGPRT : Hypoxanthine-xanthine-guanine-phosphoribosyl transferase

IPTG : Isopropyl β-D-1 thiogalactopyranoside

k_{cat} : Catalytic constant

KCl : Potassium chloride

kDa : Kilo Dalton

K_m : Michaelis-menten constant

LB : Luria-bertani broth

MALDI-TOF : Matrix-assisted laser desorption ionization-time of flight

MESG : 2-amino 6-mercapto 7-methyl purine riboside

$MgCl_2$: Magnesium chloride

Min : Minute

MLH : Mut L homolog

mM : Milli molar

MMEJ : Microhomology mediated end joining

MMR : Mismatch repair

MMS : Methyl methane sulfonate

MRE11 : Meiotic recombination 11

MRN : Mre11-Rad50-Nbs1

mRNA : Messenger RNA

MRX : Mre11-Rad50-Xrs2

MS : Mass spectrometry

MSH : Mut S homolog

Mwt : Molecular weight

NaCl : Sodium chloride

NaH_2PO_4 : Sodium dihydrogen phosphate

NaOH : Sodium hydroxide

NBS1 : Nijmejen breakage syndrome 1

NEF-1 : Nucleotide excision repair factor-1

NER : Nucleotide excision repair

NHEJ : Non-homologous end joining

Ni-NTA : Nickle nitriloacetic acid

OD : Optical density

ORF : Open reading frame

PARP-1 : Poly ADP ribose polymerase-1

PCIAA : Phenyl chloroform isoamyl alcohol

PCNA : Proliferating cell nuclear antigen

PCR : Polymerase chain reaction

PEG : Polyethylene glycol

PMS-1 : Post meiotic segregation increased-1

PNK : Poly nucleotide kinase

PNP : Purine nucleoside phosphorylase

Pol	: Polymerase
PVDF	: Polyvinylidene Fluoride
RAD50	: Radiation sensitive 50
RFC	: Replication factor C
RNA	: Ribonucleic acid
RNase	: Ribonuclease
RPA	: Replication protein A
RT-PCR	: Reverse transcriptase - polymerase chain reaction
SC	: Synthetic complete
SDS	: Sodium dodecyl sulphate
SDSA	: Synthesis dependent strand annealing
SDS-PAGE	: Sodium dodecyl sulphate poly acrylamide gel electrophoresis
Sec	: Second
SSA	: Single stranded annealing
ssDNA	: Single stranded DNA
TBE	: Tris Borate EDTA
TCA	: Trichloroacetic acid
TE	: Tris EDTA
Tris Cl	: Tris chloride
WT	: Wildtype
XLF	: XRCC4 like factor
XRCC4	: X-Ray cross complementing factor 4
XRS2	: X-Ray sensitive 2

YPD	: Yeast extract peptone dextrose
μg	: Microgram
μM	: Micromolar

CHAPTER 1

INTRODUCTION

INTRODUCTION

1.1 *Toxoplasma gondii*

Toxoplasma gondii is a widespread obligate protozoan parasite of genus *Toxoplasma* which infects humans and animals, causing Toxoplasmosis. It was first identified in the tissues of birds and mammals. It infects about one-third world population. It is an intracellular apicomplexan parasite capable of infecting humans. It may lead to congenital birth defects and encephalitis in immuno-suppressed individuals (Remington JS et al., 1989; Luft BJ et al., 1992). The parasite can be transmitted by vertical transmission of the rapidly growing tachyzoites. It is mostly acquired in humans by the oral ingestion of the tissue cysts containing bradyzoites. Ingestion of oocysts containing sporozoites which are the products of the sexual cycle in cat intestines also results in transmission. Once cysts or oocysts are ingested, they divide rapidly to form tachyzoites. These differentiate into latent bradyzoites (reviewed in Kim K and Weiss LM, 2008). Type I strains are highly virulent and result in acute outbreaks. Both GT1 and RH strain belong to this class. Type II strain, ME49 which was first *T. gondii* strain sequenced, is comparatively less virulent in mice. Type III strain, VEG strain is a type III strain which is also less virulent in mice (reviewed in Kim K and Weiss LM, 2008).

The parasite genome is 63 Mb in size and consists of 14 chromosomes (Kissinger JC et al., 2003). This protozoan parasite has a number of structural similarities to *Plasmodium, Cryptosporidium* etc. In addition to the importance of *T. gondii* as a human and veterinary pathogen, the parasite accessibility to experimental manipulation made *Toxoplasma* as a useful model for the study of other disease causing microbial pathogens and apicomplexans like *Babesia*, *Cryptosporidium*, *Eimeria*, *Neospora*, *Plasmodium*, *Sarcocystis*, *Theileria* etc. It has thus become a

model organism for genetic studies. Owing to high degree of non-homologous end joining (NHEJ), gene targeting efficiency is very low in this parasite. It is puzzling to know that although NHEJ is prevalent in eukaryotes, no functional NHEJ has been reported in any other protozoan parasite. *Toxoplasma gondii* is the only protozoan parasite which harbors NHEJ. Gene knockouts and allelic replacement is a challenge in *Toxoplasma* (Donald RG et al., 1998). High frequency of non-homologous recombination can be exploited for insertional mutagenesis but at the same time, targeted gene integration is complicated in this parasite (Donald RGK et al., 1995, 1996). In *Ku80* knock out genetic background, an increase of 300 to 400 fold was observed in targeted gene replacement efficiency in *Toxoplasma gondii*. On the other hand, integration of transgenes appears to occur exclusively by homologous recombination in other protozoan parasites, like *Leishmania*, *Trypanosoma* and *Plasmodium* (Kapler GM et al., 1990; Lee MG et al., 1990; Wu Y et al., 1996).

1.2 DNA damage

DNA damage is inevitable. It is caused by ionizing radiation, UV light, replication across a nick, oxidative free radicals, topoisomerase failures, inadvertent enzyme action at fragile sites and mechanical stress (reviewed in Lieber MR., 2010). Figure 1 displays the possible DNA damaging agents. The DNA damage signaling cascade finally reaches the effector molecules which triggers the following cellular processes- DNA repair, cell cycle checkpoint activation, transcriptional program activation, apoptosis and senescence (reviewed in Rupnik A et al., 2010).

DNA damaging agent, a type of alkylating chemical, methyl methane sulphonate (MMS) exposure results in the methylation of DNA bases and generates double strand breaks (DSBs) in a time and dose-dependent manner (Zhou C et al., 2006) upon

collision of progressing replication forks with SSB intermediates of base excision repair (Pascucci B et al., 2005; Wyatt M D et al., 2006).

1.3 DNA repair

DNA damage can occur in single or both the strands of DNA. Damaged bases or nucleotides can be repaired by excision repair mechanisms: base excision repair (BER) or nucleotide excision repair (NER). DNA mismatch is repaired via mismatch repair mechanism (MMR). Double stranded DNA breaks are more severe and difficult to repair. Double stranded DNA break repair mechanisms are of 3 types viz. homologous recombination (HR), non-homologous end joining (NHEJ) and single stranded annealing (SSA). Genomic integrity of cell depends on the fidelity of DNA repair. Figure 2 shows the consequences of the unrepaired breaks in an organism. In unicellular organisms, failure to repair DNA leads to cell death by blocking essential cellular functions; whereas in multicellular organisms, it results in apoptosis. Unrepaired DNA breaks result in genome instability, cancer, neuro-degeneration and other pathologies.

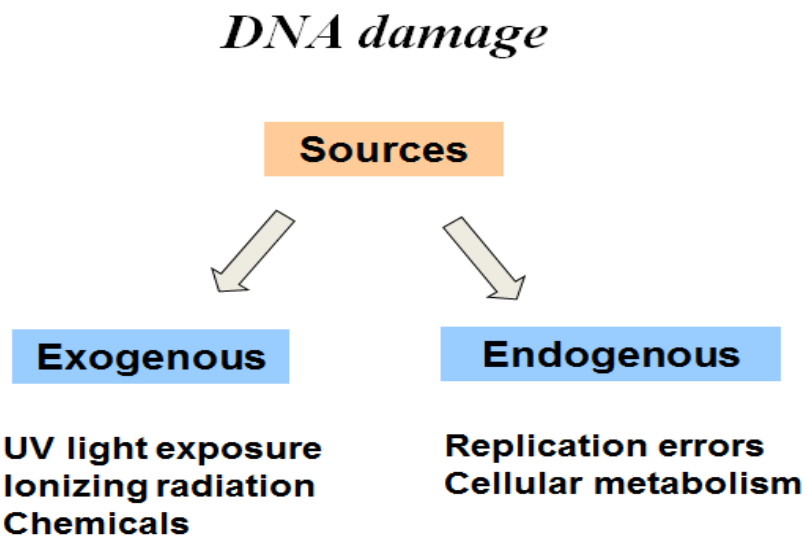

Figure 1. Different sources of DNA damage.

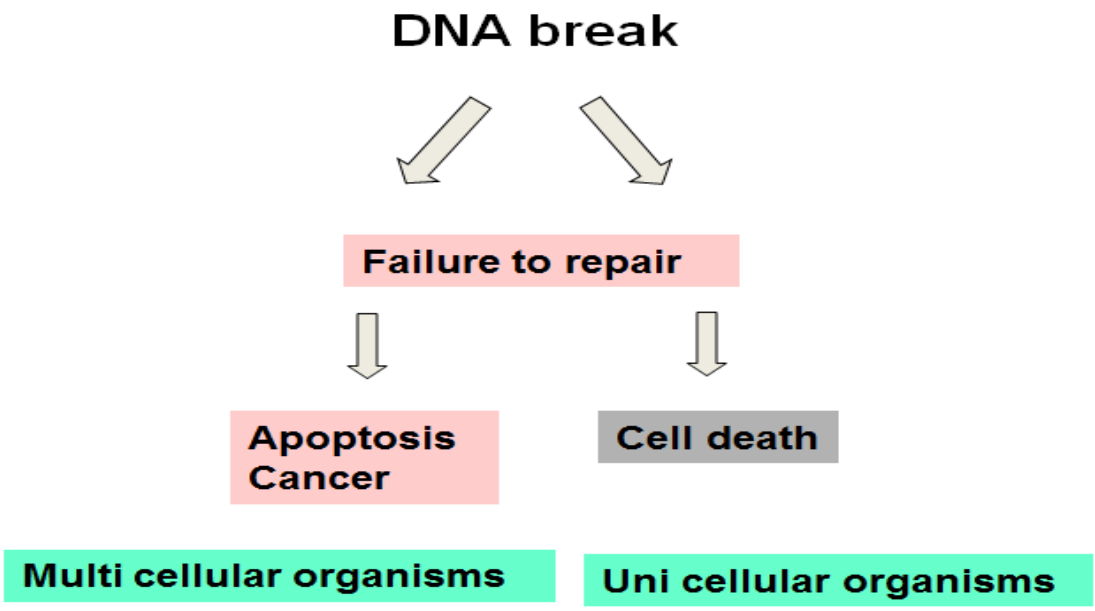

Figure 2. Illustration showing the consequences of unrepaired DNA breaks.

1.4 Repair of damaged or mismatched bases

A1) Base excision repair of DNA (BER)

Base excision repair occurs when the DNA is oxidized, deaminated or alkylated. It is well conserved from prokaryotes to eukaryotes. In general, it is of 2 types- pol β dependent short (1 nucleotide) patch and PCNA dependent long (10-12 nucleotides) patch BER.

In general, DNA glycosylases first recognize and excise the damaged base from the sugar phosphate backbone of DNA by cleaving N-glycosidic bond between the target base and deoxyribose. This creates an abasic site (AP-apurinic/apyrimidinic site). A group of enzymes called AP endonucleases recognize the abasic site and make an incision at the 5′ or 3′ phosphodiester bond of the AP site. This generates a nucleotide gap. These enzymes belong to 2 categories viz. class I and class II. Flap endonuclease (FEN-1) resolves the single stranded DNA flap intermediates that form during long patch DNA base excision repair. Long patch BER involves DNA polymerase I and DNA ligase I. DNA polymerase I fills in the abasic site and DNA ligase I catalyzes the phosphodiester bond formation.

In yeast, AP site during short patch BER pathway is created by APN1. Rad27 (FEN-1) trims the DNA at the double strand break and polymerase ε synthesizes the new strand and Cdc9 carries out ligation step (Kelley MR et al., 2003). In mammals, APE1 creates AP site. β-pol with DNA ligase I or DNA ligase III/XRCC1 carries out polymerization and nick sealing (Kelley MR et al., 2003).

A2) Base excision repair of DNA (BER) in apicomplexa

Short patch BER eliminates 80% of the damaged bases in most of the eukaryotes. Exceptionally in *Plasmodium falciparum*, BER is predominantly by long patch pathway (Haltiwanger BM et. al., 2000). In *P. falciparum,* AP endonuclease (APN1),

replication factor C (RFC), DNA polymerase I and DNA ligase are involved in BER. Homologue of DNA glycosylase is apparently absent in *Plasmodium falciparum*. But, partial purification and characterization of DNA pol β-like enzyme was done and it was observed that it can repair a stretch of 3-5 nucleotides (Nunthawarasilp P et al., 2007). DNA glycosylase and class I AP endonuclease are apparently absent in the genome of *Plasmodium*. In *P. falciparum*, class II AP endonuclease Mg^{2+} dependent and independent pathways have been reported (Haltiwanger BM et. al., 2000). Mammalian cells contain only Mg^{2+} dependent AP endonuclease (HAP1, ExoIII) which indicates that the malarial parasite Mg^{2+} independent AP endonuclease might be an important chemotherapeutic target (Haltiwanger BM et. al., 2000). The PCNA-binding domain with amino acid residues 350-359 is highly conserved in *Plasmodium* (Casta LJ et al., 2007). PfFEN-1 showed endonuclease activity which was stimulated by divalent cations (Mg^{2+} and Mn^{2+}) and inhibited by monovalent ions (Casta LJ et al., 2007). The PCNA (proliferating cell nuclear antigen) binding site is internally located in the C-domain of PfFEN-1, rather than the extreme C-terminal end as seen in FEN-1 from other organisms. PfLig I has a variable N-terminal and a constant C-terminal end. DNA binding domain is present towards the inner side of N-terminal end (Buguliskis JD et al., 2007). Onyango DO et al., 2011 reported that *Toxoplasma gondii* possesses two apurinic/apyrimdinic (AP) endonucleases, TgAPE and TgAPN that function in DNA BER. The authors reported that the knockout of TgAPN was not possible and TgAPN is critical for the parasite to recover from the DNA damage. *Toxoplasma* genome database revealed the presence of homologues of DNA glycosylase, FEN-1, PCNA, DNA ligase I and DNA polymerase I.

B1) Nucleotide excision repair

It occurs when the pyrimidine dimers are formed in DNA due to the exposure to UV light and/or environmental carcinogens. In general, UvrA protein identifies the pyrimidine dimers due to UV irradiation. Two subunits of UvrA with one subunit of UvrB binds to the DNA and unwinds it around the damaged site. UvrA subunits leave; UvrB forms a stable complex with the DNA and UvrC subunit is recruited. UvrB nicks on one side and UvrC on the other side of the DNA lesion.

In yeast, it was shown that there was no preformed repairosome complex for damage recognition and dual incision of DNA lesion but the sequential assembly of proteins takes place. Rad14p recognizes the damaged DNA and forms a ternary complex with Rad1/Rad10 proteins (Guzder SN et al., 1993). It shows strong association with Rad1 but weakly associated with Rad10 protein. This ternary protein complex is referred to as *nucleotide excision repair factor or designated as NEF-1*. Rad1/10 proteins show endonuclease activity (Sung P et al., 1993; Tomkinson AE et al., 1993). Rad3 shows ATP dependent binding to the UV damaged DNA (Sung P et al., 1994) and it associates with Rad25 resulting in unwinding of the damaged duplex DNA. This facilitates the RPA to bind to the single stranded damaged DNA. Rad2 shows endonuclease activity (Habraken Y et al., 1993). Rad2 interacts with TFIIH and RPA and makes dual incision in the damaged DNA strand. In humans, genes encoding RPA, TFIIH, XPA, XPC, PCNA, RFC and ERCC1 are present. PARP1 is involved in single strand break repair mechanisms, BER or NER, together with XRCC1, DNA ligase III, as well as polynucleotide kinase (PNK), PCNA and FEN1 (Okano S et al., 2003; Frouin I et al., 2003).

B2) Nucleotide excision repair in apicomplexa

Interestingly, the homologues of UvrA, UvrB, UvrC, XPD, XPE and XPF proteins seem to be absent in *Plasmodium falciparum* and *Toxoplasma gondii*. The genome of

P. falciparum and *T. gondii* exhibited homologues of UvrD, XPA, XPB, XPC, XPG and CSA.

C1) Mismatch repair

It is a system for recognizing and repairing the erroneous insertion, deletion and misincorporation of bases that can arise during DNA recombination and DNA replication. It is also conserved from prokaryotes to eukaryotes. It is strand specific. In eukaryotes, the newly synthesized strand is probably recognized by nascent strand discontinuities during replication (Holmes JJ et al., 1990). MSH2 first recognizes and initiates the mismatch repair process at the site of DNA lesion. MLH1 and PMS1 form a heterodimer complex. DNA polymerase synthesizes the correct bases and DNA ligase I seals the nicks to complete the mismatch repair process. PCNA (proliferating cell nuclear antigen) is a cofactor of DNA polymerase and it increases the activity of it. Single stranded binding protein binds to the unwound DNA.

In yeast, all the six MSH1 to MSH6 proteins are present. MutSα (MSH2/MSH6) and MutSβ (MSH2/MSH3) complexes recognize and bind to the mispaired bases in DNA (Acharya S et al., 1996, Habraken Y et al., 1996). Mlh3 and Pms1 are the MutL homologs in yeast. Three different exonuclease activities have been identified in yeast viz. Exo I shows 5′ to 3′ exonuclease activity; polymerases δ and ε show proofreading activity. PCNA stabilizes the MutSα and MutLα heterodimers at the mismatched sites. In humans, hMSH1 protein has not been reported till date. MutSα (hMSH2 / hMSH6) and MutSβ (hMSH2 / hMSH3) recognize the incorrect bases involved in pairing. MutSα binds to many base/base mismatches and one base pair insertion/deletion loop (IDL), whereas MutSβ binds to 2-4 basepair IDLs (Jiricny J, 1998, Kolodner RD and Marsischky GT, 1999). MutLα (MLH1/PMS2) acts as a

facilitator by coordinating all the events in mismatch repair. MutLβ (MLH1/PMS1) and MutLγ (MLH1/MLH3) complexes are less understood in humans. PCNA is required at the excision and resynthesis steps (Umar A et al., 1996, Gu L et al., 1998). EXOI/HEXI is the human homolog of Exo I. It interacts with MSH2.

C2) Mismatch repair in apicomplexa

MSH2 homologue is present in *Plasmodium falciparum.* MutH homolog is absent in *Plasmodium falciparum*. PfMLH has been cloned and biochemically characterized from *Plasmodium falciparum* 3D7 strain (Tarique M et al., 2012). It contains endonuclease and ssDNA-dependent ATPase activity and is expressed in schizont stage at a peak level in *P. falciparum* 3D7 and Dd2 strains (Tarique M et al., 2012). Bethke L et al., 2007 research group disrupted *P. berghei* MSH2-2 gene and found that this gene is not essential for parasite growth in both asexual and sexual life cycle. They identified an unusual complement of two mismatch repair genes, PfMSH2-1 and PfMSH2-2. Moreover PfMSH2-2 showed an elevation in mutation frequency and it was indirectly demonstrated that this gene played a role in mismatch repair. Antimalarial drug resistance can result from defective mismatch DNA repair (Castellini MA et al., 2011). Exo I with 5′ to 3′ exonuclease activity is present in the malarial parasite. In *Toxoplasma*, homologue of MutS DNA damage repair enzyme (TgMSH-1) is present. Lavine MD and Arrizabalaga G, 2011 reported that the antibiotic monensin caused cell death of *Toxoplasma gondii* through an unknown MSH-1 dependent cell cycle disruption.

1.5 Double strand break repair mechanisms

DNA double strand breaks (DSB) are mainly repaired either by homologous recombination (HR) pathway or by non-homologous end joining (NHEJ) mechanism.

When a DSB is formed, HR repair depends on searching of extensive homologous stretches of DNA; whereas NHEJ of broken DNA ends depends on little or no homology (Figure 3). Both pathways compete with each other for repair. The whole process of cell's decision making on which pathway to proceed for repair is not yet known. The double strand break repair pathways in *E. coli* and *S. cerevisiae* have been used as the model systems to understand the repair pathways in other organisms. Interestingly, the usage frequency of HR over NHEJ is different in different organisms (Figure 4). In higher eukaryotes, like in humans, NHEJ is predominant whereas in lower eukaryotes, like in *Saccharomyces cerevisiae*, HR is the major pathway. Interestingly, two lower eukaryotes, very closely related apicomplexan parasites; *Plasmodium falciparum* and *Toxoplasma gondii* show strikingly opposite choice of the repair pathway. *P. falciparum* apparently lacks NHEJ (based on the genome information) and shows HR and *T. gondii* uses NHEJ as the predominant pathway. *Plasmodium* is dependent on only homologous recombination in case of a double strand break repair in DNA. From the genome information, the NHEJ and SSA proteins are apparently absent in the malarial parasite (Gardner MJ et. al., 2002). In *T. gondii*, both HR and NHEJ pathways exist. This parasite surprisingly chooses NHEJ as its major repair mechanism (Fox BA et al., 2009; Huynh MH et al., 2009).

If Ku70/80 heterodimeric protein binds to DSB, the repair takes place by NHEJ. But if the ends are resected by 5' to 3' exonuclease, it forms long 3' single stranded DNA where in HR mediated repair occurs (Figure 5). In yeast, it was shown that ExoI and/or Sgs1 generate long stretches of ssDNA that are essential for HR (Mimitou EP and Symington LS, 2008). Chung W et al., 2010 demonstrated that about 2 to 10 Kb of ssDNA is generated depending on the kinetics of repair in yeast. Long ssDNA overhang inhibits recruitment of Ku protein and in turn incorporates recombination

proteins. The first step in DNA resection where in the intermediate 3' tail is formed requires Mre11-Rad50-Xrs2 complex and Sae2, while the second step where the early intermediate is rapidly processed to generate the extensive 3' overhang employs exonuclease Exo1 and/or Sgs1 (Mimitou EP and Symington LS, 2008). In eukaryotes, within S and G2 phases of cell cycle, homologous recombination is very active because the two sister chromatids are directly adjacent and further, NHEJ is active in G1 phase of the cell cycle (reviewed in Lieber MR, 2010).

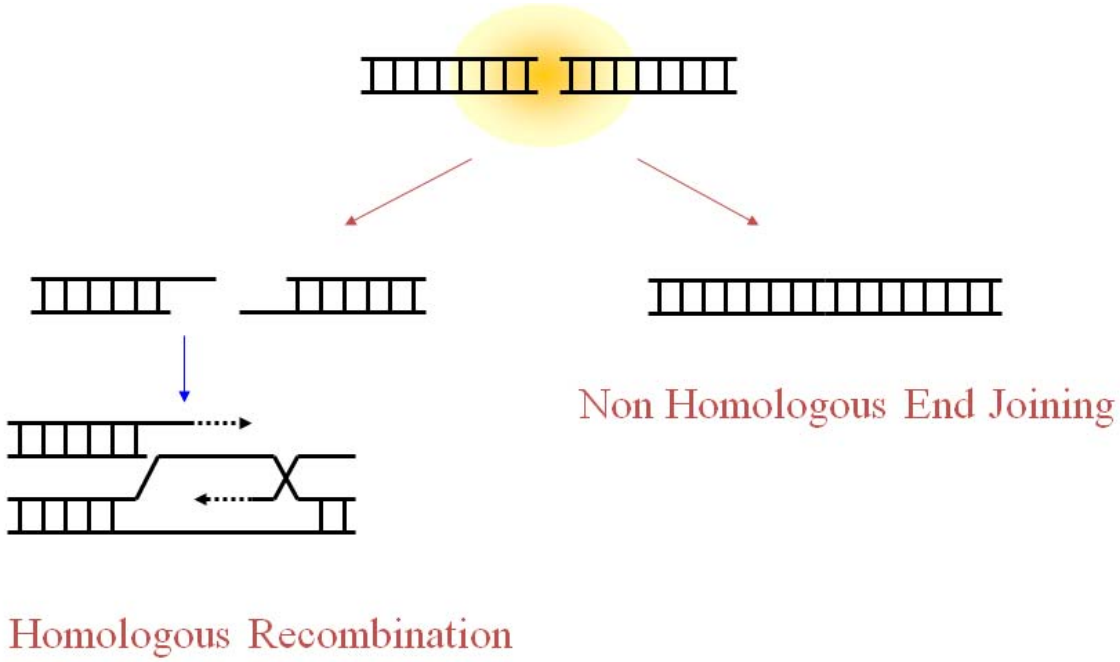

Figure 3. Illustration showing the two main DNA double strand break repair pathways: Homologous recombination (HR) and Non homologous end joining (NHEJ).

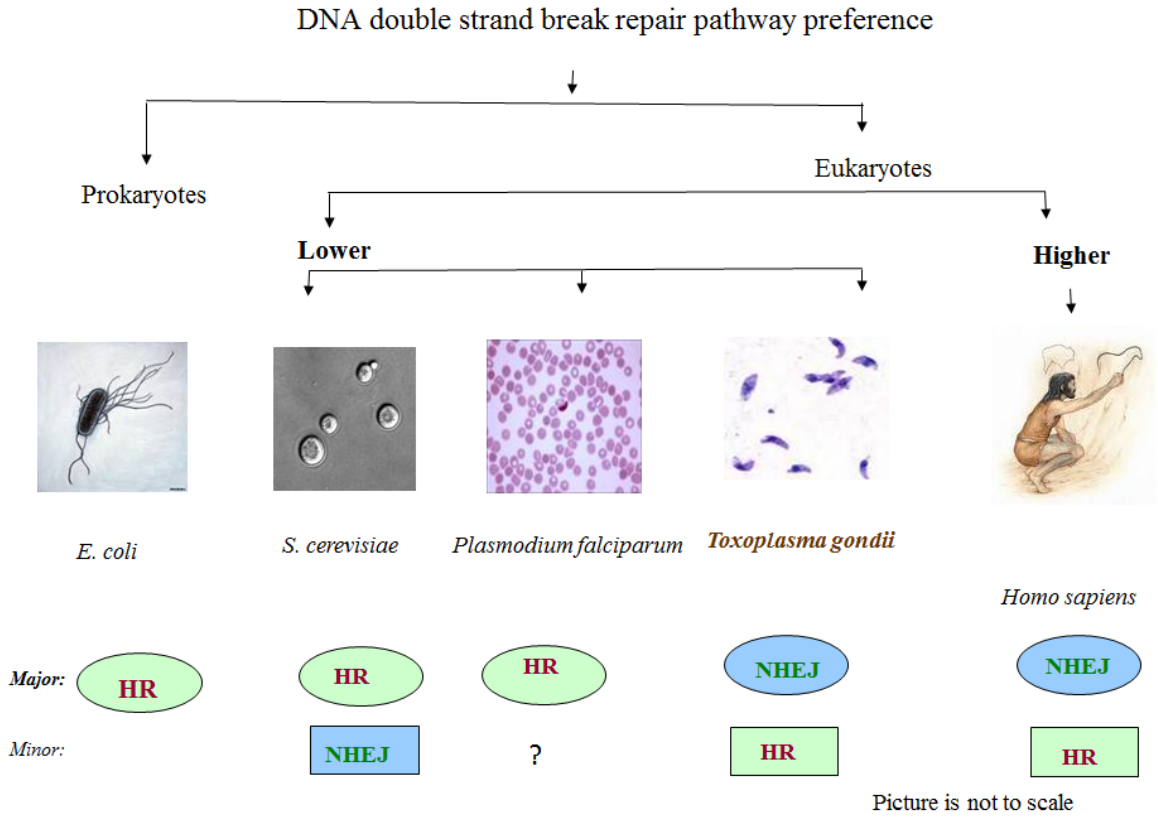

Figure 4. Illustration shows the DNA double strand break repair pathway preference in prokaryotes and eukaryotes.

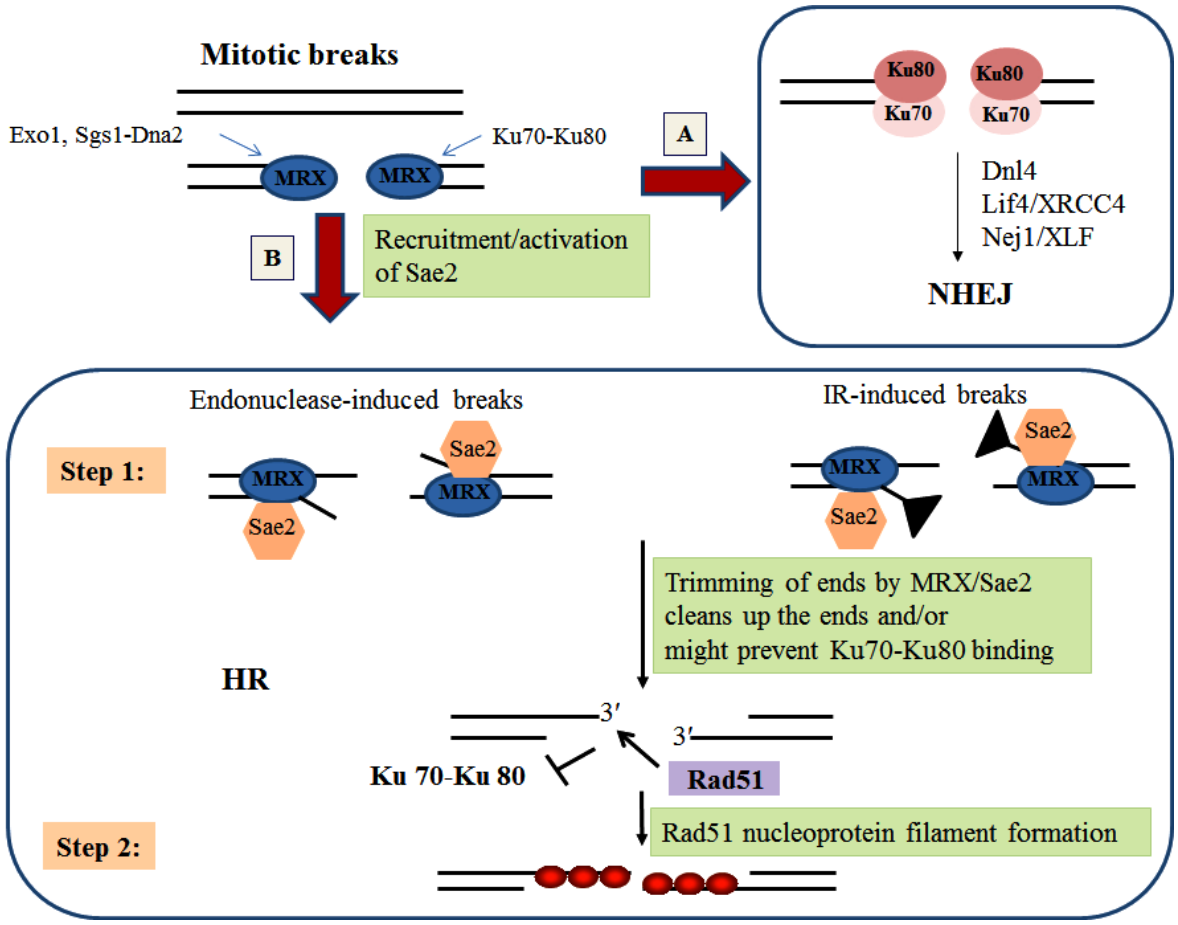

Figure 5. Choice between HR and NHEJ pathways [Modified from Mimitou EP and Symington LS. DNA repair, 2009].

In DNA repair, Mre11 (Meiotic recombination 11 protein) is recruited very early to the broken ends. Mre11p is a multifunctional player and has several roles like recognition of DNA damage, binds to the broken chromosomal ends and starts the repair process, stabilizes replisome as well as activates the DNA damage checkpoint proteins. It exists as a MRX complex which is made up of two molecules each of Mre11, Rad50 (Radiation sensitive 50) and one molecule of Xrs2 (X-ray sensitive 2) protein. The human homolog of Xrs2 is Nbs1 (Nijmegen breakage syndrome 1). Mre11 protein is known to dimerize. The Mre11 is core member of the complex and interacts with itself and both Rad50 and Xrs2/Nbs1. Mre11p resects the broken DNA ends in 5′ to 3′ direction and tethers the broken ends in the form of MRX complex. It has 3′ to 5′ exonuclease and endonuclease activities that aids in processing of DNA ends during DNA repair and activation of cell checkpoint proteins (Jazayeri A et al., 2008; Lee JH and Paull TT, 2005; Paull TT and Gellert M, 1998). Studies from the *Xenopus laevis* extracts revealed that DNA resection results in production of ssDNA oligos that associate with the MRN complex and influence cell signaling protein, ATM activity (Jazayeri A et al, 2008). Mre11 binds duplexed DNA (Hopfner KP et al., 2002) and to the ends of the linearized DNA (de Jager M et al., 2001 and Usui T et al., 1998). Once recruited to the site of DNA damage, ATM physically interacts with Mre11 and this interaction stimulates ATM to phosphorylate the downstream target substrates (Bakkenist CJ and Kastan MB., 2003; Lee JH and Paull TT., 2004). Mre11p is also known to play a role in telomeric end processing (Haber JE et al., 1995, Deng Y et al., 2009). The work done by Williams B et al., 2005 indicate that the A and B DNA binding domains of MRE11 repress telomere rapid deletion. Recently, it was proposed that MRX or MR complex disrupts the G rich quadruplex telomeric DNA in the presence of ATPase activity of Rad50. This generates the

telomeric ends which are free for the recruitment of telomerase (Ghosal G and Muniyappa K, 2007). Mre11 nuclease and C-terminal DNA damage response functions are required for initiating yeast telomere healing (Bhattacharyya MK et al., 2008). Tittel-Elmer M et al., 2009 investigated that the MRX complex of yeast is recruited to replication forks during HU-induced pausing and it stabilizes the replisome independent of the S-phase checkpoint and Mre11 nuclease activity. This study revealed the molecular function of MRX in DNA replication. In higher eukaryotes, null mutations in components of Mre11 complex proved lethal (Luo G et al., 1999; Yamaguchi-Iwai Y et al., 1999; Zhu J et al., 2001).

A1) Homologous recombination

Homologous recombination is a process in which the genetic material is exchanged between two similar or identical strands of DNA. This repair mechanism is operational when the double strand break occurs in one of the homologous chromosomes. The main proteins involved are Rad50, Rad51, Rad52, Rad54, Rad55 and RPA (replication protein A). Rad51 is one of the key proteins involved in the HR process. It is produced in larger amounts so that it immediately binds to the DNA double strand break much before other repair proteins. Repair by HR involves 5' to 3' resection of the broken DNA ends, and generates 3' single stranded DNA ends. These ends get coated by Rad51 to form the DNA-Rad51 filament. This filament searches for and base pairs with a homologous donor, forming a three stranded structure called displacement loop or D-loop. When both the ends of the broken DNA share homology with a sister chromatid, the repair takes place by gene conversion (GC). It involves the synthesis of short patch of new DNA. When homology is present with only one of the DNA DSB ends, the repair occurs by break-induced replication (BIR). The

presence of only one end can arise from stalled and/or broken replication forks (Haber JE et al., 1999) or at the telomeres that have lost their end protection (Lundblad V and Blackburn EH., 1993). When homology at both ends is present, BIR is outcompeted by GC (Malkova A et al., 2005); however what actually inhibits BIR is not clear. A DSB flanked by direct repeats can be repaired by another recombination pathway, single stranded annealing (SSA). The homologous sequences simply anneal with each other, the non-complementary tails are removed, new DNA synthesized to fill the single-stranded gaps and annealing occurs. GC, BIR and SSA are kinetically and mechanistically quite different from each other. Jain S et al., 2009 demonstrated that the initial steps of strand pairing events are similar in these three pathways. Their data suggested that a "recombination execution checkpoint" (REC) regulates the pathway choice. REC is sensitive to the location, orientation and distance between the homologous sequences used for the DSB repair.

RPA of yeast binds to DSB and is replaced by RecA homologue, Rad51. Moreover, RPA helps Rad51 in strand invasion and facilitates Rad51 binding to ssDNA and possibly by displacing dsDNA (Wang X et al., 2004). Strand invasion is the central step of homologous recombination which requires Rad51 protein. The recombinase activity of Rad51 takes place in three different phases- presynaptic, synaptic and postsynaptic phases. *RAD52* gene family is grouped into three classes- class I consists of *RAD50, MRE11* and *XRS2* genes (involved in processing of breaks); class II consists of *RAD51, -52* and *-54* (plays a role in homology search and DNA pairing) and class III consists of *RAD55* and *-57* genes (part of protein complex). Knockout of Rad51 showed embryonic lethality in mice (Tsuzuki T et al., 1996). This suggests that *RAD51* is indispensable for the homology mediated DNA repair (HDR) pathway. It interacts with RPA, Rad51, Rad54 and Rad55 proteins. Rad51 self assembles on the

single and double stranded DNA to form the nucleoprotein filament. ATP binding is required for the nucleofilament formation. It forms helical filament similar to RecA. Replication protein A (RPA) bound ssDNA is recognized by Rad52 and Rad51-Rad52 complex replaces RPA. Rad51 with its motor activity searches for homologous sequences between ssDNA and dsDNA. Rad55-Rad57 heterodimer binds to the ssDNA and shows recombination mediator activity. Rad54 acts as a translocase and DNA helicase in this recombination pathway. Rad51 binding to ssDNA requires Rad52p; the absence of Rad55 delays Rad51 binding and further Rad54 is required at or before the initiation of DNA synthesis after synapsis occurs (Sugawara N et al., 2003). The kinetics of loading of Rad51 recombinase to the single HO-induced double strand break, DNA resection, homology search, strand invasion and initiation of repair DNA synthesis at a time point of every 10 min of homologous recombination during MATa switching in *Saccharomyces cerevisiae* was elucidated by Hicks WM et al., 2011.

Benson FE et al., 1994 research group first identified human Rad51. They reported that the purified HsRad51 protein binds to both ssDNA and dsDNA and shows DNA-dependent ATPase activity. Compared to RecA, it has got low ATP hydrolysis activity. In humans**,** HsRad51 forms a nucleoprotein filament with the DNA to be repaired. In humans, there is no evidence that RPA effects HsRad51 (Gupta RC et al., 1997). Two other Rad51 related proteins viz. Xrcc2 and Xrcc3 are known to express in human cells. Both of these interact with Rad51 (Liu N et al., 1998; Pierce AJ et al., 1999).

For DNA strand exchange, DNA binding ability by *E.coli* RecA is essential and there is no requirement for ATP hydrolysis. An allosteric effector consisting of ADP and Aluminium fluoride (ADP-AlF^-_4) covalent complex can induce the binding of RecA

to DNA and can promote strand exchange up to 800 to 900 base pairs. The needed free energy for strand exchange is supplied from binding of ligand ATP to DNA-protein RecA complex and the ATP hydrolysis destroys the effector ligand (Kowalczykowski SC., 1995). Zaitseva EM et al., 1999 reported on the DNA binding properties of *Saccharomyces cerevisiae* Rad51. This protein binds to both the single stranded and double stranded DNA in an ATP and Mg^{+2} dependent manner and one protein monomer binds to three or four nucleotides. They found that ScRad51 binds to ssDNA even in absence of ATP with the stoichiometry of one protein monomer for every seven to nine nucleotides. These two binding modes are not interconvertible and this behavior is similar to RecA protein. ScRad51 showed ssDNA dependent ATPase activity and strand exchange activity (Sung P., 1994). It was shown that the low efficiency of strand exchange activity of ScRad51 in *in vitro* experiments was due to the enhanced DNA binding ability of this protein.

Site-directed mutagenesis of highly conserved residues (K133R) in ATPase domain of human RAD51 and strand exchange assay revealed that human Rad51 does not hydrolyse ATP for recombinational repair. Weak mutant Rad51-K191R in human cannot hydrolyse ATP but capable of ATP binding and DNA strand exchange (Morrison C et al., 1999). Stark JM et al., 2002 research group generated mutants of hRAD51, hRAD51-K133A and hRAD51-K133R by dominant negative approach which were defective in ATP binding and ATP hydrolysis respectively. Mouse embryonic stem cells that express hRAD51-K133R were obtained which were hypersensitive to mitomycin C and ionizing radiation; showed decrease in spontaneous sister chromatid exchange and a defect in HDR pathway. Stark JM et al., 2004 reported that mammalian HDR and SSA pathways are inversely related. Expression of RAD51 ATP binding defective mutant drastically increased SSA and

Rad52$^{-/-}$ mouse cells showed no defect in HDR but SSA was decreased. The authors demonstrated that the genetic interplay of DSB repair factors is essential to limit the mutagenic potential of the repair process. Agmon N et al., 2009 showed that intrachromosomal homologous chromosomes are always preferred as donors for DSB repair in yeast. In yeast, different protein-protein interactions have been identified viz. Rad51p-Rad52p (Shinohara A et al., 1992), Rad51p-Rad54p (Clever B et al., 1997), Rad51p-Rad55p, Rad55p-Rad57p (Hays SL et al., 1995; Johnson RD and Symington LS, 1995) and Rad52p-RPA1p (Firmenich AA et al., 1995). In yeast, N-terminal amino acids of Rad54A (1-326 residues) interact with Rad51 and it appears that the C-terminal amino acids of Rad54B (327-898 residues) do not interact with Rad51 (Clever B et al., 1997). Strong interaction between Rad51 and Rad55 was identified. No interaction detected between Rad51 and Rad57 in yeast (Johnson RD and Symington LS, 1995). Rad51p, Rad52p, Rad55p, Rad57p and RPA participate in the formation of the "recombinosome" complex during double strand break repair in yeast (Hays SL et al., 1995; Johnson RD and Symington LS, 1995). Larger subunit of RPA interacts with Rad52p in the recombinosome (Hays SL et al., 1995).

A2) Homologous recombination in apicomplexa

PfRad51 protein was identified, biochemically characterized and its functional role in HR repair was demonstrated (Bhattacharyya MK et al., 2003, 2005). PfRad51 monomer binds to 3 nucleotides. PfRad51 shows 66-75% homology within the catalytic region of the Rad51p of human, yeast and other protozoan parasites viz. *Trypanosoma* and *Leishmania* (Bhattacharyya MK and Kumar N., 2003). ATPase activity of PfRad51 provides the requisite energy for the strand migration and homologous pairing. In *Plasmodium*, Rad54 acts as an accessory protein and is also incorporated after Rad51. Compared to RecA mediated homologous recombination,

PfRad51-PfRad54 show slower ATPase activity and lower strand exchange (Bhattacharyya MK et al., 2005). The purified recombinant PfRad51 protein exhibited three-strand exchange activity in presence of non-hydrolysable ATP-γS which showed ATP binding. However the kinetics of strand exchange was slower in presence of ATP-γS (Bhattacharyya MK et al., 2005). This is in accordance with the behavior of *Saccharomyces cerevisiae* and human Rad51 proteins. In presence of ATP and 10 mM $MgCl_2$, the kinetics of strand exchange reaction was slower for PfRad51 when compared to *E.coli* RecA. The products of the reaction appeared after 10 and 20 min for PfRad51. The efficiency of conversion of linear dsDNA into nicked circular dsDNA was lower in PfRad51 when compared to bacterial RecA (Bhattacharyya MK et al., 2005).

Homologues of Rad52 and RPA are apparently absent in the genome of *P. falciparum*. One possibility is that homologous recombination of *Plasmodium* does not need RPA and Rad52 proteins and another possibility is that the sequence of these 2 proteins is different. From the genome information of *Toxoplasma gondii*, putative orthologues of Rad51and Rad54 are present. This suggests that HR is operational in this parasite.

B) Single stranded annealing

It is operational when the homology is limited to only few bases. Here, large chunks of DNA are eliminated at the double-strand break. SSA can occur with a small homology of 30 bp and can go on with a maximum of 200-400bp. The efficiency of SSA improves with the increase in the homology. Rad52p recognizes and binds to the DNA lesion. It scans the damaged DNA for small stretches of homology. The damaged DNA is cleaved at the 5′ region by the action of FEN-1. This leaves 3′

overhangs. Now, the DNA is resected so that the regions of homology can align with each other. Further endonuclease activity is carried out by Rad1/Rad10 proteins wherein the overlapping flaps are removed. The ultimate step of gap filling and nick sealing is accomplished by DNA polymerase I and DNA ligase IV respectively.

In yeast, Rad52, FEN-1 (Flap endonuclease-1) and Rad1/Rad10 proteins carry out SSA mechanism. There is evidence that the mismatch repair Msh2/Msh3 proteins are involved in cutting the small portions of overlapping DNA (Paques F et al., 1997, Sugawara N et al., 1997). In humans, the proteins operational in this pathway are RAD50, MRE11, NBS1, ERCC1 and XPF. In humans, XPF/ERCC1 perform the function of yeast Rad1/Rad10. It is yet to be investigated if hMsh2/hMsh3 proteins show dual role by involving in SSA of humans. Till now there is no report on SSA in any apicomplexa parasite.

C1) Non-homologous end joining

It is simple joining of the two broken ends of the double stranded DNA. Error incorporation is more in this method as it involves little or no homology in rejoining the broken ends. Small insertions or deletions in DNA are possible. NHEJ is mainly of two types- Classical NHEJ (C-NHEJ) and alternate NHEJ (A-NHEJ). A-NHEJ pathway is poorly understood. It is also called as microhomology-mediated end joining (MMEJ) or backup NHEJ (B-NHEJ).

In yeasts, the NHEJ core proteins are Ku70/80, XRCC4, DNA PKcs and DNA ligase IV. In budding yeast, Mre11p-Rad50p-Xrs2p protein complex shows exonuclease and endonuclease activity. In fission yeast, the absence of Mre11p has little effect. NHEJ is not a very efficient repair pathway in yeast cells. XRCC4 and DNA ligase IV are recruited for the ligation step. Domain analyses of Ku protein showed that both the N-

terminal and C-terminal domains are non-essential for its interaction with XLF protein. Only heterodimeric domain is involved. The last 10 amino acids of the C-terminal domain of XLF protein are crucial for XLF-Ku interaction. Though the N-terminal globular head is the site for interaction of XRCC4, the absence of XLF-Ku interaction reduced the interaction of XLF-XRCC4 (Yano K et al., 2011). In humans, the main proteins involved in classical NHEJ of humans are Ku70/80, XRCC4, DNA PKcs (DNA protein kinase catalytic subunit), MRX complex (Mre11/Rad 50/NBS1 proteins) and DNA ligase IV. There is another Ku-independent NHEJ pathway which is the backup NHEJ. It involves DNA ligase III, PARP1 and histone 1 (Iliakis G, 2009). Ku protein exists as a heterodimer of Ku70 and Ku80 and binds to the damaged DNA. Ku70/80 serves two functions. Firstly, it interacts with the damaged DNA and tethers the two broken ends. Secondly, it serves as a platform for the recruitment of additional NHEJ factors (reviewed in Lieber MR, 2010). Ku heterodimer interacts with the catalytic subunit of DNA-PK (Sipley JD et al., 1995). Artemis is a nuclease which is the substrate for DNA-PK. Further, its activity is modulated by DNA-PK. DNA-PKcs becomes active as a kinase and phosphorylates itself and Artemis. Phosphorylation of Artemis has no functional significance but autophosphorylation results in conformational change in DNA-PKcs resulting in its translocation; thereby exposing the DNA ends at the damaged site to Artemis. Then, Artemis cleans up these ends by nuclease activity and makes them suitable for ligation. Mre11-Rad50-Nbs1 complex is involved in the processing of the 3′ ends. Rad50 has ATPase activity and Nbs1 acts as a sensor for DNA damage response. FEN-1 enzyme processes the 5′ ends. DNA ligase IV is essential for non-homologous end joining (Timson DJ et al., 2000).

In humans, both the pathways can co-exist, although C-NHEJ is the major repair mechanism and widely characterized as a Ku protein dependent pathway. NHEJ constitutes about 70% whereas HR constitutes 30% of the total double strand break repair pathways in humans (Liang F et al., 1998).

Early studies in SV40 DNA substrates tested in monkey cells revealed efficient end joining in the absence of C-NHEJ key factors and this provided evidence for A-NHEJ pathway (Roth DB et al., 1985). It is shown that Ku protein strongly represses alternate end-joining pathway (Bennardo N et al., 2008 and Fattah F et al., 2010). However, lower eukaryotes like *E. coli* use ligase A-dependent alternate end joining method for repairing double strand breaks and for horizontal gene transfer. *In E. coli*, HR is the major repair mechanism but A-NHEJ is seen in a very low proportion (Chayot R et. al., 2010). Aniukwu J et al., 2008 showed the fidelity of NHEJ mediated repair of the two broken DNA ends in *Mycobacterium* by DNA sequencing. Hela cell extracts depleted of Ku and DNA-PKcs rejoined blunt ends and homologous ends with 3' or 5' single stranded protruding ends with similar efficiency but addition of Ku suppressed joining of blunt ends and homologous ends with 3' overhang (Wang H et al., 2003). This study presented the biochemical evidence for the presence of Ku-independent backup NHEJ pathway. Audebert M et al., 2004 found an end joining activity in mammalian cells independent of the DNA-PK/XRCC4-ligase IV complex but that actually required the synapsis activity of PARP1 and ligation activity of XRCC1-ligase III. From the human cell extracts, polynucleotide kinase (PNK) was identified as the 5'-DNA kinase associated with the poly (ADP-ribose) polymerase-1 (PARP1) dependent end-joining pathway (Audebert M et al., 2006). Wang H et al., 2005 identified DNA ligase III as the main component of backup NHEJ pathway in extracts of mammalian cells. Liang L et al., 2008 showed that human DNA ligase I or

III are required in micro-homology mediated end joining (MHEJ), and in contrast DNA ligase IV is not required for MHEJ. Their data indicated that the MHEJ and NHEJ pathways require different ligases. Ligase III functions with XRCC1 as a complex and is regulated by PARP1 (reviewed in Mladenov E and Iliakis G, 2011).

Cheng Q et al., 2011 established in human cell lines that Ku is the main factor that counteracts the mobilization of PARP-1 and MRN to the damaged chromatin and also limits the synthesis of poly (ADP-ribose) (PAR) and ssDNA formation in response to DSBs. They showed that there is a shift to B-NHEJ in case of C-NHEJ defective condition. PARP-1 binds to DNA DSBs in direct competition with Ku and PARP-1 function is essential in backup NHEJ (Wang M et al., 2006).

Couëdel C et al., 2004 studies in mice revealed that the double knockout *Rad54*$^{-/-}$ *ku80*$^{-/-}$ was severely compromised in survival than single knockouts. They were very sensitive to DNA damaging agents and ionizing radiation. It implied that cooperation between two DNA DSB repair pathways was required for survival and genomic integrity in the animal.

C2) Non-homologous end joining in apicomplexa

NHEJ apparently seems to be absent in *Plasmodium falciparum*. The parasitic genome does not give any evidence for the existence of the homologues of NHEJ proteins.

Fox BA et al., 2009 first reported that a significant KU-dependent NHEJ pathway is operational in *Toxoplasma gondii*. Knockout at *KU80* locus disrupts the NHEJ DSB repair pathway. *Toxoplasma* genome revealed the genes encoding *KU70* (50.m03211), *KU80* (583.m05492), DNA ligase IV (TGGT1_073840) and DNA-dependent protein kinase (57.m01765) components of eukaryotic NHEJ (www.toxodb.org). *KU80* knockouts retained normal tachyzoite growth and were

viable in murine infection. They exhibited increased sensitivity to DNA double strand breaks induced by phleomycin or γ-irradiation.

D) DNA double strand break repair in other protozoan parasites

Leishmania major Rad51 exhibited ATPase activity and showed magnesium dependent binding to both single stranded and double stranded DNA. When the *Leishmania* promastigotes were treated with phleomycin, elevated expression of LmRad51 at mRNA and protein level was observed (McKean PG et al., 2001). Kinetoplastids, like *Trypanosoma* and *Leishmania* species, naturally exhibit efficient HR of exogenous DNA (Cruz A et al., 1990; Eid J et al., 1991). *Trypanosoma brucei* Ku70/80 functions in telomere maintenance and regulation of telomerase (Janzen CJ et al., 2004) and Ku dependent NHEJ was not observed in *T. brucei* cell extract (Burton P et al., 2007; Glover L et al., 2008). Burton P et al., 2007 showed from *T. brucei* cell extracts that Ku independent end joining of linear DNA molecules occurs based on sequence microhomology. Further based on the bioinformatic analysis, it was found that XRCC4 and DNA ligase IV homologues may not be encoded by the genome of *T. brucei*, raising the possibility that diverged ligase factors may be present.

1.6 Gene targeting in different organisms

The chinese hamster cells defective in *ku80* were shown to exhibit gene targeting frequencies comparable to that of wildtype cells (Liang F et al., 1996). Further mouse embryonic stem cells defective in *ku70*, *Xrcc4* or *DNA-PKcs* did not exhibit increased gene targeting efficiencies (Pierce AJ et al., 2001; Dominguez-Bendala J et al., 2006). Iiizumi S et al., 2008 examined the frequencies of random and targeted integration in

chicken DT-40 and human Nalm-6 cell lines. Loss of NHEJ led to increased targeted integration in chicken DT-40 cell lines. However, in human Nalm-6 cell lines, the results demonstrated that NHEJ is not the sole mechanism of random integration. Nevertheless, the authors presented evidence that NHEJ inactivation can lead to enhanced gene targeting through a reduction of random integration.

Ninomiya Y et al., 2004 research group identified and disrupted *Neurospora crassa* Ku70 and Ku80 homologues mus-51 and mus-52 genes and studied the frequency of the integration of the exogenous DNA into the genome of this disruption strains. They found that the suppression of NHEJ increases the frequency of homologous recombination. In order to investigate the relationship between the targeting frequency and the length of the homologous sequence, 50, 100, 500 and 1000 base pairs homologous to 5′ and 3′ flanking ends of the mtr gene was replaced by bar gene. The authors concluded that more than 1000 bp homology was required for complete gene replacement in *Neurospora crassa*.

Gene targeting studies in *Plasmodium falciparum* show that the targeted gene reaches right locus. This is in accordance as HR is the main DNA DSB repair mechanism in *P. falciparum*. Owing to very high NHEJ, it is mostly random gene integration in *Toxoplasma gondii* (Fox BA et al., 2009; Huynh MH et al., 2009). However, the efficiency of gene targeting in these organisms is very low. Table 1 shows the accuracy and efficiency of gene targeting in *Plasmodium falciparum* and *Toxoplasma gondii*.

Table 1) The gene integration and the efficiency of gene targeting in apicomplexan parasites:

Organism	Efficiency of gene targeting	Gene integration
Plasmodium falciparum	Very low	accurate
Toxoplasma gondii	Very low	random

1.7 Significance of studying homologous recombination mechanism in *T. gondii:*

Understanding DNA break repair mechanisms can help in eradication of the deadly diseases caused by the pathogens. In an effort to curb Toxoplasmosis in humans, we tried to understand the DNA double strand break repair mechanisms in *Toxoplasma gondii*. To address why *Toxoplasma gondii* chooses NHEJ, we identified and cloned *MRE11*, *KU80* and *RAD51* genes of *T. gondii* and in order to understand the interplay between these three proteins during DSB repair. Understanding the functional role of TgMre11 would give insights on pathway choice as it binds to the broken DNA ends much before the decision is made between HR and NHEJ. To understand NHEJ mechanism, we choose the central player TgKu80 for characterization. Further to understand HR pathway and the targeted gene disruption in this parasite, we choose the key player TgRad51 for our study. We have cloned, purified and characterized TgRad51 biochemically as well as genetically. *T. gondii* and *P. falciparum* show surprisingly opposite choice in DNA repair pathways. *Plasmodium* genome apparently lacks the key components of NHEJ pathway and exhibits high frequency of homologous recombination. Here we report mechanistic insights for differential repair choices between two closely related lower eukaryotes. Though, TgRad51 and PfRad51 have 82% identity in their catalytic domain, they differ markedly in ssDNA dependent ATP hydrolysis activity. Since it is extremely difficult to perform the functional studies in *Toxoplasma*, we choose yeast as a surrogate model system. We hypothesize that the compromised ATPase activity of TgRAD51 leads to inefficient gene targeting and poor gene conversion efficiency in *T. gondii*.

The interplay between genomic investigations and those targeted at chemotherapy is essential for new drug designing to treat Toxoplasmosis. The overall goal of my PhD

research was to understand the DSB repair pathways in *T. gondii* by comparative analysis and gain evolutionary insights into the choice of pathways during DNA DSB repair.

1.8 Objectives of the study

Toxoplasma gondii is a lower eukaryote which exceptionally shows preference to non-homologous end joining during DNA double strand break repair. In contrast, in all other known eukaryotes, homologous recombination is the major DSB repair pathway. *Plasmodium falciparum*, a closely related apicomplexan parasite to *T. gondii* and *Saccharomyces cerevisiae* are the best examples of lower eukaryotes which prefer HR. Recent work done by two research groups (Huynh MH et al., 2009; Fox BA et al., 2009) revealed that the non-homologous end joining pathway is the predominant pathway in *T. gondii*. The authors had shown that only in the *Ku80* (an important protein recruited very early during classical NHEJ repair) knockout background, homologous mediated gene targeting studies were successful and increased by several fold. However, it does not increase the integration efficiency at the correct locus (Fox BA et al., 2009). We wanted to know why *Toxoplasma gondii*, despite being a lower eukaryote chooses NHEJ. In order to investigate this, we choose TgMre11 protein for our study which is recruited within minutes of DNA damage and it binds to the broken ends much before the decision is made between HR and NHEJ. To understand NHEJ mechanism, we choose the central player TgKu80 for characterization. Further to understand HR pathway in this parasite, we choose the key player TgRad51 for our study. Inefficient gene targeting can be due to weak ATPase activity which results in inefficient motor activity and poor homology searching. We speculated that the ATPase activity of *Toxoplasma gondii* Rad51 was weak which led to inefficient gene targeting and poor gene conversion efficiency in *Toxoplasma gondii.*

Questions addressed

1) Why *Toxoplasma gondii*, despite being a lower eukaryote chooses NHEJ over HR?
2) What determines the pathway choice in this organism?
3) How does the cell enable the chosen pathway once the decision is made?

Aims of the Ph.D work

i. Why *Toxoplasma gondii,* despite being a lower eukaryote chooses NHEJ over HR? In this regard, we have chosen three proteins TgMre11, TgKu80 and TgRad51 for characterization. We wanted to understand the interplay between these three proteins.

ii. Why is the gene targeting less efficient in *Toxoplasma gondii*? Is it because of less efficient TgRad51 protein. Our approach was to clone, express and purify TgRad51 protein and we tried to look at its ATPase activity.

CHAPTER 2

MATERIALS AND METHODS

2.1 Yeast Methods

2.1.1 Preparation of yeast competent cells

A single colony from plate was inoculated in 5 ml of medium and incubated at 30°C, 200 rpm for overnight. After 1:10 dilution, the overnight culture OD_{600} was measured. The volume of secondary inoculum in 40 ml of final culture volume was calculated by the following formula:

Volume of inoculum = (Final volume of secondary culture * 0.5 OD) / (Initial OD * 2^2).

The culture was incubated at 30°C, 200 rpm for 3 hours (2 generations) till the final OD reached 0.6-0.8. The cells were harvested by 5 min spin at 3000 rpm. The cells were washed by resuspending them in 10 ml sterilized water. The tubes were centrifuged for 5 min at 3000 rpm to pellet the cells. The cells were resuspended in 300 µl Lithium solution. About 200 µl of these competent cells were used for each transformation.

2.1.2 Yeast transformation

The sample DNA was mixed with 10 µg of carrier DNA. About 200 µl of competent yeast cells was added to each transformation tube. To each tube, 1.2 ml of PEG solution was added and incubated at 30°C for 30 min in shaker incubator. The cells were exposed to heat shock at 42°C for 15 min. The cells were collected by centrifugation at 3000 rpm for 1 min. The pellet was resuspended in 200 µl 1X TE buffer and then spread on synthetic deficient medium plates. The plates were incubated at 30°C for 72 hours.

2.1.3 RNA isolation and Semi-quantitative RT-PCR

Total RNA was isolated from yeast strains SSY7, SSY8 and SSY11 after growing at 30°C by acid phenol method as described (Schmitt M E et al,, 1990) with some modifications. About 10 ml culture was grown till OD_{600} reached 1 which is the mid log phase. The cells were pelleted down and the pellet was resuspended in 400 µl TES buffer (10 mM Tris Cl pH=7.5, 10 mM EDTA and 0.5% SDS). Equal volume of acid phenol (pre-equilibrated with water) was added, vortexed vigorously for 10 sec and incubated 65°C for 60 min with intermittent vortexing at every 15 min. The mixture was quickly chilled on ice for 5 min and centrifuged at 14000 rpm for 10 min at 4°C. The top aqueous layer was transferred to new tube and about 400 µl of chloroform was added and vortexed briefly for 10 sec. It was centrifuged at the same speed and temperature mentioned above. The extracted aqueous layer was precipitated by adding $1/10^{th}$ volume of 3M sodium acetate (pH=5.2) and 2.2 volume of chilled absolute ethanol. The precipitate was collected by centrifugation as above and was washed with 70% alcohol. The RNA was air-dried for 5 min and suspended in 30 µl of DEPC treated water. The concentration of RNA was determined spectrophotometrically (JASCO spectrophotometer EMC-709) at 260 nm. About 15 µg of RNA was subjected to DNase I treatment at room temperature for 30 min. The reaction was stopped by addition of 25mM EDTA and heat inactivation at 65°C for 10 min. The sample was rapidly chilled on ice and then reverse transcribed with oligo dT primer (Sigma Aldrich) using reverse transcriptase (Omniscript kit, Qiagen). Resulting cDNA was subjected to PCR (35 cycles) using gene specific primers. We used the oligos OMKB 209 and OMKB 56 to amplify 258 bp at the 3' end of the *TgMRE11* transcript. Primers OSB 50 and OMKB 85 were used to amplify chimeric *MRE11* transcript of size 300 bp. We used OMKB 210 and OMKB 75 to PCR amplify the 220 bp long *TgKU80* transcript. The details of primers sequences used in

the Semi-quantitative RT-PCR are given in table 5. The PCR products were run on 0.8% gel stained with ethidium bromide.

2.1.4 Protein isolation by TCA method

Cells were grown overnight in 5 ml of appropriate medium. Cells were inoculated into 20 ml of appropriate medium. The volume of secondary inoculum was determined by the following formula:

Volume of inoculum = (Final volume of secondary culture * 0.5 OD) / (Initial OD * 2^2)

The culture was grown for 2 generations so that the final OD reached 0.5 (For yeast, it takes about 3.5 hrs). The culture was spinned down (total 10 OD) at 3500 rpm for 5 min, washed once with double distilled water (500 µl). The cell suspension was transferred to a 2 ml microfuge tube and spinned again (can be stored at -80°C at this point) for 12000 rpm for 5 min. The pellet was washed with 500 µl 20% TCA. It was centrifuged 12000 rpm for 5 min. The pellet was resuspended in 200 µl 20% TCA, added glass beads upto 2/3rds volume. It was vortexed with intermittent cooling on ice at 4°C for half an hr in cold room. The supernatant was transferred into new 1.5 ml tube. It was washed up to 1 ml recovery with 5% TCA by mixing thoroughly and collecting into the same tube. It was spinned for 10 min at 3000 rpm at RT. The supernatant was discarded and the precipitate was solubilized by vortexing in 180 µl SDS-PAGE loading buffer. About 20 µl of 1M DTT and 100 µl Tris pH=9 were added. The sample was boiled for 3 min and centrifuged for 5 min at top speed.

2.2 Recombinant DNA methodology

2.2.1 Bacterial competent cell preparation

Overnight culture with a single colony of bacteria was kept in 5 ml of LB containing appropriate antibiotics at 200 rpm, 37°C. About 500 µl of overnight culture was added as inoculum to the secondary culture containing 25 ml of LB containing appropriate antibiotics. It was incubated at 37°C till the OD_{550} reached 0.5 to 0.8. The culture was centrifuged at 8000 rpm for 8 min at 4°C. The pellet was gently resuspended in 13 ml of 0.1 M calcium chloride (ice cold). It was centrifuged at 8000 rpm for 8 min at 4°C. The pellet was gently resuspended in 6.5 ml of 0.1 M calcium chloride and incubated on ice for 4-8 hours. It was centrifuged at 8000 rpm for 8 min at 4°C. Pellet was gently resuspended in 1070 µl of ice cold 0.1 M calcium chloride. 176 µl of glycerol was added. 100 µl of competent cells was added in each pre-chilled micro centrifuge tubes. Then all the tubes were frozen in liquid nitrogen and preserved at -80°C.

2.2.2 Bacterial transformation

Competent *E. coli* (DH5α strain) cells were thawed at 4°C. 1-2 µl of plasmid DNA or 5 µl of ligation mixture was added to 100 µl of competent cells. The tubes were stored on ice for 30 min and then transferred to 42°C for exactly 90 sec. Then they were rapidly transferred to ice and kept for 2 min. About 1 ml of LB broth was added to each tube and incubated at 37°C for 1 hr at 200 rpm. The samples were centrifuged at 5000 rpm for 1 min and the cell pellet was resuspended in 200 µl of LB broth and plated on LB medium containing appropriate antibiotics. The plates were incubated at 37°C for 12-16 hours.

2.2.3 Plasmid DNA isolation by alkaline lysis method

A single bacterial colony was inoculated in 5 ml of LB containing the appropriate antibiotics and kept at 37°C, 200 rpm for 16 hours. The overnight culture was

centrifuged at 4,500 rpm for 15 min. To the pellet, about 200 µl of solution I (containing 25 mM Tris (pH=8) and 10 mM EDTA) was added and mixed thoroughly and transferred to a new 1.5ml microfuge tube. The tube was incubated on ice for 5 min. Then the tube was transferred to room temperature and 200 µl of solution II (containing 0.2 M NaOH, 1% SDS) was added and inverted 4-5 times. Then, 150 µl of solution III (containing 3M sodium acetate) was added and inverted 4-5 times on ice. It was incubated for 5 minutes on ice with intermittent mixing. A white precipitate was formed. Then it was centrifuged at 13,000 rpm for 15 min at RT. Supernatant was transferred to a new 1.5 ml tube and 2.2 volumes of 100% alcohol was added and kept for DNA precipitation at -20°C for 1hr and centrifuged at 13,000 rpm for 30 min at 4°C. The supernatant was decanted and 500 µl of 70% alcohol was added. It was centrifuged at 13,000 rpm for 5 min at 4°C. Supernatant was decanted and the pellet was allowed to air-dry for 5 min. The pellet was resuspended in 30 µl of 1X Tris-EDTA buffer. About 5 µl of 10 mg/ml RNase was added and incubated in water bath at 37°C for 20 min. Equal volumes of 1X TE and phenol-chloroform-isoamyl alcohol (PCIAA in 25:24:1) was added, vortexed till a homogenous mixture is formed and centrifuged at 12,000 for 12 min. The upper aqueous layer was removed into a new 1.5 ml microfuge tube and one-tenth volume of solution III and 2.2 volumes 100% alcohol was added and kept for precipitation at -80°C for 2 hrs. It was centrifuged at 13,000 rpm for 30 min at 4°C. Supernatant was decanted and 500 µl of 70% alcohol was added and centrifuged for 5 min. The supernatant was decanted and the pellet was air-dried for 5 minutes till all the traces of alcohol are removed. The pellet was resuspended in 30 ul of IX TE.

2.2.4 Construction of plasmids

The *TgMRE11* gene was amplified by nested PCR using the PCR extender enzyme (Eppendorf) from the cDNA library of *Toxoplasma gondii* RH strain. This ORF of size 3.4 Kb was cloned into the intermediate vector TOPO2.1 TA vector (Invitrogen). It was released by *BamHI* and subcloned into pTA plasmid (Miller III CA et al., 1998). To generate chimeric *MRE11*, the N-terminal *TgMRE11* of size 2.8 Kb was amplified from full length *TgMRE11* gene by using PCR extender enzyme. It was cloned into TOPO2.1 TA vector and then the insert was released by *BamHI* and *SacI* enyzmes. The C-terminal *ScMRE11* gene of 1128 bp was amplified with PCR extender enzyme. It was cloned into TOPO4 PCR vector (Invitrogen) and finally released by *SacI* and *PstI*. Both the inserts were ligated into pTA plasmid at *BamHI* and *PstI* sites.

The *TgKU80* gene was PCR amplified using PCR extender enzyme from the genomic library of *T. gondii* RH strain. The product of size 5183 bp was cloned into the pTZ57R/T vector (Fermentas). It was released from the pTZ57R/T-*TgKU80* plasmid by *BglII* and *XbaI* digestion and subcloned into the *BamHI* and *XbaI* sites of pTA. Similarly, the *ScKU80* gene of 1890 bp was PCR amplified using PCR extender enzyme from the W303a genomic DNA. It was cloned into the pTZ57R/T vector. Then the insert was released by *BglII* and *SalI* digestion and subcloned into the *BamHI* and *SalI* sites of pTA.

The open reading frame encoding *Toxoplasma gondii* Rad51 was amplified from the cDNA library (provided by Prof. Vern Carruther) of *T. gondii* RH strain using PCR extender enzyme as described by the manufacturer. The PCR product of size 1065 bp was first cloned into TOPO2.1 TA vector (Invitrogen). The resultant plasmid was digested with *BamHI* to release the insert containing *TgRAD51* gene and was ligated with pTA plasmid. The insert was released by *BamHI* digestion and further cloned

into pET28a plasmid under the T7 promoter. The *ScRAD51* gene was PCR amplified from the genomic DNA of W303a strain. The PCR product of size 1203 bp was cloned into the TOPO2.1 TA vector. The insert was further released by *BamHI* and *PstI* and finally cloned into pTA plasmid under GPD promoter. All the plasmids used in this study are described in Table 3. The oligos used for gene cloning and their sequences are listed in Table 4.

2.3 Methods in Biochemical studies

2.3.1 Protein Purification

The plasmid TgRad51:pET28a was transformed into *Escherichia coli* host strain BL21 and Rosetta. About 200 ml culture was grown at 37°C in Luria broth containing 10 mg/ml kanamycin (Sigma) and 34 ug/ml chloramphenicol to an OD_{600} of 0.8. Expression of TgRad51 protein was induced by 1 mM isopropylthiogalactoside (Sigma) and the incubation continued for 4 hours. The cells were harvested by centrifugation at 4000 g for 30 min at 4°C. The pellet was weighed and stored at -80°C. The cell pellet was thawed and resuspended in 3 ml lysis buffer per gm of cell pellet. A pinch of protease inhibitor, phenylmethanesulfonyl fluoride (PMSF) and 1 mg/ml final concentration of lysozyme (Amresco) was added. Sonication was carried out by giving thirteen 40 second bursts at 200W using sonicator (SONICS Vibra ^{m} Cell ^{TM}) with intermittent cooling. The lysate was centrifuged at 10,000 g for 45 min at 4°C to remove the insoluble material. The supernatant containing the proteins was mixed with 50% Ni-NTA slurry (0.875 ml of slurry per 4 ml of cleared lysate) and the loading flowthrough was collected. The resin was resuspended in wash buffer W1 (containing 50 mM NaH_2PO_4, 300 mM NaCl, 20 mM Imidazole, 30% glycerol, 20mM β-mercapitoethanol, 2% Tween 20). Wash buffer W2 (same as W1) was added

to the resin. Further the column was washed with wash buffer W3 (containing 50 mM NaH_2PO_4, 300 mM NaCl, 50 mM Imidazole, 30% glycerol, 20mM β-mercapitoethanol, 2% Tween 20). Elution buffer E_{250} (containing 50 mM NaH_2PO_4, 300 mM NaCl, 250mM Imidazole, 30% glycerol, 20mM β-mercapitoethanol, 2% Tween 20) was added slowly along the side walls of the column and the fractions were collected from the column and pooled. Then elution buffer E_{400} (50 mM NaH_2PO_4, 300 mM NaCl, 400mM Imidazole, 30% glycerol, 20mM β-mercapitoethanol, 2% Tween 20) was added to the column and the fractions were collected. All the wash and elution buffers pH was adjusted to 8. Recombinant TgRad51 was eluted and identified by 10% SDS-PAGE.

The dialysis bag (Himedia) was boiled for 10 min in a solution containing 5 mM EDTA and 200 mM sodium bicarbonate. Then the bag was washed thoroughly with autoclaved double distilled water and tied at the bottom. The protein fractions eluted with 250 mM and 400 mM imidazole was pooled separately and dialysed with dialysis buffer (containing 20 mM Tris HCl (pH=8), 1 mM dithiothreitol (DTT), 5% glycerol). Two changes of this buffer was given to each protein sample. The protein was purified to near homogeneity and identified by 10% SDS-PAGE, aliquoted and stored at -80°C.

The concentration of the purified recombinant TgRad51 protein was determined by UV absorbance at 280 nm by using the extinction coefficient of 21025 M^{-1} cm^{-1} calculated from the amino acid sequence by using ExPaSy Protparam tool. The purified protein was further confirmed by MALDI-TOF analysis and by MS sequencing.

2.3.2 ATP hydrolysis activity

The rate of ATP hydrolysis of TgRad51 was determined using Enz Chek Phosphate assay kit (Molecular Probe, E-6646). A typical reaction mixture was composed of 2 μM TgRad51 along with 30 fold molar excess of φxssDNA (New England Biolab) (60 μM) in presence of ATP containing reaction buffer (25 mM Tris-HCl pH 8, 5% glycerol, 1 mM DTT and 1 mM $MgCl_2$). The reaction was incubated in presence and absence of φxssDNA at 37°C for 35 min. At every 5 min interval (0-35 min), a measured volume of the reaction mix was added to the 22°C pre-warmed PNP reaction mixture (manufacturer description) and incubated at 22°C for 30 minutes. The PNP reaction mixture consisted of MESG substrate, reaction buffer, 0.1 M purine nucleoside phosphorylase (PNP), 0.1 M KCl, 0.1 M NaCl and water. The ATP hydrolysis resulted in generation of inorganic phosphate which then reacted with another substrate, MESG (2-amino-6-mercapto 7-methyl purine riboside) generating 2-amino-6-mercapto 7-methyl purine having an absorbance at 360nm. This reaction is catalyzed by PNP. At a typical ATP concentration, the amount of inorganic phosphate released was plotted against different time interval and from the slope of the straight line, the rate of the reaction was calculated. For determining the kinetic parameters of ssDNA-dependent ATPase activity of TgRad51 protein, various concentrations of ATP (20-600 uM) were taken. The rate of the reaction were plotted against each concentration of ATP to determine the k_{cat} (catalytic constant) and K_m (michaelis-menten constant) of the enzymatic reaction catalyzed by TgRad51 protein using Graph pad Prism software, version 5.04.

2.3.3 Western blotting

The purified recombinant TgRad51 protein was detected by using anti-His antibody (Santa Cruz Biotechnology Inc., CA) by enhanced chemiluminescence (ECL) based method.

We isolated the protein from NRY1, NRY2, MVS26, SSY1, SSY2 and NA14Δ*rad51* strains by trichloroacetic acid protocol as described in Laskar S et al., 2011. Appropriate volumes of the protein samples were loaded on 12% SDS polyacrylamide gel. The gel was then transferred onto polyvinylidene difluoride (PVDF) membrane and immuno blotted for TgRad51, ScRad51 and Actin. The anti PfRad51 antibody (a kind gift from Dr. Nirbhay Kumar) was used in 1:4000 dilution to detect TgRad51 protein. The anti ScRad51 antibody (Santa Cruz Biotechnology Inc., CA) and anti Actin antibody (Abcam) were used in 1:5000 dilution. Horseradish peroxide-conjugated rabbit IgG (Santa Cruz Biotechnology Inc., CA) was used as the secondary antibody for ScRad51 at 1:10,000 dilution and HRP conjugated mouse IgG (Promega) was used as a secondary antibody for PfRad51 and actin blot at the same dilution. Western blots were developed by the manufacturer protocol (pierce). All the normalizations in the blots for yeast proteins were done against Actin.

2.4 Methods for studying DNA repair and recombination assays

2.4.1 Multiple DSB by MMS treatment

a) Spotting assay

All the five strains NRY3, SBY1, SBY2, SSY7 and SSY8 were grown in 5 ml of synthetic complete medium lacking tryptophan and incubated overnight at 30°C, 200 rpm. Secondary inoculum was given in 5 ml of fresh medium with the expected OD_{600} as 3 and was grown for two generations.

Volume of inoculum = (Final volume of secondary culture * 0.5 OD) / (Initial OD * 2^2)

The OD of different strains were determined and was normalized to the least OD. Each culture was serially diluted till 10^{-4}. About 5 ul of culture was spotted directly on both untreated and MMS (0.005%) treated SC-trp plates. Spotting was done from lower to higher dilution. The plates were incubated for 3 days at 30°C and the growth in MMS treated and untreated plates were recorded.

b) Return to growth experiment

All the three strains NRY1, NRY2 and MVS26 were grown in 5 ml of tryptophan deficient medium. About 5 ml secondary culture of each strain was kept in duplicates. They were grown till OD_{600} reached 1. One untreated set of cultures was further incubated for 2 hours. Second set of cultures were subjected to 0.005% MMS treatment for 2 hours. By subsequent wash steps, the MMS was removed and equal number of cells from each culture was spread on the synthetic complete plates lacking tryptophan. The plates were incubated for 3-4 days at 30°C and the percent viability was calculated as the ratio between the survivors on the treated to the untreated plates. These return to growth experiments were performed in triplicate with 0.005% MMS concentration.

2.4.2 NHEJ assay in yeast

The pRS426 parent plasmid was linearized with *EcoRI* enzyme and the DNA was precipitated. This plasmid has *URA3* selectable marker. This single break cannot be repaired by homologous recombination. About 500 ng of both uncut and cut pRS426 parent plasmid was transformed into each of the three strains *Δku80* (SSY9), SSY10 and SSY11 were grown in tryptophan deficient medium. The transformed mixture

was plated on SC-trp-ura double drop-out plates. The cells grown on cut DNA transformant plate should be repaired by NHEJ only. Hence, the ratio of the cut DNA transformants to uncut DNA transformants gave the efficiency of NHEJ in each of these strains. This assay was repeated thrice and the graph was plotted.

2.4.3 Gene conversion assay

In NA14 strain, there is a cassette integrated in chromosome V where two consecutive *URA3* genes are separated by *KANMX6*. The *URA3* genes are separated by 3 Kb distance. The first *ura3* gene is having a HO endonuclease recognition site in it (Agmon N et al., 2009). This strain expresses HO endonuclease when grown on galactose medium. The two strains SSY1 and SSY2 containing pTA plasmid with *ScRAD51* or *TgRAD51* respectively were assayed. The SSY1 and SSY2 strains were streaked on synthetic complete (SC) plates lacking tryptophan with glycerol as the carbon source. The wild type NA14 strain and negative control *NA14*Δ*rad51* strains were grown on synthetic complete plates containing glycerol as the carbon source. Individual colonies were resuspended in water, serially diluted and about 1000 cells from each strain were plated on two kinds of medium; one of them containing glucose and the other containing galactose as carbon sources. The plates were incubated at 30^{o}C for 3 days. Upon induction with galactose, a single DSB was created at the HO cut site. The repair efficiency was calculated by measuring the ratio of the number of colonies grown on medium containing galactose and that grown on medium containing glucose (no HO induction) was determined. The repair could be achieved either by GC or SSA. To determine the efficiency of repair choice, the survivors on galactose plate were patched on YPD glucose plate containing G418 sulphate. The ratio of G418 sulphate survivors to the total colonies on galactose medium gave the

repair efficiency by GC. The ratio of G418 sulphate non-survivors to the total colonies on galactose medium gave the repair efficiency by SSA. This assay was done thrice and the mean value was plotted using Graph Pad Prism5 software.

2.4.4 Gene targeting assay

KANMX6 cassette (1536 bp) containing the upstream P_{TEF} region and the downstream T_{TEF} region was released from pFA6a-KANMX6 (Longitine MS et al, 1998) by *NotI* digestion (*NotI* site at 671 bp) and further sub cloned into pSD158 (Diede SJ et al, 1999) at *NotI* site. The cassette comprised of KANMX6-Adh4-Ade2-Adh4 and was released by *SalI* digestion. *SalI* site was present at 37 bp in P_{TEF} and 6501 bp of pSD158 plasmid. About 1 ug of the *SalI* digested pSD158-KanMX6 plasmid was individually transformed into three strains *Δrad51* (NRY1), NRY2 and MVS26. The efficient gene targeting at the *ADH4* locus by homologous recombination resulted in expression of *ADE2* gene but did not confer resistivity to G418 sulphate. Hence, the selection was done on the SC plates lacking adenine and tryptophan amino acids. All the transformants obtained on these plates were patched on YPD agar plate and grown for 2 days at 30°C. These patched colonies were replica plated on YPD plate containing G418 sulphate and grown for 3 days at 30°C. Colonies sensitive to G418 sulphate represents targeted integration, while G418 sulphate resistant colonies correspond to random integration. Finally we calculated the % gene targeting efficiency by using the following formula:

% Gene Targeting Efficiency = (Number of G418 sulphate sensitive colonies) / (Total number of $ADE2^+$ colonies) × 100.

In each case the value obtained for *Δrad51* strain was subtracted. This assay was repeated three times.

In another gene targeting assay, we monitored the gene targeting efficiency with increasing stretches of homologous flanking sequences (200 bp, 500 bp or 1000 bp). For that purpose we started with SLY3 strain where the *SBA1* gene was already insertionally inactivated by a *KANMX6* cassette (Laskar S et al., 2011). We had amplified that *KANMX6* cassette by three different sets of primers positioned at 200 bp/ 500 bp/ 1000 bp upstream and 200 bp/ 500 bp/ 1000 bp downstream of the *KANMX6* insertion. The primer pair OSB5 and OSB6 resulted in 200 bp flanking sequence on each side of *KANMX6* cassette. Similarly OMKB193 and OMKB194 produced 500 bp flanking sequences of *SBA1* on either side of *KANMX6* and OMKB191 and OMKB192 amplified 1 Kb flanking regions on either side of *KANMX6* cassette. These three different PCR products were individually transformed to *Δrad51* (NRY1), NRY2 and MVS26 strain to knock out *sba1* by homologous recombination and selected on SC-trp containing G418 sulphate. The number of G418 sulphate resistant colonies was measured. Each of the transformations were repeated at least three times and was plotted using Graph Pad Prism with error bars.

We compared the efficiency of targeted integration in the *rad51Δ* and *rad51Δku80Δ* strains. The gene targeting efficiency was monitored in these strains background with the increasing length of homologous flanking sequences (500 bp or 1000 bp). In NKY8 strain, the *CHL1* gene was already insertionally inactivated by a *HIS* cassette. We had amplified the *HIS* cassette by two sets of primers positioned at 500 bp/1000 bp upstream and 500 bp/1000 bp downstream of the *HIS* insertion. The primer pair OMKB 211 and OMKB 212 resulted in 500 bp flanking sequence on each side of *HIS* cassette. Similarly OMKB 213 and OMKB 214 produced 1 Kb flanking sequences on each side of the *HIS* cassette. These two PCR products were individually transformed into NRY1, NRY2 and MVS26 strain to knock out *CHL1* gene by homologous

recombination and the transformants were selected on SC-trp-his plate. The number of colonies were measured. This experiment was repeated twice. Similar experiments were done in SSY4, SSY5 and SSY6 strain to replace *CHL1* gene with *HIS* cassette by homologous recombination and the transformants were selected on Sc-trp-his plate. The resistant colonies were measured and each of these transformations was repeated two times and was plotted using Graph Pad Prism. The details of primer sequences used in the gene targeting experiments are described in Table 6.

2.5 Yeast strains

The yeast strains used in this work are tabulated in Table 2. Plasmids pTA containing full length *TgMRE11* gene and chimeric *MRE11* were transformed into *Δmre11* strain of yeast to generate SSY7 and SSY8 strains. Empty vector pTA was transformed into wildtype and *Δmre11* strain of yeast to create NRY3 and SBY1. The construct pTA/*ScMRE11* was transformed into *Δmre11* strain and this new strain was designated as SBY2.

The yeast expression vectors harboring *ScKU80* and *TgKU80* were transformed into the *Δku80* strain to create SSY10 and SSY11 strains respectively. The empty vector pTA was transformed in *Δku80* strain to generate SSY9.

The yeast expression vectors harboring *TgRAD51* and *ScRAD51* were transformed into the *Δrad51* (LS402) strain to create MVS26 and NRY2 strains respectively. The empty vector pTA was transformed in *Δrad51* strain to generate NRY1. NA14 and *NA14Δrad51* strains were used (Agmon N et al., 2009) for gene conversion study. We transformed the above mentioned yeast expression vectors containing *TgRAD51* and *ScRAD51* into *NA14Δrad51* strain to generate SSY1 and SSY2 strains respectively.

KANMX gene was amplified with Taq polymerase (Jonaki) from yeast *ku80Δ* strain. The primer pair OSB127 and OSB128 resulted in 200 bp and 220 bp flanking sequence of *KU80* on each side of *KANMX* gene. This linear *KANMX* cassette of 2220 bp was used to knockout *KU80* from the *rad51Δ* (LS402) strain. The yeast expression vector pTA was transformed into the *Δrad51Δku80* (SSY3) to generate SSY4 strain. The yeast expression vectors harboring *ScRAD51* and *TgRAD51* were transformed into the *Δrad51Δku80* (SSY3) strain to create SSY5 and SSY6 strains respectively.

Table 1: Yeast strains used in this study

Strains	**Genotype**
W303a	*MAT*a *15ade2-1, ura3-1, 112 his 3-11, trp1, leu2-3*
NRY1	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* [*phi*$^{+}$] *RAD51:: LEU2 pTA*
NRY2	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* [*phi*$^{+}$] *RAD51::LEU2 pTA/ScRAD51*
NRY3	*MAT*a *15ade2-1, ura3-1, 112 his 3-11, trp1, leu2-3 pTA*
MVS26	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* [*phi*$^{+}$] *RAD51::LEU2 pTA/TgRAD51*
NA14	*MAT*a-*inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100*
NA14Δrad51	*MAT*a-*inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100 RAD51::LEU2*
SSY1	*MAT*a-*inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100 RAD51::LEU2 pTA/ScRAD51*
SSY2	*MAT*a-*inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100 RAD51::LEU2 pTA/TgRAD51*
SSY3	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* [*phi*$^{+}$] *RAD51::LEU2 KU80::KANMX*
SSY4	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* [*phi*$^{+}$] *RAD51::LEU2 KU80::KANMX pTA*
SSY5	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* [*phi*$^{+}$] *RAD51::LEU2*

	KU80::KANMX pTA/ScRAD51
SSY6	*MAT*a *leu2-3,112 trp1-1 can1-100 ura3-1 ade2 -1 his3-11,15* [*phi*$^{+}$] *RAD51::LEU2 KU80::KANMX pTA/TgRAD51*
SBY1	*MAT*a *leu2-3,112his3-11,15ade2-1,trp1,ura3-1,MRE11::URA3,VIIL::ADE2 (pSD160), LEU2(pSD158) pTA*
SBY2	*MAT*a *leu2-3,112his3-11,15ade2-1,trp1,ura3-1,MRE11::URA3,VIIL::ADE2 (pSD160), LEU2(pSD158) pTA/ScMRE11*
SSY7	*MAT*a *leu2-3,112his3-11,15ade2-1,trp1,ura3-1,MRE11::URA3,VIIL::ADE2 (pSD160), LEU2(pSD158) pTA/TgMRE11*
SSY8	*MAT*a *leu2-3,112his3-11,15ade2-1,trp1,ura3-1,MRE11::URA3,VIIL::ADE2 (pSD160), LEU2(pSD158) pTA/chimeric MRE11*
SSY9	*MATα leu2-3 112 trp1-1 can1-100, ura3-1 ade2-1 his 3-11,15 adh4::Ade2 rad5+ KU80::KANMX pTA*
SSY10	*MATα leu2-3 112 trp1-1 can1-100, ura3-1 ade2-1 his 3-11,15 adh4::Ade2 rad5+ KU80::KANMX pTA/ScKU80*
SSY11	*MATα leu2-3 112 trp1-1 can1-100, ura3-1 ade2-1 his 3-11,15 adh4::Ade2 rad5+ KU80::KANMX pTA/TgKU80*
SLY3	*MAT*a *SBA1:: KAN*^{r}
Yku80Δ	*MATα leu2-3 112 trp1-1 can1-100, ura3-1 ade2-1 his 3-11,15 adh4::Ade2 rad5+ KU80::KANMX*
NKY8	*MAT*a*,SPC29-CFP::KAN,mRFP-TETR,URA3::TETO,GFP-LACI::LEU2 CHL1::HIS3*

Table 3) List of the plasmids used in this study

Name of the plasmid	**Brief description**
pET28a	Bacterial protein expression vector containing N-terminal His-tag
pTA	Plasmid containing TRP marker
pRS426	Plasmid containing URA marker
pSD158	Cloning of *KANMX* gene at *NotI* site of plasmid (It contains 5' *ADH4* and 3' *ADH4* gene stretches)
pFA6a-KanMx6	*KANMX* gene for cloning

Table 4) List of primers and their sequences used for PCR in this study

Name of the primer	**Primer sequence (5' to 3')**	**Purpose (Gene cloning)**
OMKB 78 F	CTCTATCTTTCTCTGACTTTCCAG	pTA/ *TgMRE11*
OMKB 79 R	CGGATGAACGAGCGAACACTGG	
OMKB 72 F	GGATCC ATGGCTTCTGCTTTGCGCAGCG	
OMKB 73 R	GGATCC CGCCTAGAAGGATAAAAACGCAAC	
OMKB 108 R	ATCCGA GCTCCGGCCGCGAGTGGAGCAAGTC	pTA/ *TgMRE11* N-terminal
OMKB 109 F	ATACGAGCTCCATTTGAGGCCTCACGATAAAGATG	pTA/ *ScMRE11* C-terminal
OMKB 85 R	CTGCAG CTATTTTCTTTTCTTAGCAAGGAGAC	
OMKB 136 F	ATACAGATCTATGGCGCTTCCTGGGCAAAGC	pTA/ *TgKU80*
OMKB 138 R	ATCCAGATCTTCTAGACTACTCGACGAGATCGAGTAAGTC	
OMKB 155 F	ATAC AGATCT ATGTCAAGTGAGTCAACAAC	pTA/ *ScKU80*
OMKB 156 R	ATCC GTCGAC TTAATTATTGCTATTGTTTGGAC	
OMKB 65 F	GGATCCATGAGCGCCGTCTCTCTTCAG	pTA/ *TgRAD51* pET28a/ *TgRAD51*
OMKB 82 R	GGATCCTCAGTTGTCTTCGTAGTCGCC	
OMKB 90 F	GGATCC ATGTCTCAAGTTCAAGAAC	pTA/ *ScRAD51*
OMKB 88 R	CTGCAG CTACTCGTCTTCTTCTC	

Table 5) List of primers and their sequences used for Semi-quantitative RT-PCR in this study

Name of primer	**Primer sequence**	**Purpose (Semi-quantitative RT-PCR)**
OMKB 209 F	CAGGGAGACGAAGGCCG	*TgMRE11*
OMKB 56 R	GGATCCCTAGAAGGATAAAAACGCAAC	*TgMRE11*
OSB 50 F	CAAACGAGTGCGAACTGCAAC	Chimeric *MRE11*
OMKB 85 R	CTGCAGCTATTTTCTTTTCTTAGCAAGGAGAC	Chimeric *MRE11*
OMKB 210 F	GCGGCCGACGACAGCTTCCG	*TgKU80*
OMKB 75 R	CTACTCGACGAGATCGAGTAAGTC	*TgKU80*

Table 6) List of primers and their sequences used for PCR in this study

Name of the primer	**Primer sequence (5' to 3')**	**Length of flanking ends**	**Purpose (gene targeting)**
OSB 5 F	TGCTACCCGCCTTCGAGTG	200 bp	*KANMX* gene with homologous ends on either side of *SBA1* gene
OSB 6 R	CACATACAGTTCCATTACTTGAC		
OMKB 193 F	CTCAGAAGAATTTCGTAAATCGG	500 bp	
OMKB 194 R	GGAGATGGTACCGGTTAAGCG		
OMKB 191 F	TCACACGTCCGTCATGTCTAC	1 Kb	
OMKB 192 R	GTCCTGCAGGAGACTTATTAGC		
OMKB 211 F	CAGCTCTCTAGCCAACAGCAG	500 bp	*HIS* cassette with homologous ends on either side of *CHL1* gene
OMKB 212 R	CTTGCGTATTATCTATAGCGGC		
OMKB 213 F	CACTCGTTGACTAGTTCAGAGG	1 Kb	
OMKB 214 R	GACGAACTTCATGTGACGGCTG		
OSB 127 F	AGTCTATTAGCGGAAGTACC	40 bp	Knockout of *KU80* gene with *KANMX* cassette in *Δrad51* strain of yeast
OSB 128 R	GAACGTCCTCTACCCACG		

CHAPTER 3

Development of surrogate yeast assay system for studying double strand break repair proteins of *Toxoplasma*

3.1 Introduction

Toxoplasma gondii shares a number of structural similarities with disease causing parasites namely *Plasmodium*, *Cryptosporidium* etc. Thus it has become a model system to do functional genomics (Kim K et al., 2004). However efficient gene targeting is a challenge in this parasite. It is interesting to note that *T. gondii* is the only protozoan parasite that harbors non homologous end joining mediated DNA break repair mechanism.

When a DSB is formed, HR repair depends on searching of extensive homologous stretches of DNA. Non-homologous end joining of broken DNA ends depends on little or no homology. Both pathways compete with each other for repair. In DNA repair, Mre11 (Meiotic recombination 11 protein) is recruited very early within minutes to the broken ends. If Ku70/80 heterodimeric protein binds to DSB, the repair takes place by NHEJ. But if the ends are resected by nucleases such as Sae2, it forms long 3' single stranded DNA wherein HR mediated repair occurs. During this repair process, Rad51 is recruited which forms nucleoprotein filament (Mimitou EP and Symington LS, 2009). During HR, Rad51 plays an important role in searching of homologous template and promotes strand invasion.

In DNA repair, Mre11 (Meiotic recombination 11 protein) is recruited to the broken ends much before the decision is made between HR and NHEJ. Mre11p is a multifunctional player and has several roles like recognition of DNA damage, binding to the broken chromosomal ends, initiating the repair process by resecting the broken ends and activates during the DNA damage checkpoint proteins. It exists as a MRX complex which is made up of two molecules each of Mre11, Rad50 (Radiation sensitive 50) and one molecule of Xrs2 (X-ray sensitive 2) protein. Mre11 protein is known to form a U-shaped dimer. The dimer interacts with Rad50 and Xrs2. The N-

terminally located phosphoesterase motifs contribute to the nuclease activity of the Mre11 protein (Rupnik A et al, 2010). Mre11p resects the broken DNA ends in 5' to 3' direction and bridges the broken ends in the form of MRX complex. It has 3' to 5' exonuclease activity and endonuclease activity (Williams RS et al., 2008). Mre11 nuclease and C-terminal DNA damage response functions are required for initiating yeast telomere healing (Bhattacharyya MK et al., 2008).

Results from two research groups (Fox BA et al., 2009; Huynh MH et al., 2009) showed that *KU80* is an essential component of NHEJ mechanism in *T. gondii*. The genome of this parasite (http://beta.toxodb.org/toxo5.0/) also revealed genes encoding putative DNA ligase IV and DNA-dependent protein kinase components of NHEJ. However, readily identifiable *KU* genes are absent in *Plasmodium* species (http://plasmodb.org/plasmo). Ku proteins are present but they do not participate in NHEJ in protozoans such as *Leishmania* and *Trypanosoma* species (Burton P et al., 2007; Glover L et al., 2008). Ku70 and Ku80 form a heterodimer that binds to the broken DNA ends and bridges the two ends. Ku protects these ends from the exonucleolytic processing and recruits DNA-PKcs, a serine/threonine protein kinase (Meek K et al., 2008). The complex of Ku and DNA-PKcs triggers the DNA-PKcs kinase activity, which in turn phosphorylates and activates a nuclease, Artemis that cleans up the ends. Final step is ligation involving ligase IV-XLF-XRCC4 complex. Ku also plays an essential role in telomere maintenance. It protects the telomere ends from inappropriate degradation, and contributes to the tethering of the telomeres to the nuclear periphery and the regulation of telomerase (Riha K et al., 2006). The function of Ku is to prevent the elimination of telomeres and the formation of extrachromosomal telomere-circles (t-circles).

Homologous recombination creates genetic diversity by cross-over and exchange of genetic material (Symington L S, 2002). In eukaryotes, Rad51 is the conserved protein involved in homologous pairing and strand exchange reaction (Liang F et al, 1998, Stark JM et al, 2002, 2003, 2004). Previously, *Plasmodium falciparum* Rad51 protein was identified, biochemically characterized and its role in homologous repair was elucidated (Bhattacharyya M K et al, 2003, 2005). Rad51 self assembles on the single and double stranded DNA to form the nucleoprotein filament. ATP binding is required for the nucleofilament formation. Replication protein A (RPA) bound ssDNA is recognized by Rad52 (Hays SL et al., 1998) and Rad51-Rad52 complex replaces RPA. In yeast, the direct association between the Rad51 and Rad54 recombination proteins was demonstrated (Jiang H et al., 1996). Rad54 acts as a translocase and DNA helicase in the recombination pathway. Rad55-Rad57 heterodimer binds to the ssDNA and shows recombination mediator activity. Direct *in vivo* interactions between Rad51 and Rad55 proteins, and between Rad55 and Rad57, have also been identified and these results indicate that all these four proteins form a part of the complex "recombinosome" (Hays SL et al., 1995).

Since, *T. gondii* is not easily amenable to genetic manipulation, the first objective was to develop a surrogate yeast assay system to study DNA double strand break repair proteins of *Toxoplasma gondii.* To address why *Toxoplasma gondii* chooses NHEJ, we identified and cloned *MRE11*, *KU80* and *RAD51* genes of *T. gondii.* Understanding the functional role of TgMre11 would give insights on pathway choice as it binds to the broken DNA ends much before the decision is made between HR and NHEJ. To understand NHEJ mechanism, we choose the central player TgKu80 for characterization. Further to understand HR pathway and the targeted gene disruption in this parasite, we choose the key player TgRad51 for our study. The main

focus of this objective was to understand the interplay of three very important parasitic DNA repair genes in yeast; *TgMRE11*, *TgKU80* and *TgRAD51*.

3.2. RESULTS

3.2.1. Development of surrogate yeast assay system to study *TgMRE11* gene

In order to understand the mechanism of DNA damage recognition and how the DSB repair pathway choice is made, we choose TgMre11 for our study. BLAST analysis of *Toxoplasma* database revealed putative ortholog of *TgMRE11* from GT1 strain. The *TgMRE11* gene is of size 2196 bp (Table 7) and has 24% sequence identity with *ScMRE11*. Full length *TgMRE11* gene was amplified by nested PCR from the cDNA library of *Toxoplasma gondii* RH strain. The obtained PCR product was of size 3.4 Kb (Appendix 1.1). DNA sequencing revealed that there is an additional 1.3 Kb stretch located in the middle of *TgMRE11* gene of RH strain (Figure 6). To express TgMre11 in yeast model system, the full length *TgMRE11* gene was initially cloned into pCR 2.1 TOPO vector and the insert was released by *BamHI* digestion and subcloned into yeast expression plasmid, pTA at the same site under the GPD promoter (Appendix 1.1).

First, we tried to investigate if *ScMRE11* gene can be functionally complemented by *TgMRE11* gene by MMS sensitivity assay. Methyl methane sulphonate (MMS) is a chemical mutagen which generates numerous DNA double strand breaks inside a cell and a functional Mre11 is required for the repair of these breaks inside the cell. Unrepaired breaks result in cell death. To perform this assay, pTA-*TgMRE11* plasmid was transformed into the *Δmre11* strain to create SSY7 strain and we tried to look at the percent survivability of the cells in this assay. Empty pTA plasmid was transformed into wild type and *Δmre11* strains to create the positive (NRY3) and

negative (SBY1) controls respectively. Another plasmid pTA expressing *ScMRE11* was also transformed into *Δmre11* strain which served as the positive control, SBY2 strain (Figure 7a). All the strains grew normally under untreated conditions. The strain expressing wild type *MRE11* (NRY3) and *ScMRE11* (SBY2) showed normal survivability in MMS treated conditions. Negative control (SBY1) did not grow in the MMS treated conditions. There was no survivability in cells harboring pTA-*TgMRE11* plasmid (SSY7) in MMS treated conditions in the MMS sensitivity assay (Figure 7b). This assay revealed that the *ScMRE11* gene cannot be functionally complemented by *TgMRE11* gene.

Since, both the nuclease and DNA damage response (DDR) domains should be functional for the complementation to happen, we reasoned that the lack of complementation could be due to the non-conserved C-terminal domain of TgMre11, which might not be functional in yeast. To address this possibility we generated a chimera by fusing the conserved nuclease domain (44% sequence identity) of *TgMRE11* with the C-terminal DDR domain of *ScMRE11* (Figure 8a). To construct this we have separately cloned the N-terminal *TgMRE11* and the C-terminal *ScMRE11* genes into pCR TOPO vector and the inserts were released individually and ligated into pTA plasmid at *BamHI* and *PstI* sites (Appendix 1.2). The pTA plasmid containing the chimeric *MRE11* gene was transformed into the *Δmre11* strain to generate SSY8 strain and assayed for its functional role in the yeast surrogate system (Figure 8b). The SSY8 strain grew normally under untreated conditions. But, there was no survivability in the cells expressing chimeric *MRE11* as observed by the spotting assay, under MMS treated conditions (Figure 9). MMS sensitivity assay clearly showed that *ScMRE11* cannot be functionally complemented by chimeric *MRE11*.

We concluded that the yeast model system cannot be used to assay for the functional role of *Toxoplasma gondii MRE11*. Since the *MRE11* gene was expressed in a heterologous system, we tried to look at the possibility if the protein was not getting expressed. Since we did not have the antibody against TgMre11 protein, we checked the expression at the RNA transcript level by using semi-quantitative RT-PCR. We isolated RNA from the strains expressing *TgMRE11* (SSY7) and chimeric *MRE11* (SSY8) (Figure 10a). Clearly, the PCR results showed that the RNA transcript at the 3' end of the *TgMRE11* and the chimeric *MRE11* got amplified and produced 258 bp and 300 bp bands respectively (Figure 10b and 10c). The DNase treated RNA sample was used as the template for PCR and these PCR products were loaded as the negative controls. As expected, bands were not observed in the negative controls (Figure 10b and 10c).

Table 7) Primary sequence information of *MRE11* gene from *Toxoplasma gondii GT1* strain and *Saccharomyces cerevisiae*:

STRAINS	*MRE11* gene (bp)	Protein length (aa)
Tg GT1	2196	731
Sc	2079	692

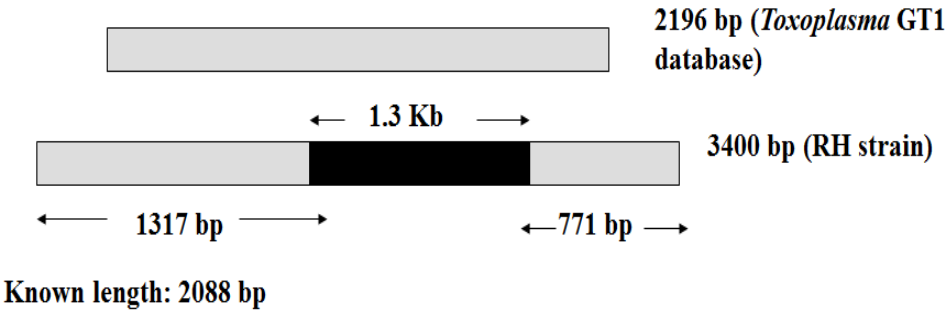

Figure 6. Schematic diagram showing the *MRE11* gene from GT1 and RH strains of *Toxoplasma gondii*. The shaded area in black indicates the additional 1.3 Kb gene stretch present in the middle of *TgMRE11* of RH strain.

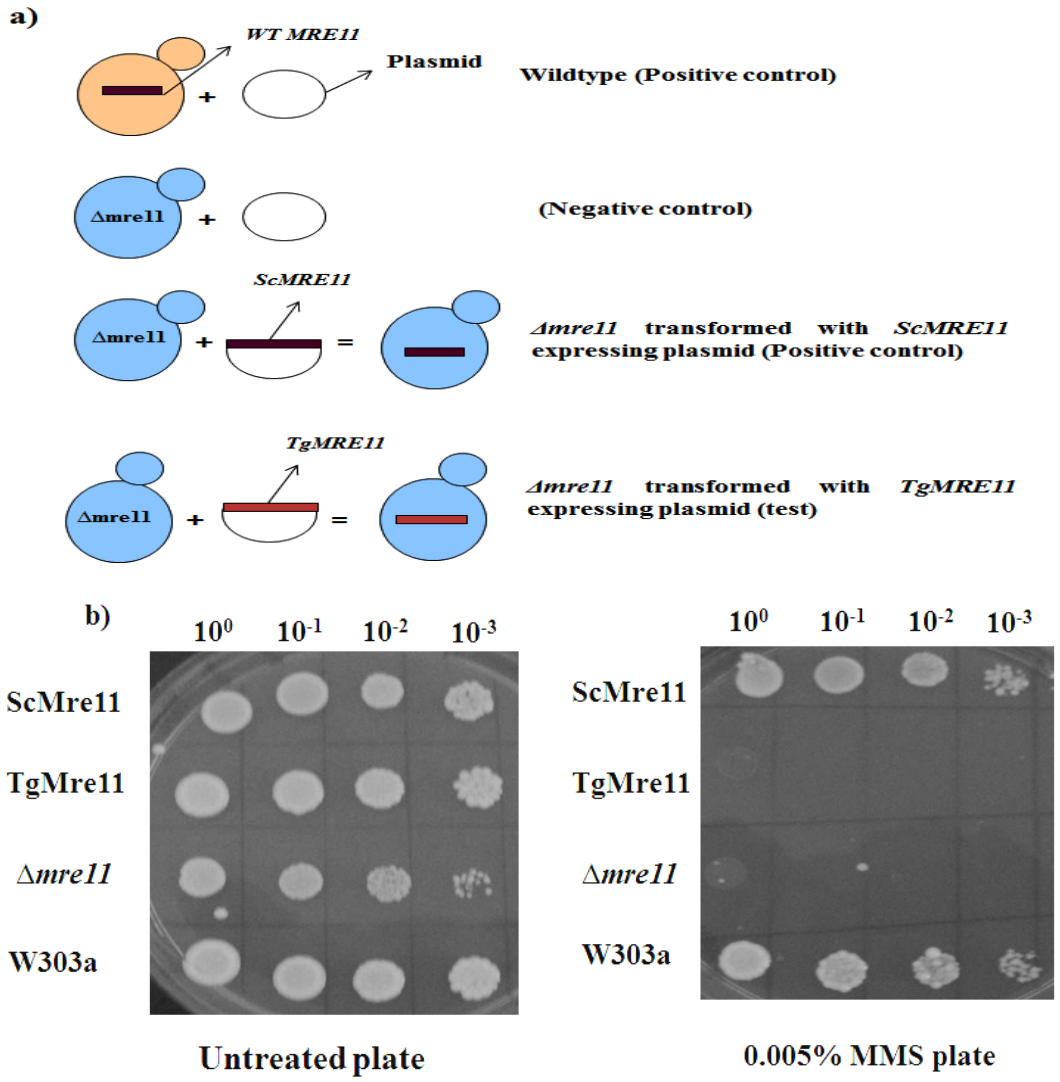

Figure 7. Functional complementation of *TgMRE11* in yeast model system. **a)** Illustration depicted the strains used in the MMS sensitivity assay.

b) showed the survivability of different strains under untreated and MMS (0.005%) treated conditions.

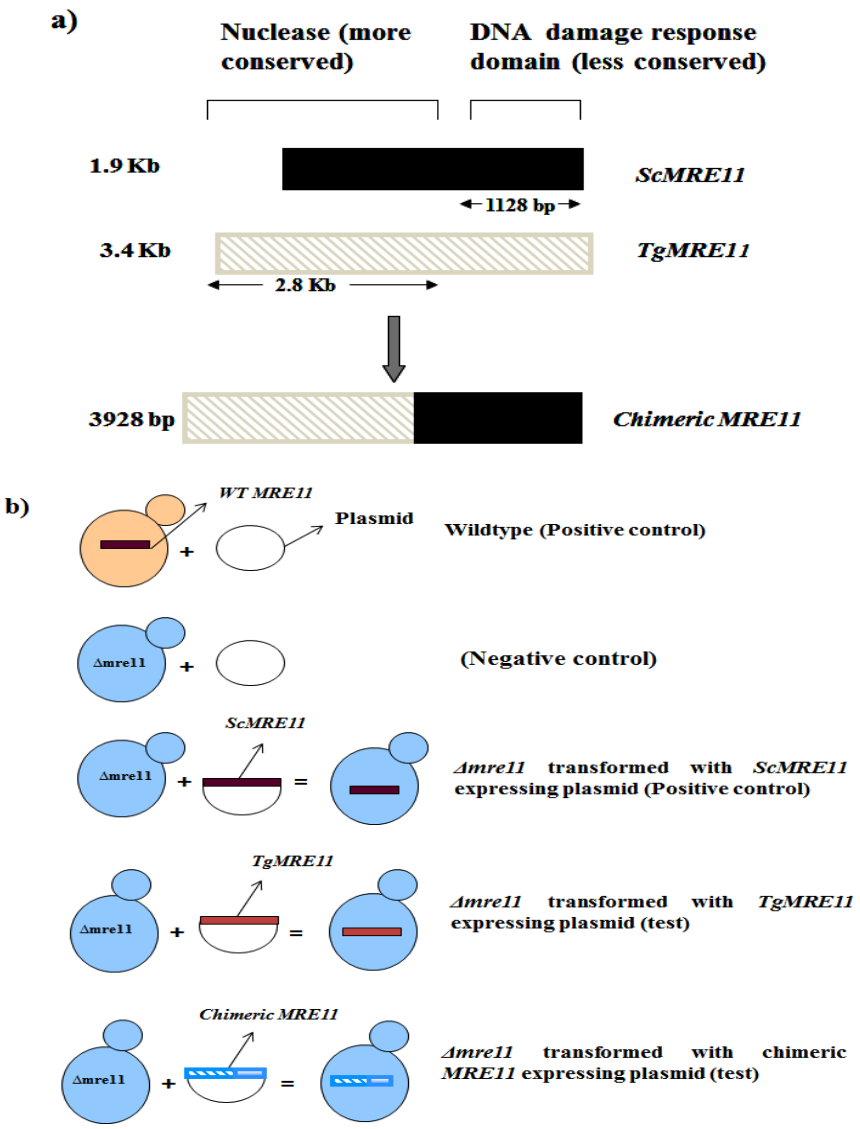

Figure 8. Functional complementation of chimeric *MRE11* in yeast. **a)** Schematic diagram of *TgMRE11* and *ScMRE11* chimera construct. The black box shows the *ScMRE11* and gray shaded box shows *TgMRE11* **b)** Illustration showing the different strains used for functional complementation of chimeric *MRE11* in yeast.

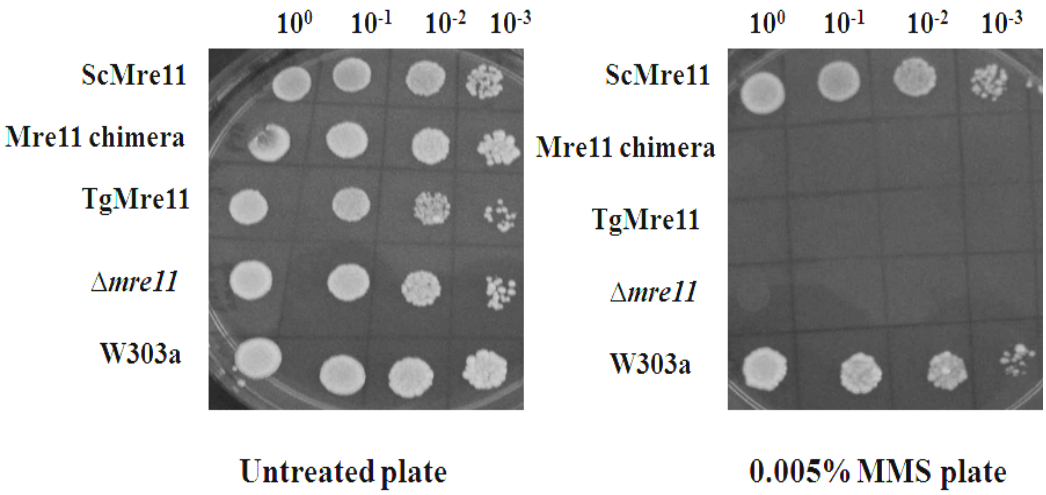

Figure 9. The survivability of different strains under untreated and MMS (0.005%) treated conditions was revealed.

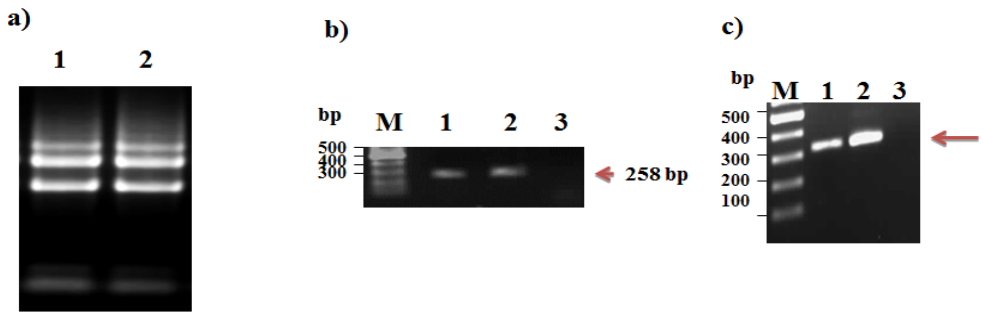

Figure 10. Semi-quantitative RT-PCR from RNA of strains harboring *TgMRE11* or *chimeric MRE11*. **a)** Lane 1 and 2 depicts the isolated RNA from strains containing *TgMRE11* or chimeric *MRE11* respectively **b)** and **c)** Reverse transcriptase PCR of strain containing *TgMRE11* or chimeric *MRE11* produced bands of 258 and 300 bp respectively (corresponding to their 3' transcript) as observed in lane 1. The PCR products from the positive controls observed in lane 2 respectively. Band was not observed in the negative control as observed in lane 3. Standard DNA marker was loaded in lane M of Figure 10b and 10c and the bands are indicated by arrowhead in left panel.

3.2.2 Development of surrogate yeast assay system to study *TgKU80* gene

To understand NHEJ mechanism of *T. gondii*, we choose the central player TgKu80 for characterization. ToxoDB database revealed putative ortholog of *TgKU80* gene from GT1 strain and the full length gene was of size 2817 bp (Table 8). Since we wanted to assay for the functional role of TgKu80 in yeast model system, we looked at the protein sequence identity. When aligned, full length sequences of TgKu80 and ScKu80 proteins revealed only 10% sequence identity.

TgKU80 gene was PCR amplified from the genomic DNA of *Toxoplasma gondii* RH strain. A single band of size 5183 bp was observed (Appendix 1.3). This gene was cloned into pTZ57R/T plasmid. The *TgKU80* gene was confirmed by DNA sequencing. This insert was released and subcloned into the yeast expression plasmid pTA at *BamHI* (616 bp) and *XbaI* (609 bp) sites, under GPD promoter (Appendix 1.3).

For the TgKu80 functional complementation assay in yeast, we included another positive control, wherein *ScKU80* gene was expressed in *Δku80* background. In this regard, *ScKU80* gene was PCR amplified from the genomic DNA of wild type yeast strain. The PCR product obtained was of size 1890 bp (Appendix 1.4). The *ScKU80* gene was first cloned into pTZ57R/T plasmid, the insert was released by *BglII* and *SalI* digestion and finally subcloned into pTA plasmid (Appendix 1.4).

To investigate if *ScKU80* gene can be functionally complemented by *TgKU80*, we have performed NHEJ assay. In the NHEJ assay, pRS426 plasmid was digested with *EcoRI* enzyme and the linear plasmid was transformed into yeast. The two ends of a linear plasmid are regarded as a DSB and are repaired by NHEJ within the cell giving rise to a circularized plasmid. In case the cell fails to repair such breaks, the linear DNA fragments are degraded within the cell. Thus, the ratio of the transformants from

the cut DNA over the uncut DNA gives the efficiency of NHEJ repair. These assays were performed in *Δku80* strain (SSY9), *ScKU80* strain (SSY10) and *TgKU80* strain (SSY11) (Figure 11a). A graph was plotted with the NHEJ efficiency in the different strains of the yeast (Figure 11b). The cell expressing *ScKU80* (SSY10) showed about 100% NHEJ efficiency. The negative control (SSY9) strain showed only 3% NHEJ efficiency. The cell expressing *TgKU80* (SSY11) gave 9.36% as NHEJ efficiency. Since there is no significant difference, we concluded that *ScKU80* cannot be functionally complemented by *TgKU80*. Therefore, yeast model system cannot be used for the study of *TgKU80* gene. We tried to look at the possibility that the TgKu80 protein may not express in a heterologous system. Since we did not have the antibody against TgKu80, we did semi-quantitative RT-PCR and looked for the expression at the RNA level. We isolated RNA from the strain expressing TgKu80 (Figure 11c). We observed a 220 bp band at the 3'end of the RNA transcript of *TgKU80* (Figure 11d). Clearly, the positive control which is the plasmid containing *TgKU80* showed the expected band of 220 bp (Figure 11d). The DNase treated RNA samples were used as the template for PCR and the PCR product was loaded as the negative control. As expected, band was not observed in the negative control (Figure 11d).

Table 8) The gene size and protein length of Ku80 in *Toxoplasma gondii* and *Saccharomyces cerevisiae*:

Gene	Size (bp)	Protein length (aa)	Protein molecular weight (kDa)
TgKU80	**2817**	**938**	**104.17**
ScKU80	**1890**	**629**	**71.24**

- **10 % sequence identity in protein sequences from ClustalW**

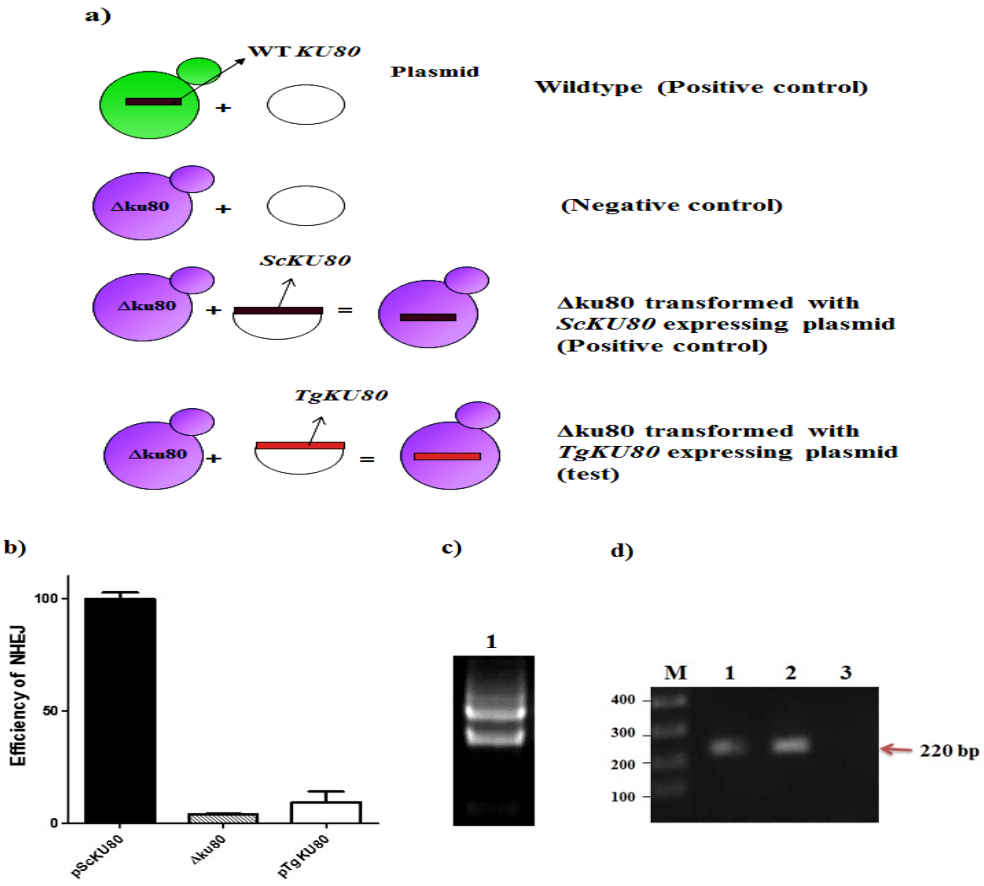

Figure 11. Functional complementation of *TgKU80* in yeast model system. **a)** Schematic diagram depicted all the test and control strains used in the NHEJ assay of *TgKU80* in yeast background. **b)** Graph showed the efficiency of NHEJ of *TgKU80* in yeast assay system. Different strains are indicated on the X-axis and the efficiency of NHEJ depicted on Y-axis. **c)** RNA isolated from the cell harboring *TgKU80* shown in the gel. **d)** RNA transcript of 220 bp from the 3' end of the *TgKU80* gene was observed in lane 1 of the gel. The positive control produced the PCR product of 220 bp as shown in lane 2. Band was not observed in the negative control as shown in lane 3. Standard DNA marker was loaded in lane M of Figure 11d and the bands are indicated by arrowhead in left panel.

3.2.3 Development of surrogate yeast assay system to study *TgRAD51* gene

A BLAST search of ToxoDB database (www.toxodb.org) have identified three ORFs corresponding to putative *RAD51* orthologs from ME49 strain (TGME49_072900), VEG strain (TGVEG_020310) and GT1 (TGGT1_112080) strain. Based on these sequences we have designed PCR primers and amplified TgRad51 ORF from RH strain using cDNA library as template (a kind gift from Dr. Vern Carruthers). Sequence comparison of the amplified ORF indicated that it is an ortholog of Rad51. We found little sequence polymorphism of *RAD51* gene among the four strains of *T. gondii* (Figure 14). The deduced amino acid sequence of TgRad51 protein contains 354 amino acids with a predicted molecular mass of 38,796 and isoelectric point (pI) value of 5.77. Further sequence comparison of TgRad51 protein with other known eukaryotic Rad51 orthologs revealed a high degree of overall identity (52 to 72%) and even higher degree of conservation (65 to 82%) in the core catalytic domain (Table 10). Computer alignment using Clustal method (Meg align, DNA star) revealed a near perfect alignment and identified the two nucleotide binding Walker motifs (Figure 12). Phylogenetic analysis revealed that TgRad51 is closer to another apicomplexan parasite PfRad51 (Figure 13).

The *TgRAD51* gene of 1065 bp (Table 9) was PCR amplified from *T. gondii* cDNA library of RH strain. A single PCR band of *TgRAD51* was observed. It was cloned into pCR 2.1 TOPO vector (3.9 Kb). The *TgRAD51* insert was released by *BamHI* digestion and then cloned into pTA plasmid (6282 bp) at *BamHI* site under GPD promoter (Appendix 1.5).

To investigate the functional role of TgRad51 in HR pathway, we first looked at its functional complementation in yeast surrogate system. To perform this assay, pTA-*TgRAD51* plasmid was transformed into the *Δrad51* strain to generate MVS26 strain.

Wild type strain containing empty pTA plasmid (NRY3) was used as the positive control. Empty pTA plasmid was transformed into *Δrad51* strain to create the negative control (NRY1). Another plasmid pTA expressing *ScRAD51* was also transformed into *Δrad51* strain which served as the positive control, NRY2 (Figure 15a). We performed the MMS sensitivity assay. MMS generates numerous DSBs which if repaired, enabled the cell to survive. So, we tried to look at the percent repair efficiency of TgRad51 using yeast as the surrogate assay system. The data generated from the MMS sensitivity assay was plotted with Graph Pad Prism. The percent repair efficiency of ScRad51 (NRY2) was 90% and that of TgRad51 (MVS26) was 50% (Figure 15b).

Table 9) Primary sequence information of *RAD51* of *Toxoplasma gondii* and *Saccharomyces cerevisiae*:

Gene	Size (bp)	Protein length (aa)	Protein molecular weight (kDa)
TgRAD51	**1065**	**354**	**38.7**
ScRAD51	**1203**	**400**	**42.9**

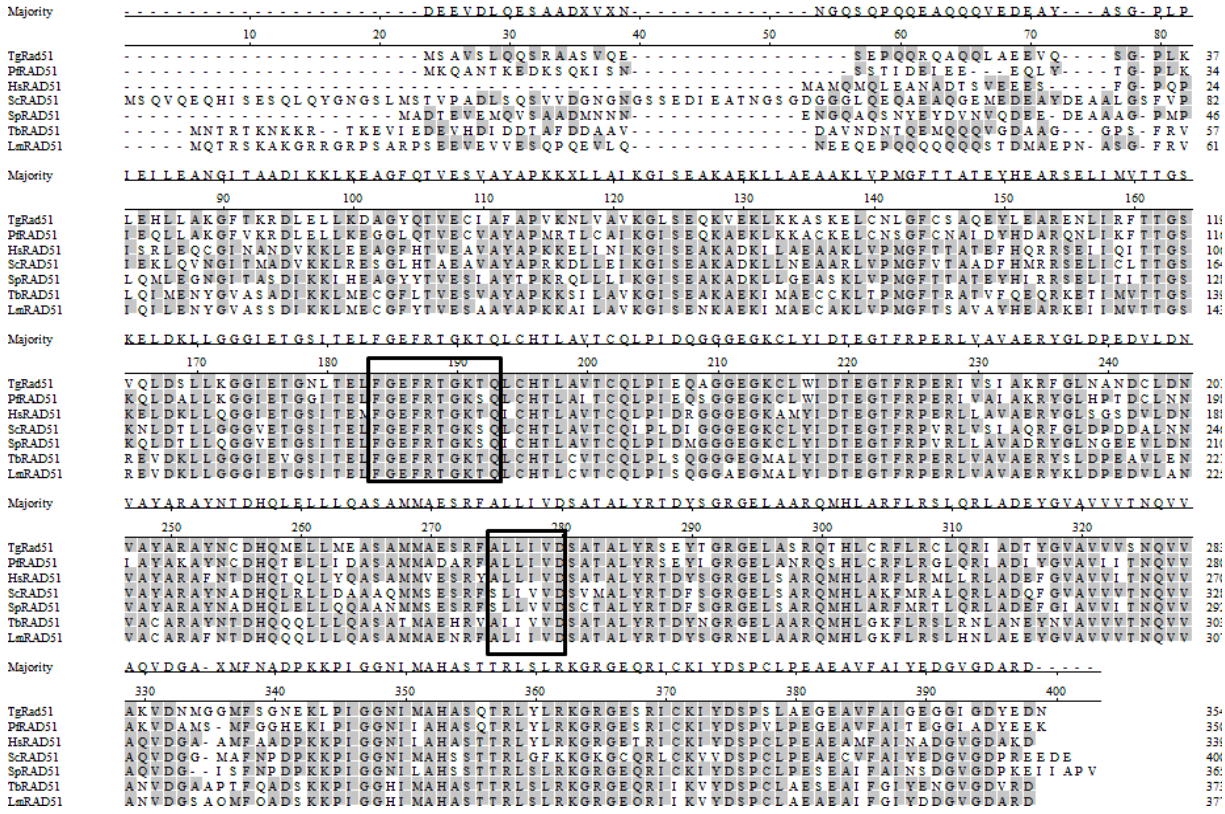

Figure 12. Sequence alignment of TgRad51 protein with other Rad51 proteins: Alignment of TgRad51 with *Plasmodium falciparum* (PfRad51), human (HsRad51), *Saccharomyces cerevisiae* (ScRad51), *Schizosaccharomyces pombe* (SpRad51), *Trypanosoma brucei* (TbRad51) and *Leishmania major* (LmRad51) using Clustal method (Meg align, DNA star). Shaded areas represent identical amino acids. The two DNA binding Walker motifs are boxed.

Table 10) The homology score of TgRad51 with Rad51 proteins in model organisms:

	TgRad51	PfRad51	HsRad51	ScRad51	SpRad51	TbRad51	LmRad51
TgRad51	100	72 (82)	58 (70)	52 (65)	55 (68)	53 (65)	57 (68)
PfRad51		100	60 (71)	55 (64)	56 (68)	52 (63)	54 (66)
HsRad51			100	66 (71)	73 (80)	64 (73)	69 (77)
ScRad51				100	66 (79)	54 (66)	54 (66)
SpRad51					100	56 (68)	58 (69)
TbRad51						100	76 (90)
LmRad51							100

Numbers in parenthesis indicate homology within the catalytic domain of Rad51 proteins

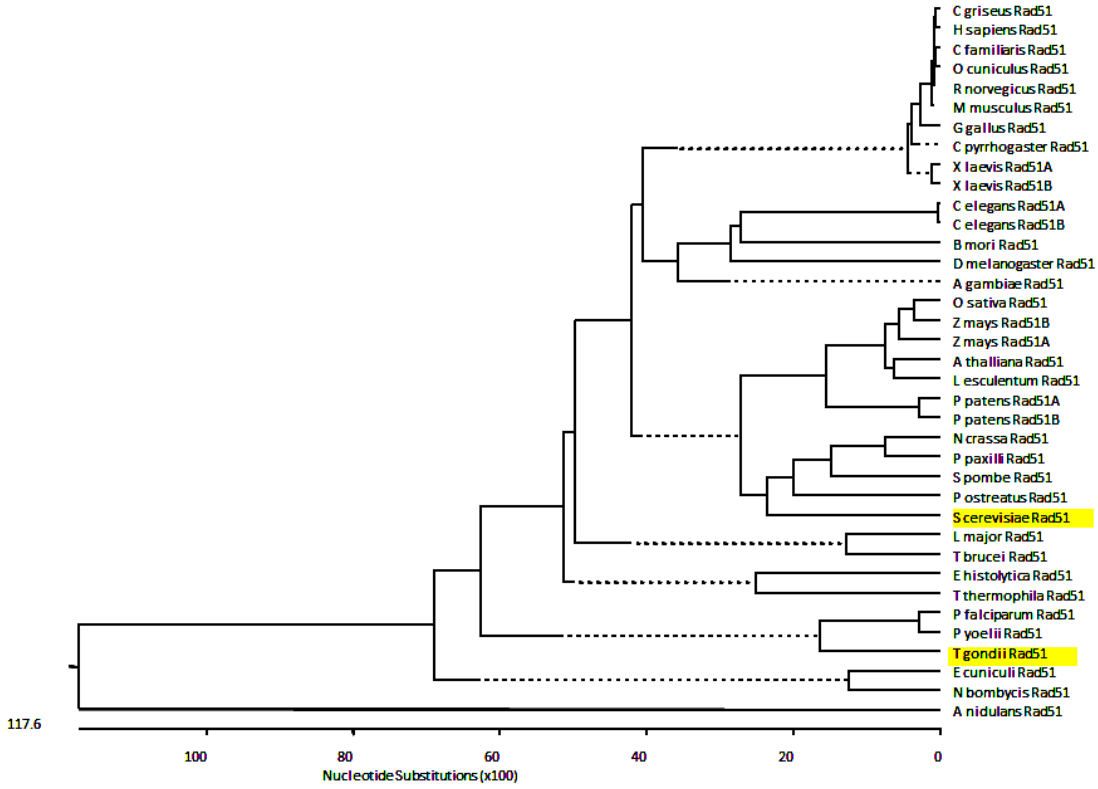

Figure 13. Phylogenetic tree of Rad51 proteins: Phylogenetic analysis of eukaryotic Rad51 proteins using Clustal method (Meg align, DNA star). *T. gondii* Rad51 and *S. cerevisiae* Rad51 are highlighted.

Figure 14. Alignment of Rad51 proteins from the different strains of *T. gondii* – RH, GT1, ME49 and VEG using Clustal method (Megalign, DNAstar).

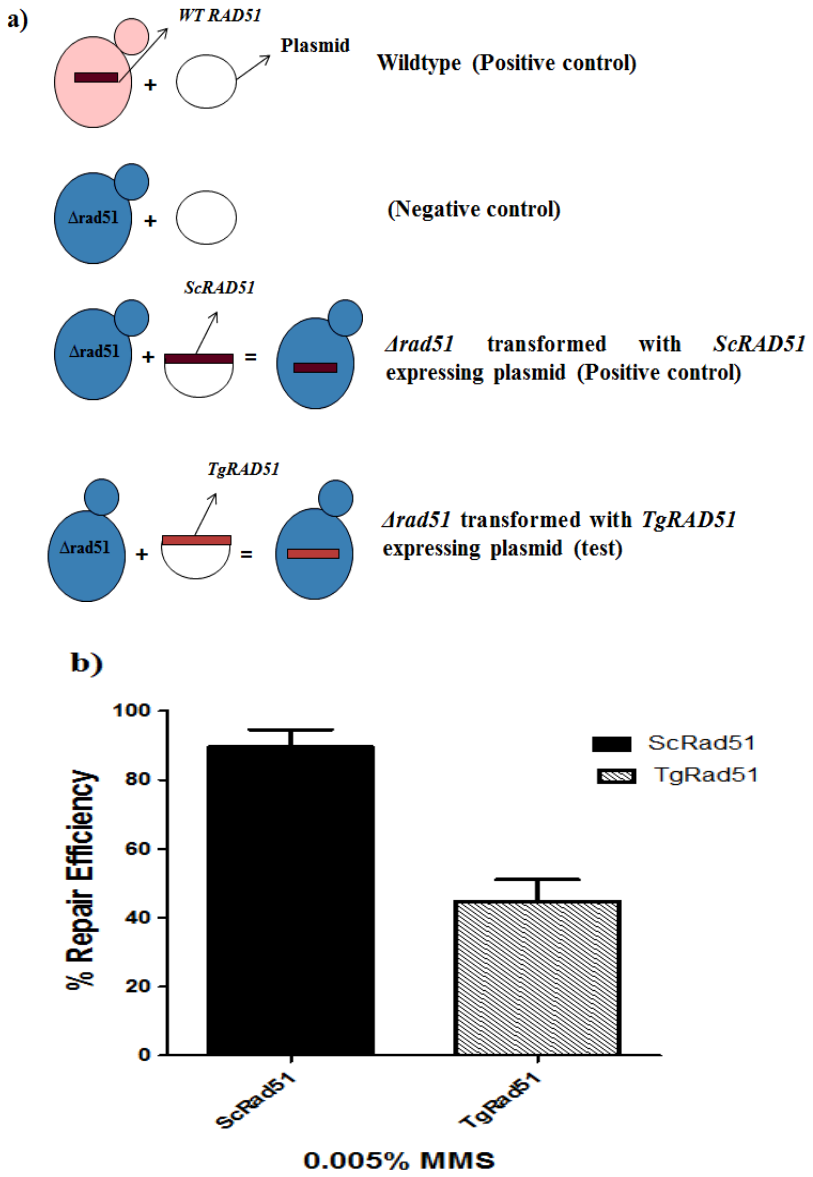

Figure 15. Functional complementation of *TgRAD51* in yeast model. **a)** illustration depicted the strains used for the functional complementation of *TgRAD51* in yeast system by MMS sensitivity assay. **b)** Bar diagram showing the percent repair efficiency in cells harboring ScRad51 or TgRad51.

3.3 Discussion

The *TgMRE11* gene was cloned and expressed in the yeast surrogate system. Chimeric *MRE11* was generated by fusing the N-terminal *TgMRE11* gene with the C-terminal *ScMRE11* gene. This fusion gene was also cloned and expressed in yeast. From the MMS sensitivity assay, it was concluded that *ScMRE11* cannot be functionally complemented by full length *TgMRE11* or chimeric *MRE11*. Hence the yeast surrogate assay system cannot be used for the study of *Toxoplasma gondii* Mre11. In order to study NHEJ pathway of *T. gondii* in yeast system, the *TgKU80* gene was chosen. It was PCR amplified, cloned and expressed in yeast. Our studies demonstrated that *ScKU80* gene cannot be functionally complemented by full length *TgKU80*. By semi-quantitative RT-PCR, we found that *TgMRE11*, chimeric *MRE11* and *TgKU80* were expressed in yeast.

To understand HR and the efficiency of gene targeting in *T. gondii*, we did functional complementation of ScRad51 with TgRad51. Return to growth experiments with MMS as the DNA damaging agent clearly exhibited that *ScRAD51* was functionally complemented by *TgRAD51*. From this part of the study, we concluded that the surrogate yeast assay system was developed for the study of TgRad51 but it cannot be used for the study of TgMre11, chimeric Mre11 and TgKu80.

CHAPTER 4

Biochemical characterization of *TgRAD51*: an implication for inefficient gene targeting

4.1 Introduction

Rad51 protein has ATP dependent DNA binding activity, it multimerizes on single stranded DNA to form helical filament similar to that formed by bacterial RecA protein (Ogawa T et al., 1993). In an ATP hydrolysis dependent manner its motor activity searches for the homologous sequences between a single strand DNA and double stranded DNA and catalyses the strand exchange reaction. During this process it interacts with replication protein A (RPA), Rad52, Rad54 and Rad55 (Jiang H et al., 1996; Sung P., 1994; Hays SL et al., 1995).

The Ku70/80 heterodimer binds at the DSB along with DNA ligase IV-Xrcc4 complex and DNA-PKc to catalyze the break repair (Walker JR et al., 2001). In a ku80 knock out background there is a 300-400 fold increase in targeted gene disruption in *T. gondii* (Fox BA et al., 2009). The key protein, TgRad51, involved in targeted gene disruption has not been characterized yet. We have cloned, purified and characterized TgRad51 biochemically as well as genetically. It is observed that *T. gondii* and *Plasmodium falciparum* show strikingly opposite choice in DNA repair pathways. While *Plasmodium falciparum* depends solely on HR and apparently lacks NHEJ, *T. gondii* prefers NHEJ. Here we report the mechanistic insights for differential repair choices between these two closely related lower eukaryotes. PfRad51 has been identified as a DNA repair protein and has been speculated to play major role during mitotic recombination (Bhattacharyya MK et al., 2003; 2004). PfRad51 has also been characterized biochemically (Bhattacharyya MK et al., 2005) and despite having 82% identity in the catalytic domain, the kinetics of ssDNA dependent ATP hydrolysis activity differs markedly between PfRad51 and TgRad51. We hypothesize that compromised ATPase activity of TgRAD51 leads to inefficient gene targeting in *T. gondii*.

4.2 Results

4.2.1 Expression and purification of recombinant TgRad51 protein

The expression system Rosetta (DE3)/pET:*TgRAD51* was used for the purification of TgRad51 protein. In the BL21plysS background, upon induction, no induced protein band was observed (data not shown). The expression vector pET:*TgRAD51* was transformed into Rosetta (DE3) cells, which has the pRARE plasmid that supplies the tRNA for six codons that are rarely used in *E. coli* (AUA, AGG, AGA, CUA, CCC and GGA). The whole transformation mixture was added to 10 ml LB media containing chloramphenicol and kanamycin and incubated at 37°C for overnight. Next morning 5% of the primary culture was added to 100 ml fresh LB media (containing chloramphenicol and kanamycin) and tested for the IPTG induced expression of the TgRad51 protein. Under T7 promoter, a strongly induced band of His-tagged TgRad51 protein at 41.97 kDa size was observed. Figure 16a (lane 2) shows that TgRad51 protein corresponds to about 25% of the total proteins in the crude cell extract after IPTG induction. The cells were collected by centrifugation and lysed by sonication. The induced recombinant protein was predominantly detected in the soluble fraction (lane 3). Figure 16a shows the purification profile of recombinant TgRad51 protein through Ni-NTA column. Glycerol (30%) was included in the wash buffer in order to remove the unwanted hydrophobic interactions between nonspecific proteins attached with TgRad51. His tag is uncharged and it allows immobilization of protein on metal chelating substances. NTA has six ligand binding sites four with coordination sphere of Ni ion and two with His tag. His tagged protein is eluted with increasing concentration of imidazole. Competition is created between His tag and imidazole to bind with Ni-NTA column. The elution at 250 mM imidazole resulted in about 90% purified protein with minor protein contaminants (lane 6). However, 400

mM imidazole containing elution buffer resulted in greater than 99% purified protein as judged by SDS-polyacrylamide gel electrophoresis (lane 7). Finally the purified protein was dialyzed thoroughly against Tris-HCl (pH=8) to remove excess imidazole.

Mass spectroscopic analysis confirmed the presence of TgRad51 protein and yielded a molecular mass of 39,398 Da in agreement with the molecular mass of 38,796 Da predicted by ExPasy Protparam tool. A single peak was obtained which indicated that there was no other contaminating protein in the purified sample. MS-MS analysis yielded sequence of 5 peptides corresponding to the predicted amino acid sequence of TgRad51 (Figure 17). The western blot analysis with anti histidine antibody confirmed the presence of His-tagged TgRad51 (Figure 16b).

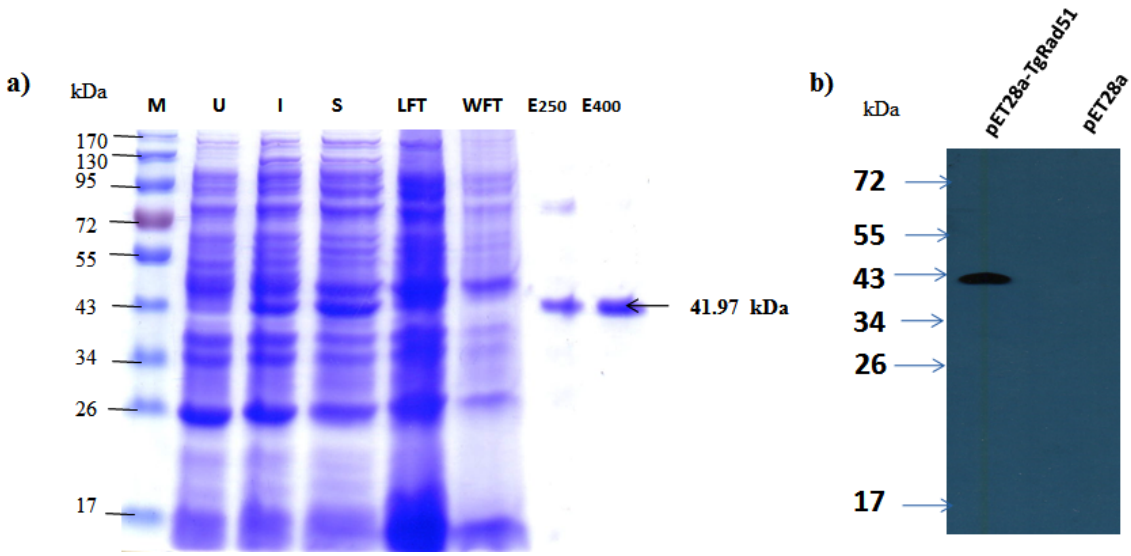

Figure 16. Purification of recombinant TgRad51 protein and western blot analysis. a**)** M corresponds to molecular weight standards, the sizes of marker proteins in kDa are indicated, Lane 1 and 2 are uninduced and induced cell free extracts respectively, Lane 3 corresponds to the proteins in soluble fraction, Lane 4 being the loading flow through and Lane 5 is the washing flow through. Lane 6 and 7 correspond to the eluted fractions with increasing imidazole concentrations i.e. 250mM and 400mM respectively. **b)** Western blot analysis shows the presence of Histidine tagged TgRad51 in lane 1 from the cells bearing pET-*TgRAD51* and lane 2 shows the absence of the protein in the cell bearing the empty vector pET.

>TgRad51 protein sequence from RH strain

```
  01 MSAVSLQQSRAASVQESEPQQRQAQQLAEEVQSGPLKLEHLLAKGFTKRD
  51 LELLKDAGYQTVECIAFAPVKNLVAVKGLSEQKVEKLKKASKELCNLGFC
 101 SAQEYLEARENLIRFTTGSVQLDSLLKGGIETGNLTELFGEFRTGKTQLC
 151 HTLAVTCQLPIEQAGGEGKCLWIDTEGTFRPERIVSIAKRFGLNANDCLD
 201 NVAYARAYNCDHQMELLMEASAMMAESRFALLIVDSATALYRSEYTGRGE
 251 LASRQTHLCRFLRCLQRIADTYGVAVVVSNQVVAKVDNMGGMFSGNEKLP
 301 IGGNIMAHASQTRLYLRKGRGESRICKIYDSPSLAEGEAVFAIGEGGIGD
 351 YEDN
```

Figure 17. Amino acid sequence of TgRad51 protein from RH strain. The underlined peptide sequences (red) were generated from MS-MS analysis.

4.2.2 ssDNA dependent ATP hydrolysis of TgRAD51

The concentration of recombinant His-tagged TgRad51 protein was determined by Bradford assay. The increase of absorbance at 595 nm was proportional to the amount of the bound dye and thus to the amount of the protein present in the sample. Different concentrations of the known protein standard, Bovine serum albumin (BSA) protein was used and the standard curve was plotted (Figure 18a). The concentration of the purified recombinant TgRad51 protein was determined from the standard graph.

In order to investigate whether the low gene targeting efficiency in *T. gondii* is due to a compromised Rad51 protein activity, we determined the ATP hydrolysis activity of purified TgRad51 protein. We used EnzChek Phosphate Assay kit (Molecular Probes) to determine the rate of ATP hydrolysis of TgRad51. In this assay the inorganic phosphate when combines with the substrate MESG (2-amino-6-mercapto-7-methylpurine riboside) in presence of the enzyme PNP (purine nucleoside phosphorylase) produces ribose 1-phosphate and 2-amino-6-mercapto-7-methylpurine (Figure 18b). Formation of the product changes the absorbance of the substrate from 330 nm to 360 nm which can be measured spectroscopically. Figure 19a shows the steep increment of the absorbance at 360 nm with increasing concentration of the inorganic phosphate. The slope of the graph with the different concentrations of standard phosphate on X-axis and corresponding absorbance values on Y-axis was determined to be 0.013. When we incubated TgRad51 with ATP at 37^{o}C in presence of φxssDNA, ATP is hydrolyzed to release phosphate. At different time interval aliquots of this reaction mixture were withdrawn and added to a second reaction mixture containing MESG along with PNP. This second reaction mixture was incubated at 22°C. Accordingly we monitored the kinetics of ATP hydrolysis of

TgRad51 protein in presence of ssDNA. Figure 19b shows that TgRad51 is incapable of ATP hydrolysis in absence of ssDNA. However in presence of 30 fold molar excess of ssDNA, ATP hydrolysis took place. At 300 μM ATP concentration, the rate of ATP hydrolysis as calculated from the slope of the curve was found to be 0.027 μM/ min. The ATPase activity of TgRad51 was measured at different ATP concentrations (20 to 600 μM) and used to plot Michaelis Menten curve (Figure 20). The data was also used to generate Lineweaver Burk plot (Figure 20). The k_{cat} value ($0.034 min^{-1}$) and K_m value for TgRad51 (81.53 μM) were calculated using Graph Pad Prism. The michaelis-menten constant, K_m of TgRad51 was compared to PfRad51 protein which was 28.37 uM (Bhattacharyya M K et al, 2005). It revealed that TgRad51 protein had much lower affinity to substrate ATP than PfRad51.

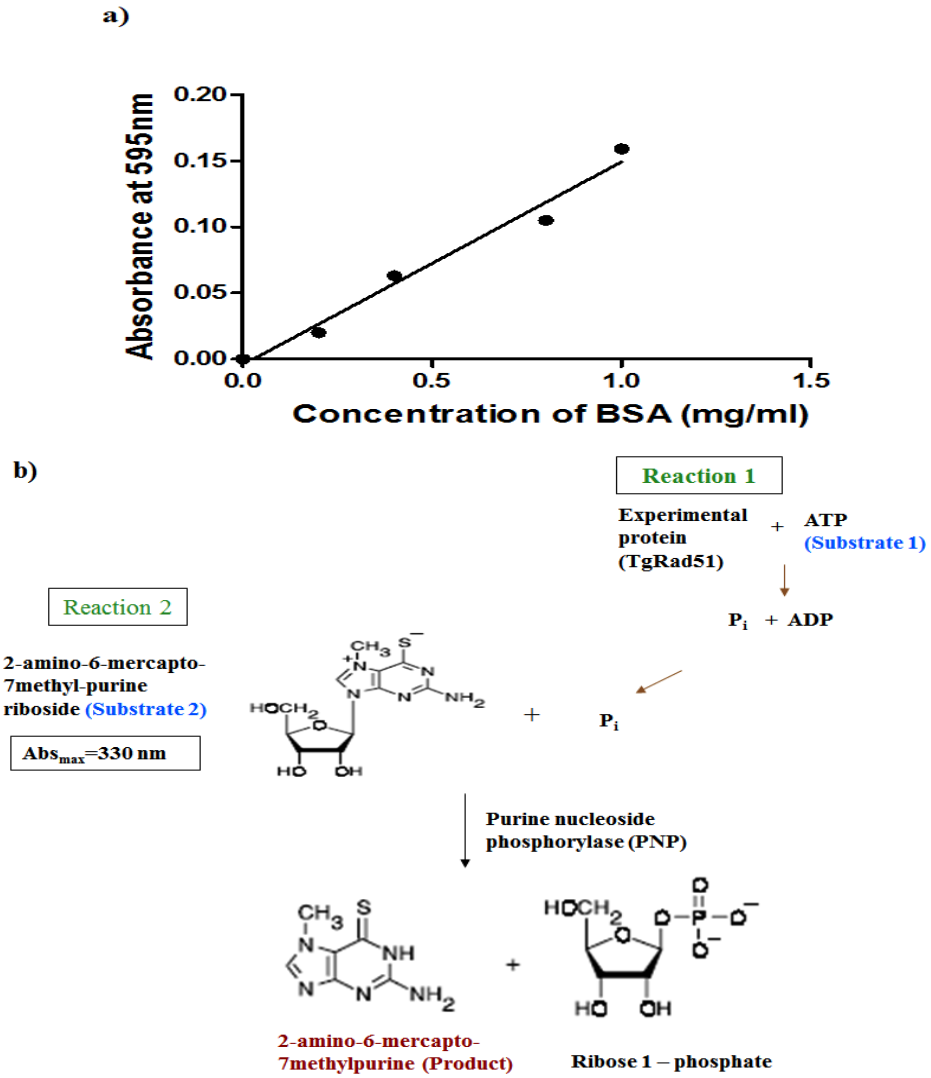

Figure 18. ATP hydrolysis reaction of TgRad51 protein. **a)** Concentration of the purified TgRad51 protein was determined by Bradford assay using BSA as the standard. A linear increase was observed in the graph. **b)** In ATPase assay, TgRad51 protein reacts with the first substrate ATP to liberate inorganic phosphate by ATP hydrolysis. This phosphate reacts with another substrate MESG (2-amino 6-mercapto 7-methyl-purine riboside) to form 2-amino-6-mercapto-7 methyl purine as the product. This reaction is catalysed by purine nucleoside phosphorylase (PNP). There is a shift in maximum absorbance from 330 to 360 nm when substrate was converted to product.

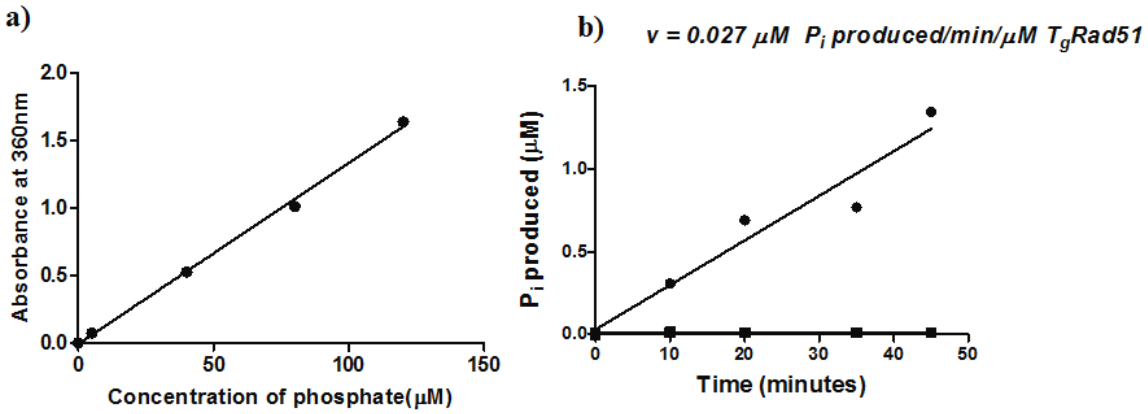

Figure 19. Standard curve and ssDNA dependent ATP hydrolysis of TgRad51. **a)** The standard curve showing linear relationship between inorganic phosphate added and increase in absorbance at 360 nm in response to the addition to MESG as substrate and the enzyme PNP (EnzChek phosphate assay kit). **b)** ssDNA dependent ATP hydrolysis of TgRAD51 using EnZChek phosphate assay kit. At 300 μM ATP concentration, with 2 μM TgRAD51 and 60 μM φxssDNA, the rate of the reaction as calculated from the slope of the curve is 0.027μM P_i produced/min (•) where as in absence of ssDNA (■) no ATP hydrolysis occurs.

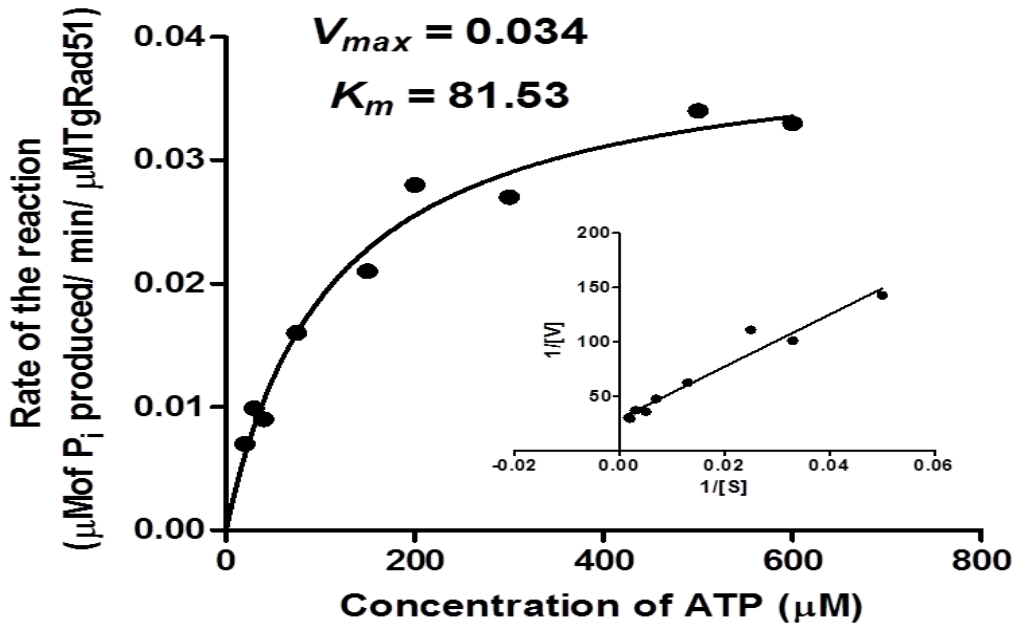

Figure 20. Michaelis Menten curve of TgRad51. Graph plotted with ATP concentrations in the range of 20 μM to 600 μM and kinetic parameters are derived from the curve.

Table 11) The catalytic ability constants and gene targeting efficiency of Rad51 proteins from prokaryotes to eukaryotes were compared.

Protein	k_{cat} values (min^{-1})	Gene targeting efficiency
EcRecA	18	High
LmRad51	2.9	High
ScRad51	0.7	High
HsRad51	0.16	Low
PfRad51	0.082	Low
TgRad51	0.034	Low

4.3 Discussion

Double strand breaks (DSBs) can cause damage to the genomic integrity of a cell. Repair of such DSBs are essential for cell survival. We have cloned, expressed and purified recombinant TgRad51 and have shown that it possesses weak ATP hydrolysis activity (k_{cat} is 0.034 min^{-1}). TgRad51 is the first member of the recombination machinery of *T. gondii* to be characterized. The catalytic ability constants of PfRad51, ScRad51 and LmRad51 were 2, 20 and 85 times more than TgRad51 protein. These results clearly indicated that the ATP hydrolysis activity of TgRad51 protein is very much compromised. Low ATPase activity meant that the targeted gene integration was lower at right locus. A strong ATPase activity of Rad51 is an indicative of a robust HR system. This notion is supported by studies with mutant hRad51 (K133R) incapable of ATP hydrolysis, where gene targeting is severely compromised. However, this mutant does not have any effect on strand exchange between homologous templates. Moreover, the DNA repair ability of the mutant cell is not affected (Morrison C et al., 1999).

T. gondii despite being a lower eukaryote shows preference to NHEJ pathway. TgRad51 possesses weak ATP hydrolysis activity which fits very well with the choice of DSB repair pathway in this organism. Although, it appears that HR might not be a general mechanism for repair of DSB in *T. gondii*, whether HR activity is required to repair specific types of DSB ends remains as an open question. It is likely that the primary function of HR in *T. gondii* could be during sexual reproduction.

CHAPTER 5

Genetic characterization of *TgRAD51*: an implication for inefficient DNA repair and gene targeting

5.1 Introduction

Depending upon the nature and orientation of the homologous sequences with respect to the break, a DSB can be repaired by any of the three different HR pathways: the single strand annealing (SSA), break induced replication (BIR) and the gene conversion (GC), also known as synthesis dependent strand annealing (SDSA) (Haber JE et al., 2000; Jain S et al., 2009; Agmon N et al., 2009). Presence of homology at both the DSB ends leads to GC. When homology is present only at one DSB end, BIR follows and when a DSB is flanked by direct repeats, SSA is the mechanism of choice. In the case of SSA, a minimum of 30 base pair homologous sequence on either side of the DSB is sufficient and such repair leads to the deletion of intervening sequences. All of the above mentioned pathways are driven by several proteins that belong to Rad52 epistatis group (Symington LS., 2002). Strand invasion is the central step of HR mechanism and it requires Rad51 protein. However, Rad51 is essential only for GC, not for BIR or SSA (Paques F et al., 1999; Szostak JW et al., 1983). Efficient gene targeting is a challenge in *T. gondii*. It has been observed that in this parasite gene targeting efficiency is very low as they demonstrate high degree of non homologous end joining (Donald RG et al., 1998). It is interesting to note that *T. gondii* is the only protozoan parasite that harbors non homologous end joining mediated DNA break repair mechanism. In a ku80 knock out background, gene replacement efficiency was markedly increased in *T. gondii* and target DNA flanks of only 500 bp were sufficient for efficient gene targeting (Fox BA et al., 2009). We hypothesize that compromised ATPase activity of TgRAD51 leads to inefficient gene targeting and poor gene conversion efficiency in *T. gondii*. With increase in

homologous flanking ends, we also observed an increase in targeted gene integration similar to the trend observed with ScRad51.

5.2 Results

5.2.1 Gene conversion efficiency of TgRAD51 is poor compared to that of ScRAD51

The ATP hydrolysis activity of Rad51 protein is essential for its motor activity, which is instrumental for the genome-wide search for a homologous template during the pre-synapsis stage of homologous recombination. In order to investigate how the weak ATP hydrolysis activity of TgRad51 affects homology search, we have assayed gene conversion (GC) efficiency of TgRad51. Repair of a DSB by GC relies on Rad51 mediated homology search. To this end we have used yeast as a surrogate model system where GC could be assayed accurately and quantitatively. We used a yeast strain NA14 having two copies of *URA3* gene in chromosome V with a *KANMX6* selectable marker in between (Figure 21a). One of the *ura3* genes is made nonfunctional by incorporating HO endonuclease site in it. The functional *URA3* gene is present in the same chromosome about 3 kb apart. The strain harbors *HO* endonuclease from GAL1 promoter, which is activated by growing the cells in presence of galactose as a carbon source. When a DSB is generated in *ura3* gene by induction of *HO* endonuclease it can be repaired either by GC using the homologous template (*URA3*) present in the same chromosome 3 kb apart, or it can be repaired by single strand annealing (SSA), which results in deletion of a stretch of intervening DNA sequence. We created two strains SSY1 and SSY2 by transforming the expression vector pTA*ScRAD51* and pTA*TgRAD51* respectively to *NA14*Δ*rad51* strain. In order to assay the efficiency of DNA repair activity of *TgRAD51* we created a single DSB in the *ura3* gene by HO induction and counted the number of cell

survival on galactose containing plate and compared to the number of cells on glucose containing plate. Our result shows that the repair efficiency of TgRAD51 (83%) is comparable to that of the wild type (96%), while *NA14*Δ*rad51* being 63% (Figure 21b). In order to monitor the contribution of GC among the survivors, we looked for G418 sulphate resistant cells. If the DSB is repaired by GC, the survivor colonies will be G418 resistant. On the other hand if the break is repaired by SSA, the survivor colonies will be G418 sensitive due to the loss of *KANMX6* cassette. We found that the majority of the survivors (77%) of SSY2 strain (TgRad51) are due to repair by SSA, which does not require the presence of Rad51 protein, gene conversion efficiency being only 23% of the total. On the contrary SSY1 strain (ScRad51) exhibited almost equal preferences for GC and SSA pathways, i.e. approximately 42% of its survival by GC and 58% by SSA (Figure 21b). Hence the GC efficiency of SSY1 (ScRad51) strain is about two fold more than that of SSY2 (TgRad51) strain. The key protein involved in homologous recombination mediated gene conversion is the Rad51 and our result shows that TgRad51 possess inefficient gene conversion activity compared to ScRad51. To rule out the possibility that inefficient gene conversion is not due to poor expression of TgRAD51 in heterologous system, we have done Western blot analysis of SSY1 and SSY2 strains and probed with anti ScRAD51 and PfRAD51 antibodies respectively (anti-PfRad51 antibody cross-reacts with TgRad51 protein). Our data (Figure 21c) shows that TgRad51 expression is comparable to that of ScRad51, actin being the loading control. We conclude that low ATPase activity of TgRad51 is responsible for poor gene conversion efficiency.

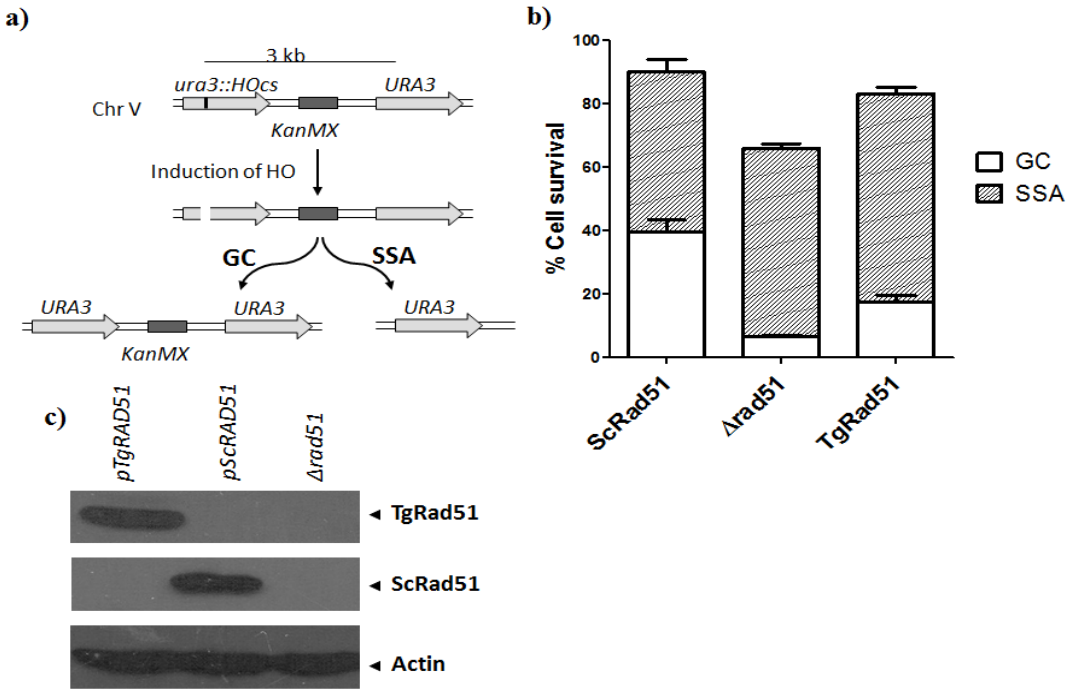

Figure 21. Gene conversion efficiency of *TgRAD51* is poor compared to that of *ScRAD51*. **a)** Schematic diagram of DSB repair choice experiment. *URA3* and *KANMX* represent wild type alleles. A HO endonuclease site is incorporated in *ura3::HOcs* mutant allele. Once the HO induced DSB is repaired by gene conversion (GC), the mutant *ura3::HOcs* allele is converted into wild type *URA3* allele. A single strand annealing (SSA) event leads to deletion of the intervening sequence (containing *KANMX* gene) and merging of *ura3::HOcs* and *URA3* alleles to create the wild type *URA3* allele. **b)** Bar diagram showing the percentage of cells survived after induction of DSB. The relevant genotypes are marked on the X-axis. The white bars represent fraction of the cells survived by repairing the DSB using GC mechanism, where as the hatched bars denote the fraction of the survivors that employed SSA mechanism. Each bars represent mean value ± SD from 4 different experiments. **c)** Western blots showing the abundance of ScRad51 and TgRad51 proteins. Different lanes are marked with the respective genotypes. Actin is the loading control.

5.2.2 Gene targeting efficiency of TgRAD51 increases with increase in stretch of homologous sequences

A linear piece of DNA bearing homology to a part of the genome can either be integrated at the correct chromosomal locus by means of homologous recombination, or it could be integrated at any random site on the chromosomes via non-homologous recombination. In *T. gondii* targeted integration is extremely less efficient. We wanted to investigate whether such less efficient gene targeting is due to the weak ATP hydrolysis activity of TgRad51, since gene targeting at the correct locus demands search of two homologous needles in the genomic hay stack. In order to appreciate the contribution of TgRad51 protein alone, from other *Toxoplasma* proteins presumably involved in HR pathway, we have used a surrogate yeast model, where the yeast *RAD51* gene is deleted and is replaced by *TgRAD51* gene. Thus, in this system, the main recombination protein (Rad51) is from *T. gondii* and all other auxiliary proteins are from yeast origin. We have engineered this system in such a way that it would lead us to investigate the frequency of targeted vs. random integration choices in the cell harboring TgRAD51. The integration construct contains two selectable markers: *ADE2* and *KANMX6*; while the *ADE2* marker scores for all the integration events (targeted or random), the *KANMX6* marker is able to differentiate between random versus targeted integration. In case of targeted integration at the *ADH4* locus the *KANMX6* marker will be lost as a result of homologous recombination. On the other hand, during NHEJ mediated random integration the *KANMX6* cassette will be retained and such transformants will be G418 resistant. Thus random integration of the cassette would result in $ADE2^{+}$ and $G418^{R}$ colonies, whereas the targeted integration would lead to $ADE2^{+}$ and $G418^{S}$ colonies (Figure 22a).

We transformed the KANMX-ADH4-ADE2-ADH4 cassette in three strains: NRY1 (Δ*rad51* containing an empty vector pTA); NRY2 strain (Δ*rad51* containing an expression vector pTA: *ScRAD51*); and MVS26 strain (Δ*rad51* containing an expression vector pTA: *TgRAD51*). The experiments with three repeats showed that NRY2 cells (*ScRAD51*) prefer to go for targeted integration at the ADH4 locus with about 85% frequency, whereas *TgRAD51* showed a preference for random integration having only 24% frequency for targeted integration (Figure 22b). We hypothesize that poor ATP hydrolysis activity of TgRad51 results in inefficient motor activity of the TgRad51 protein, as a result of which it shows three times reduction in gene targeting efficiency compared to that of *ScRAD51*. A prediction of this hypothesis would be that the increment of flanking homologous sequences would facilitate the homology search and thus would compensate for the weak ATP hydrolysis of Rad51. So, the next question we asked was whether the increase in homologous stretch of DNA has any effect on the frequency of TgRad51 protein mediated targeted integration. To test this hypothesis we designed three constructs to knockout *SBA1* gene. These constructs contains a selectable marker *KANMX6* flanked by upstream and downstream sequences (200 bp/ 500 bp/ 1000bp on either side) of *SBA1* gene (Figure 23a). These three constructs were used to disrupt the *SBA1* gene in the cells bearing *ScRAD51* (NRY2) and *TgRAD51* (MVS26). As expected, our result showed that with the increase in homologous sequence there was a gradual increase in the frequencies of the targeted gene disruption (Figure 23b). NRY1 (*Δrad51)* cells showed very less number of targeted disruptions as measured by the survival on G418 sulphate plates throughout varied degree of homology. With increase in flanking homologous sequences, NRY2 cells (*ScRAD51*) showed a steep rise in the gene targeting efficiency (Figure 23b). MVS26 (*TgRAD51*) cells behaved much like *Δrad51* strain

up to 500 base pair flanking homology. As the homologous stretches were increased to 1000 base pairs on either side of *KANMX6*, the number of $G418^R$ colonies also increased sharply for MVS26 strain. The increase in homologous flanking region causes same fold increase (approximately 2.5 fold) in the gene targeting efficiency for both *ScRAD51* and *TgRAD51*.

Works from Bzik laboratory and Carruthers laboratory (Fox BA et al., 2009; Huynh MH et al., 2009) have demonstrated enhanced gene targeting in *ku80* null *T. gondii.* However, it would not increase the integration efficiency at the correct locus. This notion is supported by the finding that targeted repair of Δ*hxgprt* became independent of TgKu80 when enough flanking homology (910 bp) was provided (Fox BA et al., 2009). To test this, we monitored the gene targeting efficiency in *Δrad51* and *Δrad5Δku80* background. Two linear constructs were designed to target *CHL1* gene locus. We genetically engineered 500/1000 bp upstream and downstream of *CHL1* gene on either side of *HIS* marker (Figure 24a). These two constructs were used to disrupt the *CHL1* gene locus in the cells bearing ScRAD51 (SSY5) or TgRAD51 (SSY6). The absence of Ku80 in yeast assay system did not increase the efficiency of TgRad51 recombinase during gene targeting. Similar results were observed in efficiency of gene targeting in *Δrad51* and *Δrad5Δku80* backgrounds (Figure 24b).

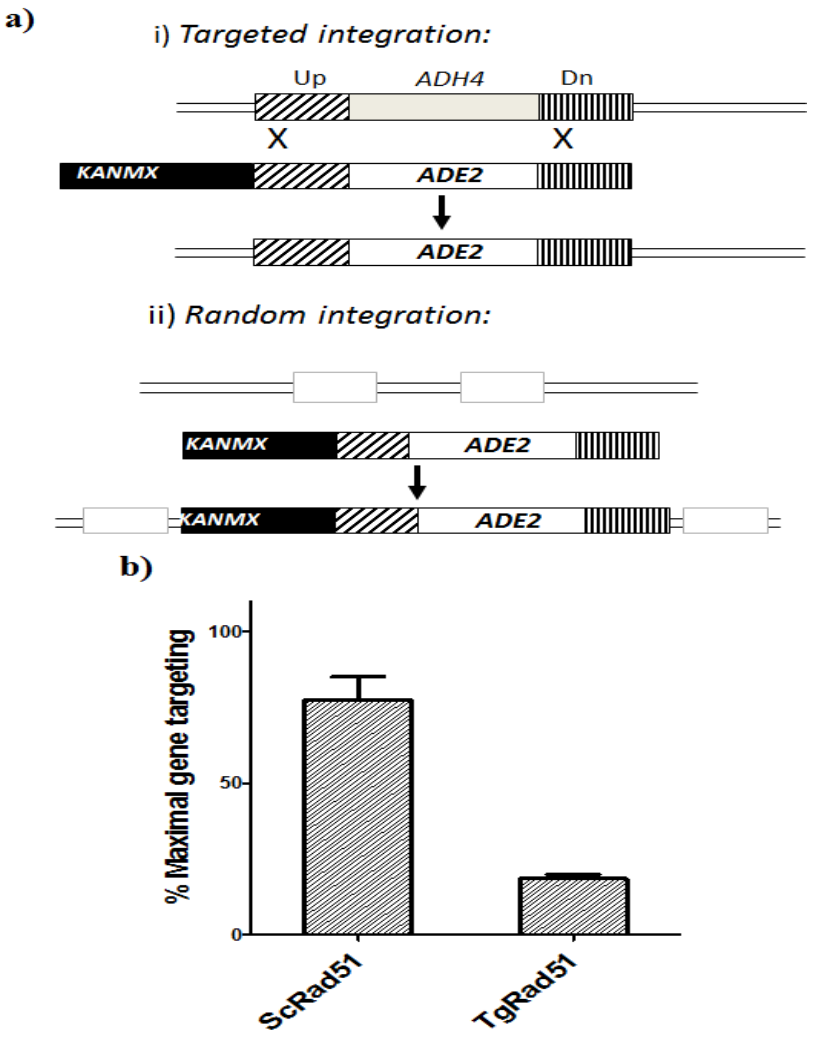

Figure 22. Gene targeting efficiency of TgRad51. **a)** It increases with increase in stretch of homologous sequences. Schematic diagram showing molecular events leading to targeted integration of *ADE2* gene at *ADH4* locus versus random integrations. *KANMX* is a second selectable marker retained only in case of random integration via non-homologous recombination mechanism. **b)** Bar diagram showing efficiency of gene targeting at the *ADH4* locus in cells harboring ScRad51 or TgRad51 as the sole recombinase. The mean value from four independent experiments is plotted with standard deviations.

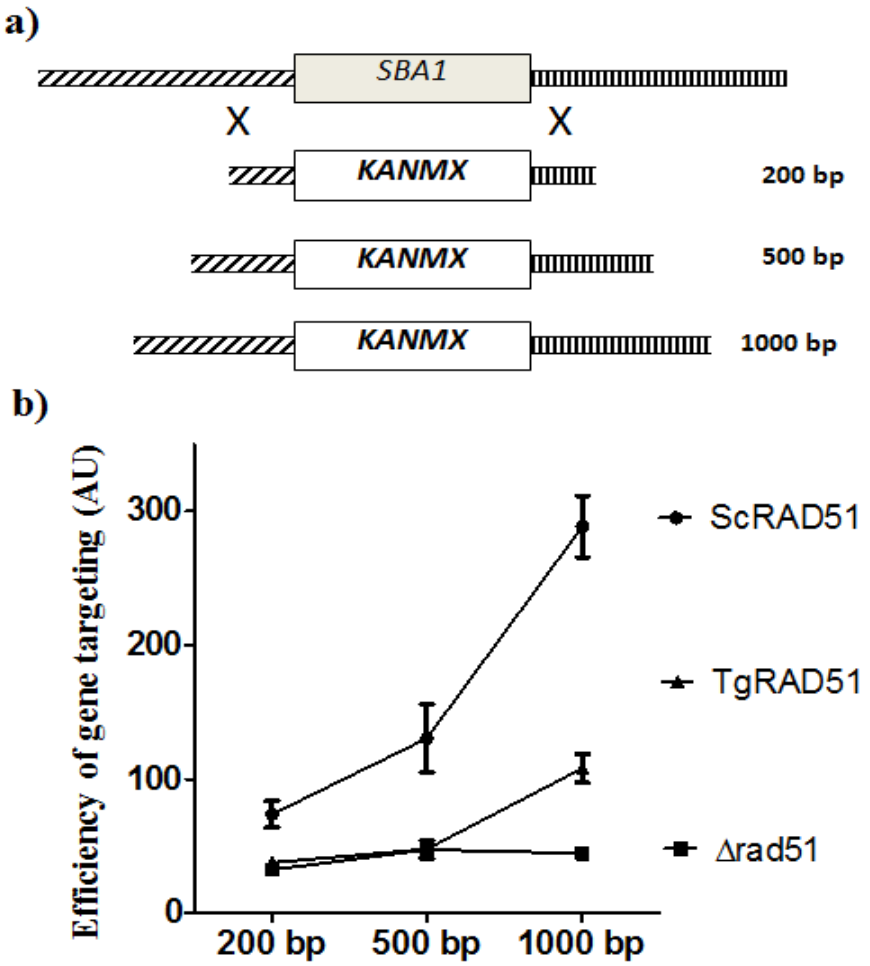

Figure 23. Gene targeting efficiency of TgRad51 with increase in homology length. **a)** Schematic diagram showing knockout strategy for *SBA1* gene. Varying lengths of flanking homologous sequences on either side are indicated. **b)** Efficiency of gene knockout with increasing flanking homology. The lengths of the flanking homologous stretches are indicated on the X-axis. These experiments are done at least three times and the mean values with standard deviations are plotted.

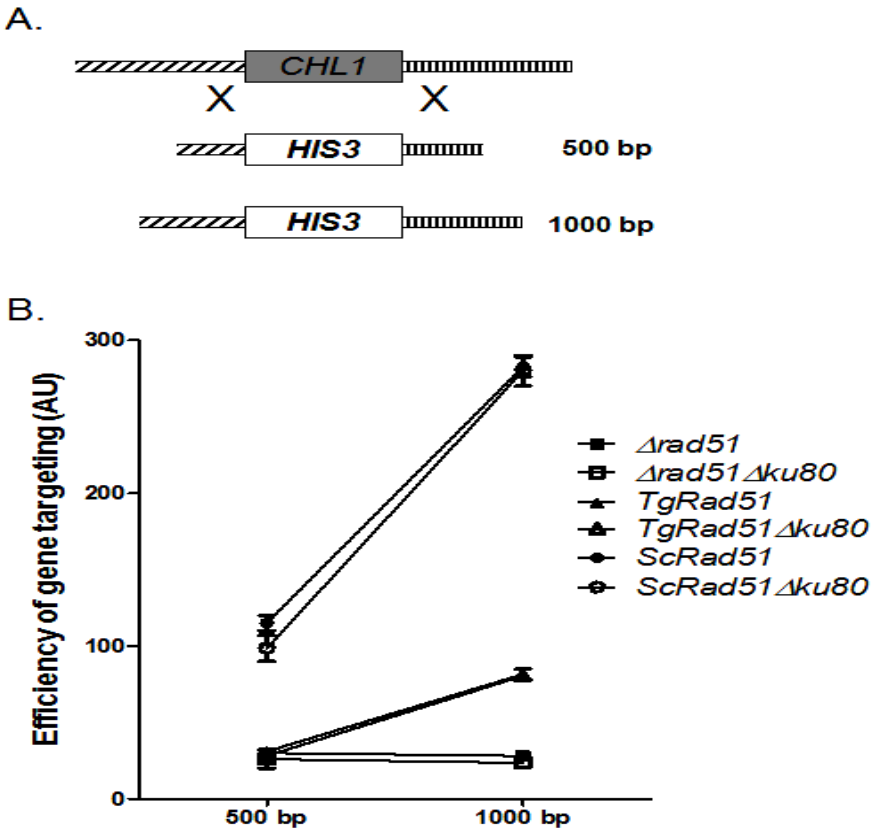

Figure 24. Gene targeting efficiency of TgRad51 is independent of Ku80 function. **a)** Schematic diagram showing knockout strategy for *CHL1* gene. Varying lengths of flanking homologous sequences on either side are indicated. **b)** Efficiency of gene knockout with increasing flanking homology in KU80 proficient (closed- circle, triangle or square) and KU80 deficient (open-circle, triangle or square) cells. The lengths of the flanking homologous stretches are indicated on the X-axis. These experiments are done at least 3 times and the mean values with standard deviations are plotted.

5.3 Discussion

Using yeast as a surrogate system we have characterized *TgRAD51* genetically. Three experiments: repair choice experiment, targeted gene knock-in experiment and targeted knock-out experiment were performed. In the repair choice experiment, the induced DSB created within *URA3* gene is repaired by any of the two HR mechanisms: gene conversion (GC) or single strand annealing (SSA). While SSA does not depend on Rad51, GC totally depends on Rad51 mediated homology search. Our results revealed that GC frequency is very low in cells containing TgRad51.

Two other assays (namely gene knock-in and gene knock-out) that also depend on Rad51 mediated homology search are also found to be affected by TgRad51. An organism that harbors these two competing pathways (HR and NHEJ) may results in either random integration of the transfected DNA *via* NHEJ mechanism or integration at the right locus *via* HR mechanism. Only one of the two types of recombinant clones (where integration took place at the right locus) is desirable. Abrogating NHEJ pathway (by knocking out *KU80* or other molecular players of NHEJ) would definitely minimize the number of those recombinant clones where random integration has taken place and thereby facilitate the screening of the desired clone. However, it would not increase the integration efficiency at the correct locus. Genome wide homology search by ssDNA bound Rad51 is the rate determining step in homology directed recombination. Thus, longer homologous stretches are likely to have great positive impact on such process. Our finding that increased flanking homology facilitate targeted gene knock out by TgRad51 is in corroboration with this hypothesis. Similarly, since ATP hydrolysis activity of Rad51 is required for such homology search a higher ATPase activity also results in more efficient gene targeting.

We investigated whether the absence of Ku80 protein has any positive effect on the efficiency of gene targeting. Since it is possible to obtain individual clonal transformants on solid medium using yeast system, it allowed us to determine the locus of integration (targeted versus random) in each of the individual colonies. As expected, there is no change in gene targeting efficiency in Ku80 null parasite lines (Fox BA et al., 2009). We found that the presence or absence of Ku80 protein does not have any significant effect on gene targeting efficiency. This is because, absence of Ku80 abrogates NHEJ pathway but does not increase the efficiency of Rad51 recombinase.

CHAPTER 6

DISCUSSION

This study revealed the underlying reasons for weak HR repair and inefficient gene targeting in *Toxoplasma gondii*. We cloned, expressed and purified a recombinant TgRad51 and performed biochemical studies to understand its functional role. Our investigation revealed that TgRad51 indeed has weak ATPase activity and thus it results in inefficient gene targeting. We showed that TgRad51 protein has ATP hydrolysis activity that is lowest among all the known eukaryotic Rad51 proteins studied so far. Three independent experiments- repair choice, targeted gene knock-in, and targeted gene knock-out were carried out in yeast to genetically characterize TgRad51. Our studies revealed that the Rad51 dependent gene conversion pathway is compromised in cells harboring TgRad51. Our results revealed that the weak ATPase activity of TgRad51 protein was responsible for poor homology searching and hence lowered gene conversion efficiency. Targeted gene knock-in and knock-out experiments revealed high frequency of random integration in presence of TgRad51. Due to its weak ATPase activity, TgRad51 showed poor homology searching and exhibited lower frequency of targeted gene integration. Hence, the homologous recombination (HR) machinery is weak in *T. gondii*.

Targeted gene disruption or tagging of endogenous genes are very less efficient in this parasite. Works from Bzik laboratory and Carruthers laboratory (Fox BA et al., 2009; Huynh MH et al., 2009) have demonstrated enhanced gene targeting in *ku80* null *T. gondii.* However, it does not increase the integration efficiency at the correct locus. Results from our gene targeting experiments suggest that Rad51 mediated homology search must be less inefficient in this parasite. As a result, TgRad51 preferred random integration. But, the increase in homology favored more of targeted gene integration. The absence of Ku80 in yeast assay system did not increase the efficiency of TgRad51 recombinase during gene targeting.

DNA double strand breaks are repaired by either homologous recombination (HR) and non-homologous end joining (NHEJ). Prokaryotes and eukaryotes show different choice of repair pathway. Both prokaryotes and lower eukaryotes prefer high fidelity HR as predominant pathway for DSB repair; whereas higher eukaryotes choose mutagenic NHEJ as the major repair pathway. An attractive hypothesis links the repair choice to gene density. Higher eukaryotes with lower gene density have greater probability of the breaks to occur in the intergenic region than the gene. Hence, they can afford to have mutagenic NHEJ as major pathway. For example, mouse or human genome has only 7 genes per megabase. On the other hand, prokaryotes and lower eukaryotes have high gene density and the probability of the random breaks to occur in the gene is higher than the intergenic region. As a result, they may not prefer mutagenic NHEJ as predominant pathway as it will result in the loss of gene function. Exceptionally, *T. gondii*, despite being lower eukaryote prefers mutagenic NHEJ as the major repair mechanism. Although, *T. gondii* has moderate gene density (100 genes per megabase) compared to its sister parasite *P. falciparum* (gene density is 231 per megabase), it is not unlikely to accumulate mutation within its genes. Whether, such genetic variations actually help the parasite to adapt to a variety of host cells, remains as an open question. Alternatively, the high frequency of NHEJ might just be a reflection of a weak HR mechanism of this parasite. Our finding that TgRad51 has weak ATPase activity and HR is weak in this parasite fits well with this notion. Our present study demonstrated that TgRad51 protein has comparatively a very low k_{cat} value of 0.034 min^{-1}. EcRecA, LmRad51 and ScRad51 with predominant HR repair have higher k_{cat} values of 18, 2.9 and 0.7 min^{-1} compared to HsRad51with major NHEJ repair (k_{cat} value of 0.16 min^{-1}). This supports the notion that the strong ATPase activity means a robust HR system in the organism.

The exact frequency of repair choice (HR versus NHEJ) in this parasite is not experimentally determined. The best way to determine such frequency is to create a DSB at a defined chromosomal locus and then follow the repair process on various counts: starting from the kinetics of the recruitments of different DNA repair proteins belonging to either HR or NHEJ pathway to the ultimate product after repair. Due to the technical limitations to perform such genetic manipulation in the parasite itself, we attempted to develop a surrogate yeast system to address such questions. To that end we have taken TgMre11, TgKu80 and TgRad51. *MRE11* plays a very essential role in DNA DSB repair. It is recruited before the choice is made between HR and NHEJ. Ku80 was reported to be an essential component of NHEJ in *Toxoplasma gondii* (Fox BA et al., 2009). Gene knockout at KU80 locus disrupted the functional NHEJ pathway in this parasite (Fox BA et al., 2009; Huynh MH et al., 2009). To understand NHEJ pathway in *T. gondii*, we tried to genetically characterize *TgKU80* gene. Unfortunately, neither TgMre11 nor TgKu80 could functionally complement the orthologous genes in yeast. Thus, a surrogate yeast model to study DSB repair choice of *T. gondii* could not be developed.

The putative orthologues of Rad50, Rad54, Dmc1 and Mre11 are present in the genome of *T. gondii*. However, the bioinformatics analysis of ToxoDB did not reveal any homologs of Rad52 and Rad55. Interestingly, the apparent lack of Rad52 orthologue in *Toxoplasma gondii*, *Plasmodium falciparum* and *Cryptosporidium parvum* is suggestive of a Rad52 independent recombination mechanism in these apicomplexan parasites.

Further, the *Toxoplasma* genome search revealed putative orthologues of *KU70* (50.m03211), DNA ligase IV (TGGT1_073840) and DNA protein kinase (57.m1765). Repeated gene knockout of DNA ligase IV and KU70 were unsuccessful in *T. gondii*

(Fox BA et al., 2009). Probably these two genes were essential and hence their knockout was not possible. This clearly strengthens the fact that NHEJ is a predominant repair pathway in this parasite. Gene knockout of *TgKU80* gene was possible but these *ku80* knockouts exhibited increased sensitivity to DSBs generated by γ-irradiation and phleomycin (Fox BA et al., 2009). From the *Toxoplasma* genome, the putative orthologue of XRCC3 is present but XRCC4 seem to be absent. The cell signaling orthologues Tel1 and Mec1 seem to be absent in *Toxoplasma*.

Findings from the present study suggest that replacement of TgRad51 by Rad51 of other organisms like *Leishmania major*, *Saccharomyces cerevisiae* might increase gene targeting efficiency in *Toxoplasma gondii.* LmRad51 and ScRad51 have 85 fold and 20 fold more ATPase activity than TgRad51. For improved targeted integration, the flanking ends of homology can be increased to a length of 1 Kb. The high gene targeting efficiency of this strain would easily facilitate genetic manipulation. Future studies might test this possibility either by expressing LmRad51/ScRad51 in *T. gondii* from an episomal plasmid or by delivering purified Rad51 proteins *via* nano particle mediated delivery systems. This transgenic strain can be used for the study of essential genes of related pathogens like *Plasmodium falciparum* and other parasites. This would greatly help in understanding the functions of unknown genes of these parasites and thus would aid in controlling the diseases caused by these deadly pathogens.

REFERENCES

1. Acharya S, Wilson T, Gradia S, Kane MF, Guerrette S. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc. Natl. Acad. Sci.* USA. 1996;93:13629-13631.
2. Agmon N, Pur S, Liefshitz B and Kupiec M. Analysis of repair mechanism choice during homologous recombination. *The Nucleic Acids Research*. 2009;37(15):5081-5092.
3. Aniukwu J, Glickman MS and Shuman S. The pathways and outcomes of mycobacterial NHEJ depend on the structure of the broken DNA ends. *Genes and Development*. 2008;22:512-527.
4. Audebert M, Salles B and Calsou P. Involvement of Poly(ADP-ribose) Polymerase-1 and XRCC1/DNA Ligase III in an alternative route for DNA double-strand breaks rejoining. *The Journal of Biological Chemistry*. 2004;279(53):55117-55126.
5. Audebert M, Salles B, Weinfeld M and Calsou P. Involvement of Polynucleotide Kinase in a Poly(ADP-ribose) polymerase-1-dependent DNA double-strand breaks rejoining pathway. *Journal of Molecular biology*. 2006;356:257-265.
6. Bakkenist CJ and Kastan MB. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature*. 2003;421(6922):499-506.
7. Bennardo N, Cheng A, Huang N and Stark JM. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genetics*. 2008;4(6):e1000110.

8. Benson FE, Stasiak A and West SC. Purification and characterization of the human Rad51 protein, an analogue of *E. coli* RecA. *The EMBO Journal*. 1994; 13(23):5764-5771.

9. Bethke L, Thomas S, Walker K, Lakhia R, Rangarajan R, Wirth D. The role of DNA mismatch repair in generating genetic diversity and drug resistance in malaria parasites. *Molecular and Biochemical Parasitology*. 2007;155:18-25.

10. Bhattacharyya MK, Kumar N. Identification and molecular characterization of DNA damaging agent induced expression of *Plasmodium falciparum* recombination protein PfRad51. *International Journal for Parasitology*. 2003;33:1385-1392.

11. Bhattacharyya MK, Mathews KM, Lustig AJ. Mre11 nuclease and C-terminal tail-mediated DDR functions are required for initiating yeast telomere healing. *Chromosoma*. 2008;117:357-366.

12. Bhattacharyya MK, Norris DE, and Kumar N. Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation. *Infection, Genetics and Evolution*. 2004;4:91-98.

13. Bhattacharyya MK, Sunanda D, Jayabalasingham B, Kumar N. Characterization of kinetics of DNA strand-exchange and ATP hydrolysis activities of recombinant PfRad51, a *Plasmodium falciparum* recombinase. *Molecular and Biochemical Parasitology*. 2005;139:33-39.

14. Buguliskis JD, Casta LJ, Butz CE, Matsumoto Y and Taraschi TF. Expression and biochemical characterization of *Plasmodium falciparum* DNA ligase I. *Mol. Biochem. Parasitol*. 2007;155(2):128-137.

15. Burton P, McBride DJ, Wilkes JM, Barry JD, and McCulloch R. Ku heterodimer-independent end joining in *Trypanosoma brucei* cell extracts relies upon sequence micro homology. *Eukaryot Cell*. 2007;61773-61781.

16. Casta LJ, Buguliskis JS, Matsumoto Y, Taraschi TF. Expression and biochemical characterization of the Plasmodium falciparum DNA repair enzyme flap endonuclease-1 (PfFen-1). *Mol. Biochem. Parasitol.* 2007;157:1-12.

17. Castellini MA, Buguliskis JS, Casta LJ, Butz CE, Clark AB, Kunkel TA, Taraschi TF. Malaria drug resistance is associated with defective DNA mismatch repair. *Molecular and Biochemical Parasitology*. 2011;177:143-147.

18. Chayot R, Montagne B, Mazel D, and Ricchetti M. An end-joining repair mechanism in *Escherichia coli*. *PNAS*. 2010;107(5):2141-2146.

19. Cheng Q, Barboule N, Frit P, Gomez D, Bombarde O, Couderc B, Ren G, Salles B, Calsou P. Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks. *Nucleic Acids Research*. 2011:1-15.

20. Chung WH, Zhu Z, Papusha A, Malkova A and Ira G. Defective resection at DNA double-strand breaks leads to *De Novo* telomere formation and enhances gene targeting. *PLoS Genetics*. 2010;6(5)e1000948.

21. Clever B, Interthal H, Schmuckli-maurer J, King J, Sigrist M, Heyer WD. Recombinational repair in yeast: functional interactions between Rad51 and Rad54 proteins. *The EMBO Journal*. 1997;16(9):2535-2544.

22. Couëdel C, Mills KD, Barchi M, Shen L, Olshen A, Johnson RD, Nussenzweig A, Essers J, Kanaar R, Li GC, Alt FW, Jarin M. Collaboration of

homologous recombination and nonhomologous end-joining factors for the survival and integrity of mice and cells. Genes and *Development*. 2004;18:1293-1304.

23. Cruz A., and S. M. Beverley. Gene replacement in parasitic protozoa. *Nature* 1990;348:171–173.
24. de Jager M, van Noort J, van Gent DC, Dekker C, Kanaar R, Wyman C. Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Molecular Cell*. 2001;8(5):1129-1135.
25. Deng Y, Guo X, Ferguson DO and Chang S. Multiple roles for Mre11 at uncapped telomeres. *Nature*. 2009;460(7257):914-918.
26. Diede SJ, Gottschling DE. Telomerase-mediated telomere addition in vivo requires DNA primase and DNA polymerases alpha and delta. *Cell*. 1999;99:723-733.
27. Dominguez-Bendala J, Masutani M and McWhir J. Down-regulation of PARP-1, but not of Ku80 or DNA-PKcs results in higher gene targeting efficiency. *Cell Biol. Int*. 2006;30:389-393.
28. Donald RG, Roos DS. Gene knock-outs and allelic replacements in *Toxoplasma gondii*: HXGPRT as a selectable marker for hit-and-run mutagenesis. *Mol Biochem Parasitol*. 1998;91(2):295-305.
29. Donald RGK, Carter D, Ullman B, Roos DS. Insertional tagging, cloning and expression of the *Toxoplasma gondii* hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation. *Journal of biological chemistry*. 1996;271:14010-14019.

30. Donald RGK, Roos DS. Insertional mutagenesis and marker rescue in a protozoan parasite, Cloning of the uracil phosphoribosyltransferase locus from *Toxoplasma gondii*. *PNAS* USA. 1995;92:5749-5753.

31. Eid J and B. Sollner-Webb. Stable integrative transformation of *Trypanosoma brucei* that occurs exclusively by homologous recombination. *Proc. Natl. Acad. Sci*. USA. 1991;88:2118–2121.

32. Fattah F, Lee E H, Weisensel N, Wang Y, Lichter N, Hendrickson E A. Ku regulates the non-homologous end joining pathway choice of double strand break repair in human somatic cells. *PLoS Genetics*. 2010;6(2):e1000855.

33. Firmenich AA, Elias-Arnanz M, Berg P. A novel allele of *Saccharomyces cerevisiae* RFA1 that is deficient in recombination and repair and suppressible by RAD52. Mol. Cell. Biol. 1995;15(3):1620-1631.

34. Fox BA, Ristuccia JG, Gigley JP and Bzik DJ. Efficient gene replacements in *Toxoplasma gondii* strains deficient for non-homologous end joining. *Eukaryotic cell*. 2009;8(4):520-529.

35. Frouin I, Maga G, Denegri M, Riva F, Savio M, Spadari S, Prosperi E and Scovassi I. Human proliferating cell nuclear antigen, poly(ADP-ribose) polymerase-1, and p21waf1/cip1. *JBC*. 2003;278:39265-39268.

36. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B.

Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*. 2002;419(6906):498-511.

37. Ghosal G and Muniyappa K. The characterization of *Saccharomyces cerevisiae* Mre11/Rad50/Xrs2 complex reveals that Rad50 negatively regulates Mre11 endonucleolytic but not the exonucleolytic activity. *Journal of Molecular Biology*. 2007;372:864-882.
38. Glover L, McCulloch R, and Horn D. Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Res*. 2008;362608-2618.
39. Gu L, Hong Y, McCulloch S, Watanabe H, Li GM. ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. *Nucleic Acids Res*. 1998;26:1173-1178.
40. Gupta RC, Bazemore RL, Golub EI and Radding CM. Activities of human recombination protein Rad51. *Proc Natl Acad Sci* USA. 1997;94:463-468.
41. Guzder SN, Sung P, Prakash L, Prakash S. Yeast DNA-repair gene RAD14 encodes a zinc metalloprotein with affinity for ultraviolet-damaged DNA. *Proc. Natl. Acad Sci* USA. 1993;90(12):5433-5437.
42. Haber JE. DNA recombination: The replication connection. *Trends Biochem. Sci*. 1999;24:271-275.
43. Haber JE. Partners and pathways repairing a double strand break. *Trends in Genetics*. 2000;16(6):259-264.
44. Haber JE. The many interfaces of Mre11. *Cell*. 1995;95:583-586.
45. Habraken Y, Sung P, Prakash L, Prakash S. Binding of insertion/deletion DNA mismatches by the heterodimer of yeast mismatch repair proteins MSH2 and MSH3. *Curr. Biol*. 1996;6:1185-1187.

46. Habraken Y, Sung P, Prakash L, Prakash S. Yeast excision repair gene RAD2 encodes a single-stranded DNA endonuclease. *Nature*. 1993;366(6453):365-368.

47. Haltiwanger BM, Karpinich NO and Taraschi TF. Characterization of class II apurinic/apyrimidinic endonuclease activities in the human malaria parasite, *Plasmodium falciparum. Biochem. J.* 2000;345:85-89.

48. Haltiwanger BM, Matsumoto Y, Nicolas E, Dianov GL, Bohr VA and Taraschi TF. DNA Base Excision repair in Human malaria parasites is predominantly by a long-patch pathway. *Biochemistry*. 2000;39:763-772.

49. Hays SL, Firmenich AA and Berg P. Complex formation in yeast double strand break repair: participation of Rad51, Rad52, Rad55 and Rad57 proteins. *Proc National Sci Academy*. 1995;92(15):6925-6929.

50. Hays SL, Firmenich AA, Massey P, Banerjee R, Berg P. Studies of the interaction between Rad52 protein and the yeast single-stranded DNA binding protein RPA. *Mol. Cell. Biol*. 1998;18(7):4400-4406.

51. Hicks WM, Yamaguchi M and Haber JE. Real-time analysis of double-strand DNA break repair by homologous recombination. *PNAS*. 2011;108(8):3108-3115.

52. Holmes JJ, Clark S, Modrich P. Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines. *Proc. Natl. Acad. Sci.* USA. 1990;87:5837-5841.

53. Hopfner KP, Craig L, Moncalian G, Zinkel RA, Usui T, Owen BA. The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature*. 2002;418(6897):562-566.

54. Huynh MH and Carruthers VB. Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. *Eukaryotic cell*. 2009;8(4):530-539.

55. Iiizumi S, Kurosawa A, So S, Ishii Y, Chikaraishi Y, Ishii A, Koyama H and Adachi N. Impact of non-homologous end-joining deficiency on random and targeted DNA integration: implications for gene targeting. *Nucleic Acids research*. 2008;36(19):6333-6342.

56. Iliakis G. Backup pathways of NHEJ in cells of higher eukaryotes: cell cycle dependence. *Radiotherapy and Oncology*. 2009;92(3):310-315.

57. Jain S, Sugawara N, Lydeard J, Vaze M, Tanguy Le Gac N, Haber JE. A recombination execution checkpoint regulates the choice of homologous recombination pathway during DNA double-strand break repair. Genes Dev. 2009;23(3):291-303.

58. Janzen CJ, Lander F, Dreesen O, and Cross GA. Telomere length regulation and transcriptional silencing in KU80-deficient *Trypanosoma brucei*, *Nucleic Acids Res*. 2004;32:6575-6584.

59. Jazayeri A, Balestrini A, Garner E, Haber JE and Costanzo V. Mre11-Rad50-Nbs1-dependent processing of DNA breaks generates oligonucleotides that stimulate ATM activity. *EMBO Journal*. 2008;27,1953-1962.

60. Jiang H, Xie Y, Houston P, Stemke-Hale K, Mortensen UH, Rothstein R and Kodadek T. Direct association between the yeast Rad51 and Rad54 recombination proteins. *J Biol Sci.* 1996;271:33181-33186.

61. Jiricny J. Replication errors: cha(lle)nging the genome. *EMBO J*. 1998;17:6427-6436.

62. Johnson RD, Symington LS. Functional differences and interactions among the putative RecA homologs Rad51, Rad55, and Rad57. *Mol. Cell. Biol.* 1995;15(9):4843-4850.

63. Kapler GM, Coburn CM, Beverley SM. Stable transfection of the human parasite *Leishmania major* delineates a 30-kilobase region sufficient for extrachromosomal replication and expression. *Mol. Cell Biol*. 1990;10:1084-1094.

64. Kelley MR, Kow YW and Wilson III DM. Disparity between DNA base excision repair in yeast and mammals: translational implications. *Cancer Research*. 2003;63:549-554.

65. Kim K, Weiss LM. *Toxoplasma gondii*: the model apicomplexan. *Int. J. Parasitol*. 2004; 34(3):423-432.

66. Kim K, Weiss LM. *Toxoplasma*: the next 100 years. *Microbes and Infection*. 2008; 10:978-984.

67. Kissinger JC, Gajria B, Li L, Paulsen IT and Roos DS. ToxoDB: accessing the Toxoplasma gondii genome. *Nucleic Acids Research*. 2003;31:234-236.

68. Kolodner RD, Marsischky GT. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev*. 1999;9:89-96.

69. Kowalczykowski S C, Krupp R A. DNA strand exchange promoted by RecA protein in the absence of ATP: Implications for the mechanism of energy transduction in protein-promoted nucleic acid transactions. *PNAS*. 1995;92:3478-3482.

70. Laskar S, Bhattacharyya MK, Shankar R and Bhattacharyya S. Hsp90 controls SIR2 mediated gene silencing. *PloS One*. 2011;6(8):e23406.

71. Lavine MD and Arrizabalaga G. The antibiotic monensin causes cell cycle disruption of *Toxoplasma gondii* mediated through the DNA repair enzyme TgMSH-1. Antimicrob Agents Chemother. 2011; 55(2):745-755.

72. Lee JH and Paull TT. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science*. 2005;308(5721):551-554.

73. Lee JH and Paull TT. Direct activation of the ATM protein kinase by the Mre11/Rad50/Nbs1 complex. *Science*. 2004;304(5667):93-96.

74. Lee MG, Van der Ploeg LH. Homologous recombination and stable transfection in the parasitic protozoan *Trypanosoma brucei*. *Science*. 1990;250:1583-1587.

75. Liang F, Han M, Romanienko PJ, and Jasin M. Homology directed repair is a major double-strand break repair pathway in mammalian cells. *Proc. Natl. Acad. Sci* USA. 1998;95:5172-5177.

76. Liang F, Romaneiko PJ, Weaver DT, Jeggo PA and Jasin M. Chromosomal double-strand break repair in Ku80-deficient cells. *Proc. Natl Acad. Sci* USA. 1996;93:8929-8933.

77. Liang L, Deng L, Nguyen SC, Zhao X, Maulion CD, Shao C and Tischfield JA. Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Res*. 2008;36(10):3297-3310.

78. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review Biochemistry*. 2010;79:181-211.

79. Liu N, Lamerdin JE, Tebbs RS, Schild D, Tucker JD, Shen MR, Brookman KW, Siciliano MJ, Walter CA, Fan W. XRCC2 and XRCC3, new human

Rad51- family members, promote chromosome stability and protect against DNA cross-links and other changes. *Mol Cell*. 1998;1:783-793.

80. Longtine MS, McKenzie A 3rd, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*. 1998;14(10):953-961.
81. Luft BJ and Remington JS. Toxoplasmic encephalitis in AIDS patients. *Clin. Infect. Dis*. 1992;15:211-222.
82. Lundblad V and Blackburn EH. An alternative pathway for yeast telomere maintenance rescues est1-senescence. *Cell*. 1993;73:347-360.
83. Luo G, Yao MS, Bender CF, Mills M, Bladl AR. Disruption of mRAD50 causes embryonic stem cell lethality, abnormal embryonic development and sensitivity to ionizing radiation. *Proc Natl Acad Sci* USA. 1999;96:7376-7381.
84. Malkova A, Naylor ML, Yamaguchi M, Ira G and Haber JE. RAD51-dependent break-induced replication differs in kinetics and checkpoint responses from RAD51-mediated gene conversion. *Mol. Cell. Biol*. 2005;25: 933-944.
85. McKean PG, Keen JK, Smith DF, Benson FE. Identification and characterisation of a RAD51 gene from *Leishmania major*. *Molecular and Biochemical Parasitology*. 2001;115:209-216.
86. Meek K, Dang V, Lees-Miller SP (2008) DNA-PK: the means to justify the ends? Adv Immunol. 2008;99:33-58.
87. Miller III CA, Martinat MA and Hyman LE. Assessment of aryl hydrocarbon receptor complex interactions using pBEVY plasmids: expression vectors with

bi-directional promoters for use in *Saccharomyces cerevisiae*. *The Nucleic acid research journal*. 1998;26(15):3577-3583.

88. Mimitou EP and Symington LS. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature*. 2008;455(7214):770-774.
89. Mimitou EP and Symington LS. DNA end resection: many nucleases make light work. *DNA Repair*. 2009;8(9):983-995.
90. Mladenov E and Iliakis G. (2011) Induction and repair of DNA double strand breaks: The increasing spectrum of non-homologous end joining pathways. *Mutation research.* 2011;711(1-2):61-72.
91. Morrison C, Shinohara A, Sonoda E, Yamaguchi-Iwai Y, Takata M, Weichselbaum RR, Takeda S. The essential functions of human Rad51 are independent of ATP hydrolysis. *Mol Cell Biol*. 1999;19(10):6891-6897.
92. Ninomiya Y, Suzuki K, Ishii C, Inoue H. Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. *PNAS*. 2004;101(33):12248-12253.
93. Nunthawarasilp P, Petmitr S and Petmitr PC. Partial purification and characterization of DNA polymerase β-like enzyme from *Plasmodium falciparum*. *Molecular and biochemical parasitology.* 2007;154:141-147.
94. Ogawa T, Yu X, Shinohara A, Egelman EH. Similarity of the yeast RAD51 filament to the bacterial RecA. *Science*. 1993;259(5103):1896-1899.
95. Okano S, Lan L, Caldecott K.W, Mori T and Yasui A. Spatial and temporal cellular responses to single-strand breaks in human cells. *Mol. Cell. Biol*. 2003;23:3974-3981.
96. Onyango DO, Naguleswaran A, Delaplane S, Reed A, Kelley MR, Georgiadis MM, Sullivan WJ Jr. *DNA repair*. 2011;10(5):466-475.

97. Paques F, Haber JE. Multiple pathways of recombination induced by double strand break in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev*. 1999;63:349-404.

98. Paques F, Haber JE. Two pathways for removal of non-homologous DNA ends during double strand break repair in *Saccharomyces cerevisiae. Mol Cell Biol.* 1997;17:6765-6771.

99. Pascucci B, Russo MT, Crescenzi M, Bignami M and Dogliotti E. The accumulation of MMS-induced single strand breaks in G1 phase is recombinogenic in DNA polymerase beta defective mammalian cells. *Nucleic Acids Res*. 2005;33:280–288.

100. Paull TT and Gellert M. The 3' to 5' exonuclease activity of Mre11 facilitates repair of DNA double-strand breaks. *Molecular Cell*. 1998;1(7):969-979.

101. Pierce AJ, Johnson RD, Thompson LH, Jasin M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev*. 1999;13:2633-2638.

102. Pierce AJ, Hu P, Han M, Ellis N and Jasin M. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes and Development*. 2001;15:3237-3242.

103. Remington JS and Desmonts G. Toxoplasmosis. In Remington JS and Klein JO (eds). Infectious diseases of the fetus and newborn infant. W. B. Saunders, Philadelphia PA. 1989;89-195.

104. Riha K, Heacock ML, Shippen DE (2006) The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology. *Annu Rev Genet*. 2006;40:237-277.

105. Roth DB, Porter TM and Wilson JH. Mechanisms of nonhomologous recombination in mammalian cells. *Molecular Cell biology*. 1985;5:2599-2607.

106. Rupnik A, Lowndes N F, Grenon M. MRN and the race to the break. *Chromosoma.* 2010;119:115-135.

107. Shinohara A, Ogawa H, Ogawa T. Rad51 protein involved in repair and recombination in Saccharomyces cerevisiae is a RecA-like protein. *Cell*. 1992;69(3):457-470.

108. Sipley JD, Menninger JC, Hartley KO, Ward DC, Jackson SP, and Anderson CW. Gene for the catalytic subunit of the human DNA-activated protein kinase maps to the site of the XRCC7 gene on chromosome. *Proc. Natl. Acad. Sci* USA. 1995;92(16):7515-7519.

109. Stark JM and Jasin M. Extensive loss of heterozygosity is suppressed during homologous repair of chromosomal breaks. *Molecular and Cellular biology*. 2003;23(2):733-743.

110. Stark JM, Hu P, Pierce AJ, Moynahan ME, Ellis N and Jasin M. ATP hydrolysis by mammalian Rad51 has a key role during homology-directed DNA repair. *The journal of Biological Chemistry*. 2002;277(23):20185-20194.

111. Stark JM, Pierce AJ, Oh J, Pastink A and Jasin M. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Molecular and Cellular Biology*. 2004;24(21):9305-9316.

112. Sugawara N, Paques F, Colaiacovo M, Haber JE. Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break induced recombination. *Proc Natl Acad Sci* USA. 1997;94:9214-9219.

113. Sugawara N, Wang X and Haber JE. *In vivo* roles of Rad52, Rad54 and Rad55 proteins in Rad51-mediated recombination. *Mol. Cell*. 2003;12(1):209-219.

114. Sung P, Reynolds P, Prakash L, Prakash S. Purification and characterization of the *Saccharomyces cerevisiae* RAD1/RAD10 endonuclease. *J. Biol. Chem*. 1993;268(35):26391-26399.

115. Sung P. Catalysis of ATP dependent homologous DNA pairing and strand exchange by yeast Rad51 protein. *Science*. 1994;265:1241-1243.

116. Symington LS. Role of Rad52 epistatis group genes in homologous recombination and double strand break repair. *Microbiol Mol Biol Rev.* 2002;66:630-670.

117. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double strand break repair model for recombination. *Cell*. 1983;33(1):25-35.

118. Tarique M, Satsangi AT, Ahmad M, Singh S, Tuteja R. *Plasmodium falciparum* MLH is schizont stage specific endonuclease. *Molecular and Biochemical Parasitology*. 2012;181:153-157.

119. Timson DJ, Singleton MR, and Wigley DB. DNA ligases in the repair and replication of DNA. *Mutation Research.* 2000;460:301-318.

120. Tittel-Elmer M, Alabert C, Pasero P and Cobb JA. The MRX complex stabilizes the replisome independently of the S phase checkpoint during replication stress. *The EMBO Journal*. 2009;28:1142-1156.

121. Tomkinson AE, Bardwell AJ, Bardwell L and Tappe NJ, Friedberg EC. Yeast DNA repair and recombination proteins Rad1 and Rad10 constitute a single-stranded-DNA endonuclease. *Nature*. 1993;362(6423):860-862.

122. Tsuzuki T, Fujii Y, Sakumi K, Tominaga Y, Nakao K, Seiguchi M, Matsushiro A, Yoshimura Y, Morita T. Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. *PNAS*. 1996; 93:6236-6240.

123. Umar A, Buermeyer AB, Simon JA, Thomas DC, Clark AB. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell*. 1996;87:65-73.

124. Usui T, Ohta T, Oshiumi H, Tomizawa J, Ogawa H. Complex formation and functional versatility of Mre11 of budding yeast in recombination. *Cell*. 1998;95(5):705-716.

125. Walker JR, Corpina RA and Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*. 2001;412(6847):607-614.

126. Wang H, Perrault AR, Takeda Y, Qin W, Wang H and Iliakis G. Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic acids research*. 2003;31(18):5377-5388.

127. Wang H, Rosidi B, Perrault R, Wang M, Zhang L, Windhofer F and Iliakis G. DNA ligase III as a candidate component of Backup pathways of Nonhomologous end joining. *Cancer Research*. 2005;65(10):4020-4030.

128. Wang M, Wu W, Wu W, Rosidi B, Zhang L, Wang H and Iliakis G. PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Research*. 2006;34(21):6170-6182.

129. Wang X and Haber JE. Role of *Saccharomyces* single-stranded DNA-binding protein RPA in the strand invasion step of double-strand break repair. *PLoS Biol*. 2004;2(1):E21.

130. Williams B, Bhattacharyya MK, Lustig AJ. Mre11p nuclease activity is dispensable for telomeric rapid deletion. *DNA repair*. 2005;4:994-1005.

131. Williams RS, Moncalian G, Williams JS, Yamada Y, Limbo O, Shin DS, Groocock LM, Cahill D, Hitomi C, Guenther G, Moiani D, Carney JP, Russell P, Tainer JA. Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell*. 2008;135(1):97-109.

132. Wu Y, Kirkman LA, Wellems TE. Transformation of *Plasmodium falciparum* malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. *Proc Natl Acad Sci USA* 1996;93:1130-1134.

133. Wyatt MD and Pittman DL. Methylating agents and DNA repair responses: methylated bases and sources of strand breaks. *Chem. Res. Toxicol*. 2006;19:1580–1594.

134. Yamaguchi-Iwai Y, Sonoda E, Sasaki MS, Morrison C, Haraguchi T et al. Mre11 is essential for the maintenance of chromosomal DNA in vertebrate cells. *EMBO Journal*. 1999;18:6619-6629.

135. Yano K, Morotomi-Yano K, Lee K, Chen D J. Functional significance of the interaction with Ku in DNA double-strand break recognition of XLF. *FEBS LETTERS* 2011;585:841-846.

136. Zaitseva E M, Zaitsev E N, Kowalczykowski. The DNA binding properties of *Saccharomyces cerevisiae* Rad51 protein. *The Journal of Biological Chemistry* 1999;274(5):2907-2915.

137. Zhou C, Li Z, Diao H, Yu Y, Zhu W, Dai Y, Chen FF and Yang J. DNA damage evaluated by gammaH2AX foci formation by a selective group of chemical/physical stressors. *Mutat. Res*. 2006;604:8–18.

138. Zhu J, Petersen S, Tessarollo L, Nussenzweig A. Targeted disruption of the Nijmegen breakage syndrome gene NBS1 leads to early embryonic lethality in mice. *Curr Biol*. 2001;11:105-109.

APPENDIX

APPENDIX- 1

A.1.1 Construction of *TgMRE11* gene in pTA vector (pTA:*TgMRE11*)

TgMRE11 gene was cloned into pTA vector containing TRP and ampicillin resistance markers. From the cDNA library of *T. gondii*, a 3.4 Kb *TgMRE11* ORF (Figure A1.B) was amplified by nested PCR with internal primers OMKB 72 and OMKB 73. *BamHI* site was incorporated in both the primers. To express TgMre11 in yeast model system, the full length *TgMRE11* gene was cloned into the pCR 2.1 TOPO vector and the insert was released by *BamHI* enzyme (Figure A1.C) and finally sub cloned into pTA plasmid at the same site under the GPD promoter (Figure A1.A). Both pTA vector backbone and the *TgMRE11* gene contained *EcoRI* site. The directionality of *TgMRE11* gene was confirmed by *EcoRI* digestion, which gave two bands of sizes 1.5 Kb and 8.1 Kb (Figure A1.D).

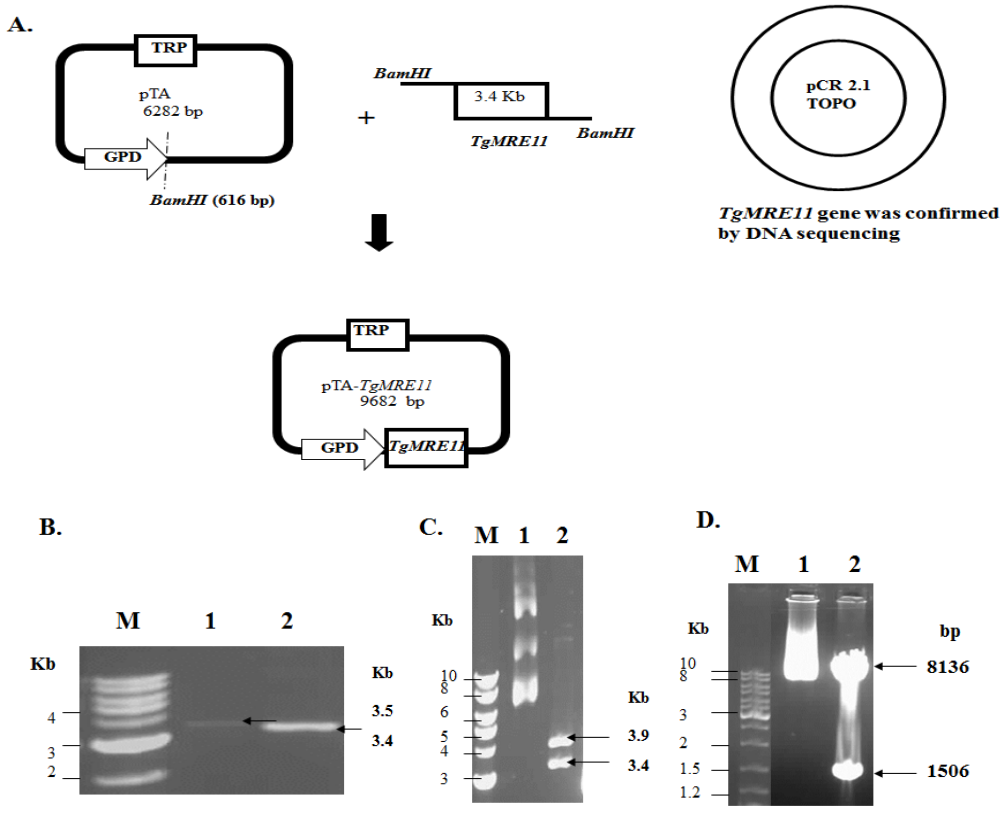

Figure A1: Construction of *TgMRE11* clone in pTA vector. A). Illustration depicted the cloning of *TgMRE11* gene into the intermediate pCR 2.1 TOPO vector and the sub cloning into pTA plasmid at *BamHI* site. B). The PCR result of *TgMRE11* from cDNA library with outer and inner set of primers shown in lanes 1 and 2 respectively. The final PCR product of *TgMRE11* gene was 3.4 Kb in size. C). Uncut plasmid was loaded in the lane 1. *TgMRE11* insert was released by *BamHI* from intermediate TOPO vector and gave two bands, 3.4 and 3.9 Kb as observed in lane 2. D). The right orientation of *TgMRE11* insert in the pTA plasmid confirmed by *EcoRI* digestion. Uncut plasmid pTA-*TgMRE11* was loaded in lane 1. Upon *EcoRI* digestion, two expected bands of sizes 8.1 and 1.5 Kb were observed in lane 2. Standard DNA marker was represented as M in Figure B to D and the bands are indicated by arrowhead in left panel.

A.1.2 Construction of chimeric *MRE11* gene in pTA vector (pTA:chimera *MRE11*)

The N-terminal *TgMRE11* of size 2.8 Kb was PCR amplified from the full length *TgMRE11* gene. The primers OMKB 72 and 108 harboured *BamHI* and *SacI* sites respectively. The C-terminal *ScMRE11* gene of size 1128 bp was PCR amplified from the full length *ScMRE11* gene. The oligos OMKB 109 and OMKB 85 contained *SacI* and *PstI* sites respectively. We cloned N-terminal *TgMRE11* into pCR 2.1 TOPO vector and the insert was released after *BamHI* and *SacI* digestion. Similarly, the C-terminal *ScMRE11* gene was cloned into pCR 4 TOPO vector and released by *SacI* and *PstI* enzymes. Both the genes were ligated with pTA plasmid at *BamHI* and *PstI* sites (Figure A2.A). The confirmation of N-terminal *TgMRE11* gene in the intermediate vector was done by *BamHI* digestion, which released the insert of size 2.8 Kb (Figure A2.B). Similarly, the cloning of the C-terminal *ScMRE11* gene in the intermediate vector was confirmed by *PstI* digestion which gave two expected bands of sizes 1149 bp and 3935 bp (Figure A2.C). The confirmation of the chimeric *MRE11* in pTA plasmid was done by *BamHI* and *PstI* digestions where in the chimeric *MRE11* insert of size 3.9 Kb was released (Figure A2.D).

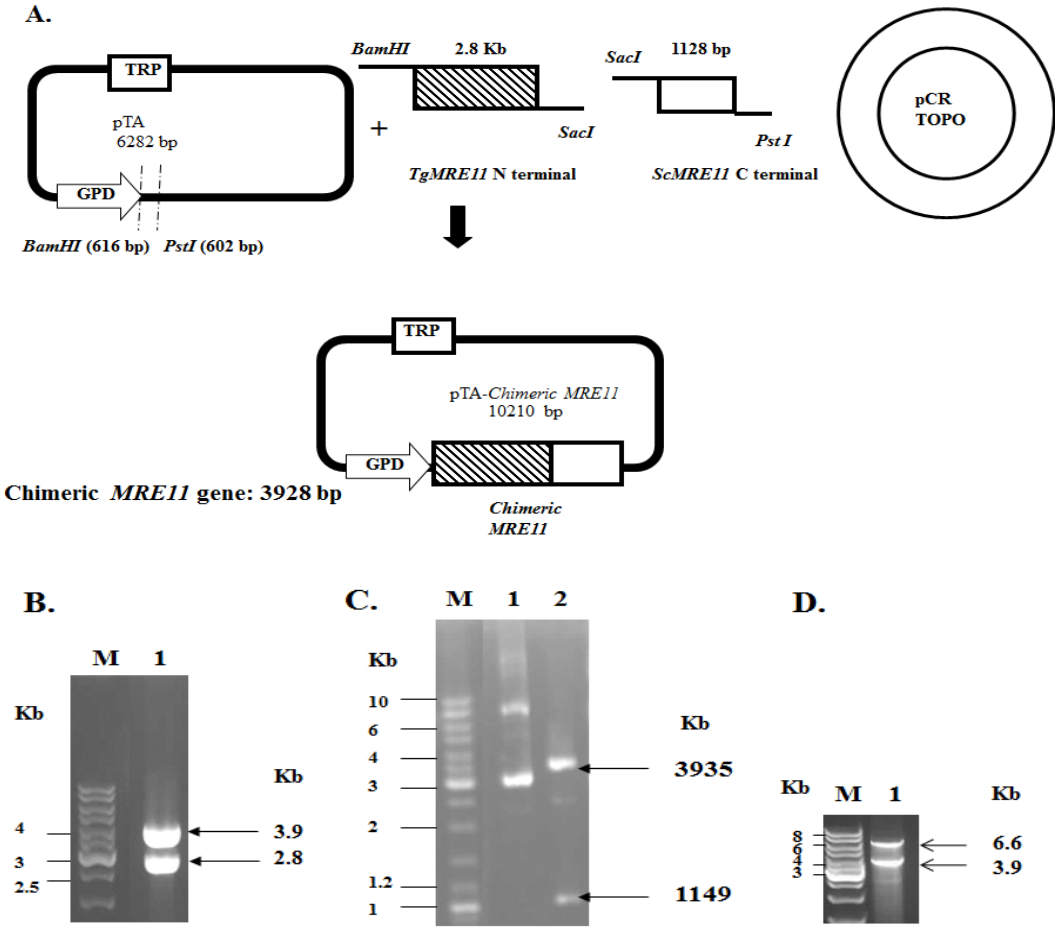

Figure A2: Construction of Chimeric *MRE11* in pTA plasmid. A). The chimeric *MRE11* was cloned into pTA plasmid at *BamHI* and *PstI* sites, under GPD promoter. B). The N-terminal *TgMRE11* of size 2.8 Kb was released from TOPO vector by *BamHI* digestion as shown in lane 1. C). Uncut TOPO vector containing C-terminal *ScMRE11* was loaded in lane 1. The insert of size 1149 bp was released from TOPO vector by *PstI* digestion as observed in lane 2. D). The chimeric *MRE11* of 3.9 Kb as shown in lane 1 was confirmed in pTA plasmid by *BamHI* and *PstI* digestion. Standard DNA marker was represented as M in Figure B to D and the bands are indicated by arrowhead in left panel.

A.1.3 Construction of *TgKU80* gene in pTA vector (pTA:*TgKU80*)

TgKU80 gene was PCR amplified from the genomic DNA of *Toxoplasma gondii* RH strain. The oligos OMKB 136 and OMKB 138 contained *BglII* and *XbaI* sites respectively. PCR yielded a single band of *TgKU80* corresponding to size 5183 bp (Figure A3.B). This gene was cloned into pTZ57R/T plasmid. The directionality of the insert in this plasmid was checked by *EcoRI* enzyme. Internally a single *EcoRI* site was present in both vector and *TgKU80* insert. As a result, *EcoRI* digestion yielded two expected bands of sizes 7.5 Kb and 539 bp (Figure A3.C). The *TgKU80* gene was confirmed by DNA sequencing. This insert was released by *BglII* and *XbaI* digestion and sub cloned into the yeast expression plasmid pTA at *BamHI* (616 bp) and *XbaI* (609 bp) sites, under GPD promoter (Figure A3.A). The confirmation of the *TgKU80* insert in pTA plasmid was done by *HindIII* enzyme. pTA vector backbone and the *TgKU80* insert each contains two *HindIII* sites. Therefore, upon digestion with *HindIII*, the clone gave four expected DNA fragments of sizes 353, 1747, 3441 and 5541 bp (Figure A3.D).

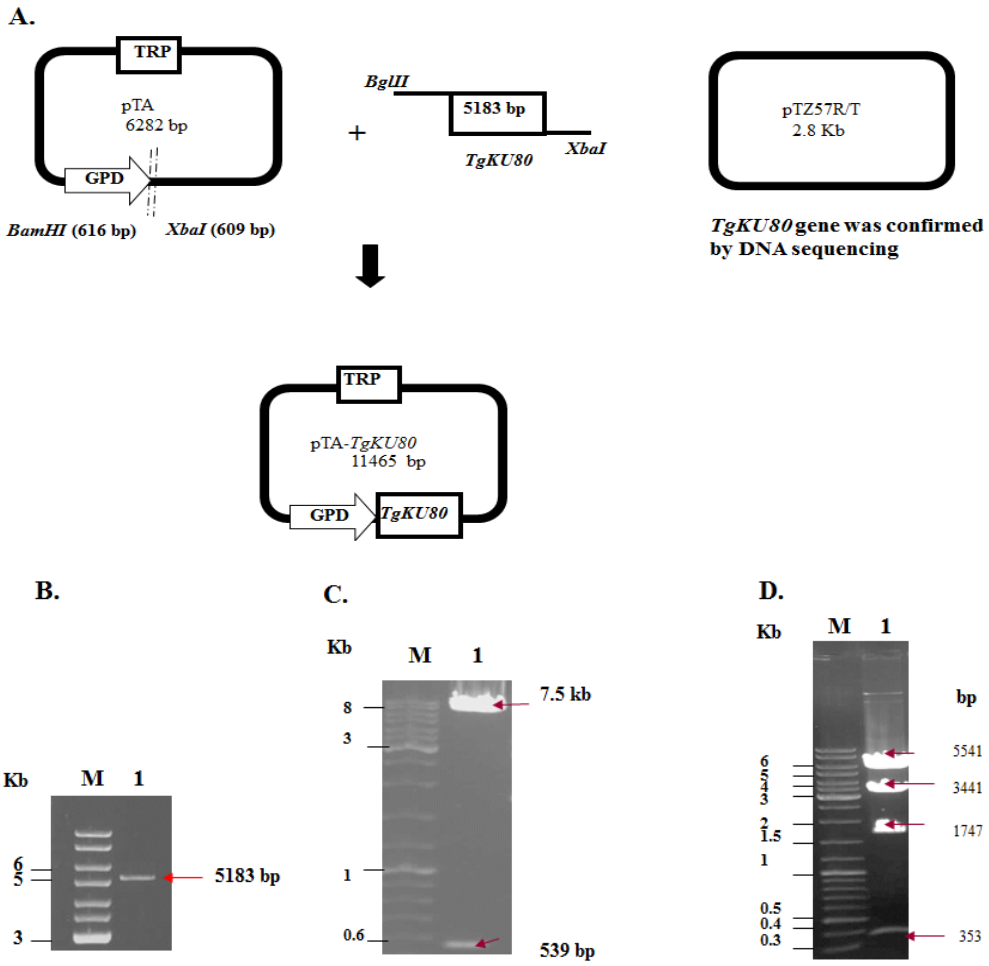

Figure A3: Construction of *TgKU80* clone in pTA plasmid. A). The strategy to clone *TgKU80* gene into pTA plasmid. B). *TgKU80* gene of size 5183 bp was PCR amplified as shown in lane 1. C). pTZ57R/T-*TgKU80* plasmid gave two expected bands of sizes 539 bp and 7.5 Kb upon *EcoRI* digestion as shown in lane 1. D). The confirmation of *TgKU80* in the yeast expression plasmid pTA produced four expected bands 353, 1747, 3441 and 5541 bp with *HindIII* as shown in lane 1. Standard DNA marker was represented as M in Figure B to D and the bands are indicated by arrowhead in left panel.

A.1.4 Construction of *ScKU80* gene in pTA vector (pTA:*ScKU80*)

ScKU80 gene was PCR amplified from the genomic DNA of wild type yeast strain. The oligos OMKB 155 and OMKB 156 contained *BglII* and *SalI* restriction sites respectively. The PCR product obtained was of size 1890 bp (Figure A4.B). The *ScKU80* gene was cloned into pTZ57R/T plasmid. This clone was confirmed by *BglII* digestion which produced the linearized plasmid at 4.7 Kb (Figure A4.C). Second digestion with *SalI* enzyme released the *ScKU80* insert and the expected bands of sizes 1.9 Kb and 2.8 Kb were observed (Figure A4.C). The insert was released after *BglII* and *SalI* digestion and finally sub cloned into pTA plasmid at *BamHI* (616 bp) and *SalI* (604 bp) sites (Figure A4.A). This insert was confirmed in pTA plasmid by *EcoRI* digestion. pTA vector contained a single *EcoRI* site and *ScKU80* gene had two *EcoRI* sites. Hence, upon digestion, three DNA fragments of sizes 6.5, 1.2 and 0.4 Kb were observed (Figure A4.D).

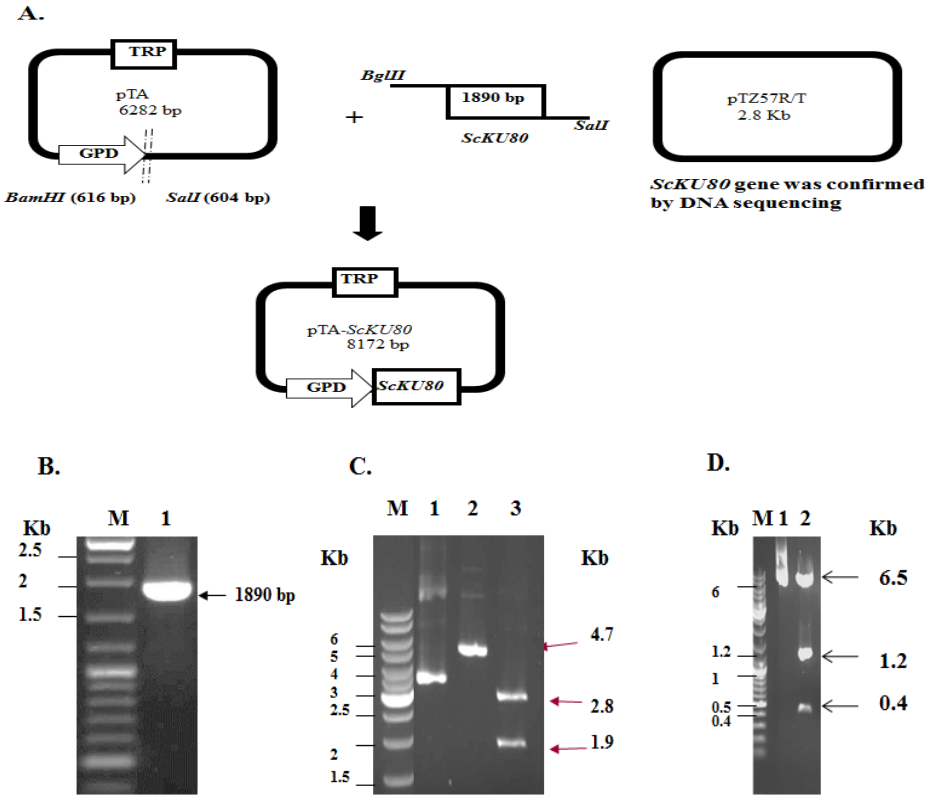

Figure A4: Construction of *ScKU80* clone in pTA plasmid. A). The strategy to clone *ScKU80* gene into pTA plasmid at *BamHI* and *SalI* sites. B). Lane 1 showed the PCR product of *ScKU80* gene (1890 bp) amplified from the genomic DNA of *Saccharomyces cerevisiae*. C). *ScKU80* gene was first cloned into pTZ57R/T. The uncut DNA was loaded in lane 1. Clone was confirmed by *BglII* digestion which produced 4.7 Kb band as shown in lane 2. Further with *SalI* digestion, the insert of 1890 bp was released and the empty vector backbone of 2.8 Kb was observed in lane 3. D). The pTA plasmid containing *ScKU80* gene was further confirmed for right orientation with *EcoRI* digestion. Uncut plasmid was loaded in lane 1. With *EcoRI* digestion, three expected bands of sizes 6.5, 1.2 and 0.4 Kb were observed in lane 2. Standard DNA marker was represented as M in Figure B to D and the bands are indicated by arrowhead in left panel.

A.1.5 Construction of *TgRAD51* gene in pTA vector (pTA:*TgRAD51*)

The open reading frame encoding *Toxoplasma gondii* Rad51 was amplified from the cDNA library (provided by Prof. Vern Carruther) of *T. gondii* RH strain using PCR extender enzyme as described by the manufacturer. The forward primer OMKB 65 and reverse primer OMKB 82 are complementary to the 5′ and 3′ ends of the coding sequence of TgRad51 gene and the *BamHI* sites was incorporated in both the primers. A single PCR band of *TgRAD51* was observed (Figure A5.B). The PCR product was first cloned into TOPO2.1 TA vector (Invitrogen). The resultant plasmid was digested with *BamHI* to release the insert containing TgRad51 gene (Figure A5.C) and was ligated with pTA plasmid (6282 bp) at *BamHI* site (Figure A5.A). The pTA-*TgRAD51* clone (7347 bp) was confirmed by *BamHI* which released the insert of size 1065 bp and the empty vector backbone corresponding to 6242 bp (Figure A5.D). The right directionality of the insert in this plasmid was confirmed by *SalI* enzyme. Both the vector pTA and the *TgRAD51* gene contained *SalI* site. In right orientation, two bands of sizes 1231 bp and 6116 bp were observed with *SalI* enzyme (Figure A5.E). The *TgRAD51* gene was confirmed by DNA sequencing.

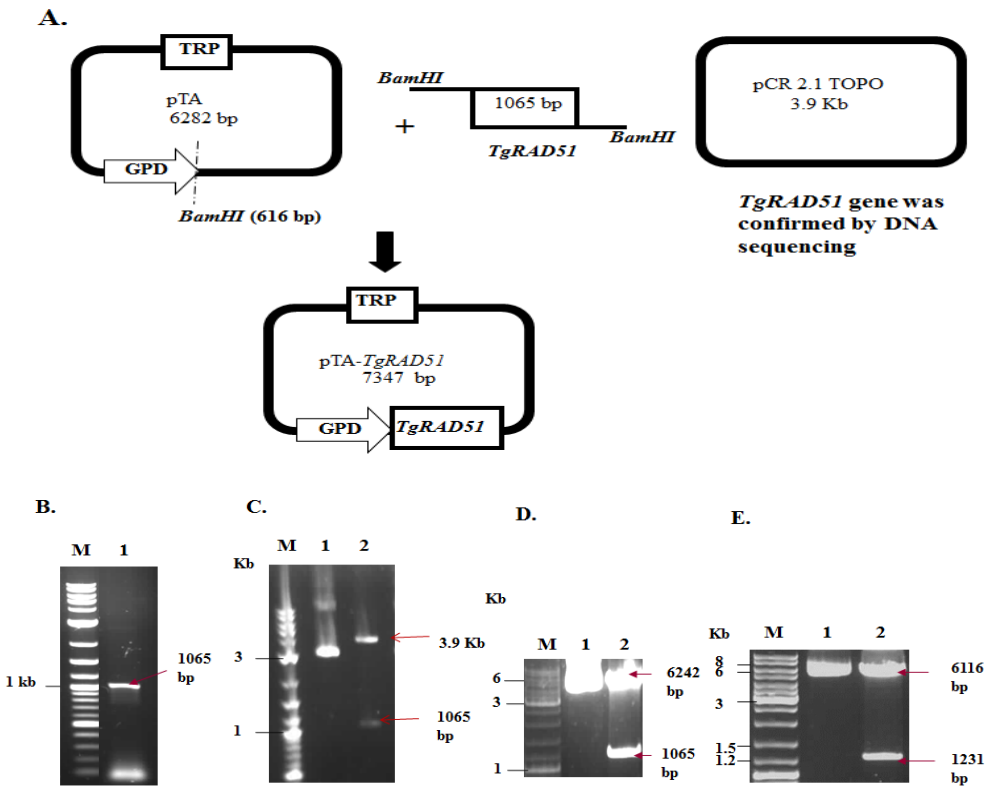

Figure A5: Construction of *TgRAD51* gene in pTA vector. A). The cloning strategy of *TgRAD51* gene from RH strain of *T. gondii* into yeast expression vector, pTA at *BamHI* site, under GPD promoter. B). Lane 1 corresponds to the *TgRAD51* PCR product of size 1065 bp. C). Uncut pCR 2.1 TOPO plasmid containing *TgRAD51* was loaded in lane 1. Lane 2 displayed the release of *TgRAD51* insert of 1065 bp from TOPO vector by *BamHI* digestion. D). Uncut pTA-*TgRAD51* plasmid was loaded in lane 1. The insert was released from pTA-*TgRAD51* plasmid by *BamHI* as observed in lane 2. E). Uncut pTA-*TgRAD51* plasmid was shown in lane 1. The orientation of *TgRAD51* insert in this plasmid was confirmed by *SalI* digestion which produced two bands of 1231 bp and 6116 bp as exhibited in lane 2. Standard DNA marker was represented as M in Figure B to E and the bands are indicated by arrowhead in left panel.

A.1.6 Construction of *TgRAD51* gene in pET28a vector (pET28a:*TgRAD51*)

The plasmid pTA:*TgRAD51* was digested with *BamHI* to release the *TgRAD51* gene and was ligated with pET28a (Novagen) (Figure A6.A). Bacterial protein expression plasmid, pET28a contained T7 promoter, kanamycin selectable marker and an N-terminal His_6 tag. Confirmation of this construct was done by *BamHI* digestion which released the insert of size 1065 bp from the pET28a plasmid (Figure A6.B). The directionality of the insert in this plasmid was checked by *EcoRI* enzyme. Internally a single *EcoRI* site was present in both vector and *TgRAD51* insert. As a result, correct orientation of the gene resulted in two bands of sizes 6.1 Kb and 650 bp (Figure A6.C).

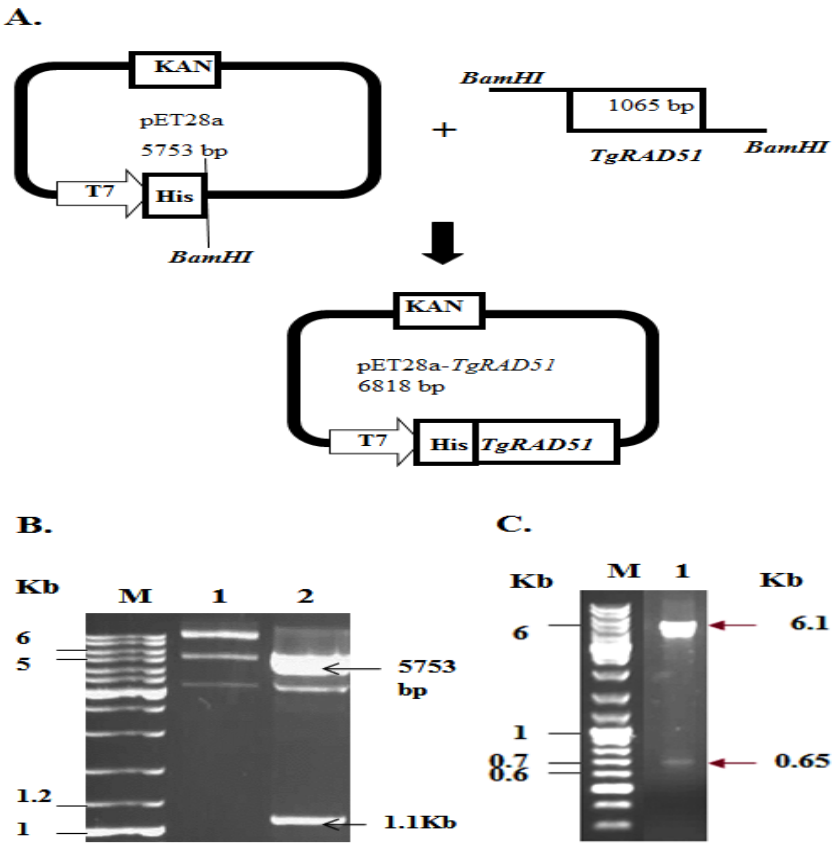

Figure A6: Construction of *TgRAD51* clone in pET28a plasmid. A). Strategy to clone *TgRAD51* gene into pET28a plasmid under T7 promoter at *BamHI* site. B). Lane 1 corresponds to uncut plasmid pET28a:*TgRAD51*. After *BamHI* digestion, pET28a:*TgRAD51* released two bands of sizes 5.7 and 1.1 Kb as shown in lane 2. C). To confirm the right orientation of the gene, pET28a:*TgRAD51* plasmid was digested with *EcoRI* and two DNA fragments of sizes 6.1 Kb and 0.65 Kb were produced as indicated in lane 1. Standard DNA marker was represented as M in Figure B and C and the bands are indicated by arrowhead in left panel.

A.1.7 Construction of *ScRAD51* gene in pTA vector (pTA:*ScRAD51*)

ScRAD51 gene was PCR amplified from the genomic DNA isolated from W303a strain using the forward primer OMKB 90 and the reverse primer OMKB 88. *BamHI* and *PstI* sites were incorporated in the forward and reverse primer respectively. The PCR product of size 1203 bp was initially cloned into pCR 2.1 TOPO vector and then the *ScRAD51* gene was released by *BamHI* and *PstI* and sub cloned into the yeast expression vector, pTA (Figure A7.A). The intermediate clone was digested with *BamHI* enzyme which released the insert of 1203 bp and the empty vector backbone of 3.9 Kb (Figure A7.B). *PstI* and *EcoRI* digestion of this construct yielded three desired DNA fragments of sizes 0.6 Kb, 0.7 Kb and 3.9 Kb (Figure A7.B). These results proved that the construct has *ScRAD51* insert. This insert was released by *BamHI* and *PstI* digestion and further sub cloned into pTA vector. Both the pTA plasmid and the insert contained *EcoRI* sites. Therefore, upon digestion, two desired DNA bands of sizes 1.6 Kb and 5.9 Kb were observed (Figure A7.C). All the cloning was confirmed by sequencing.

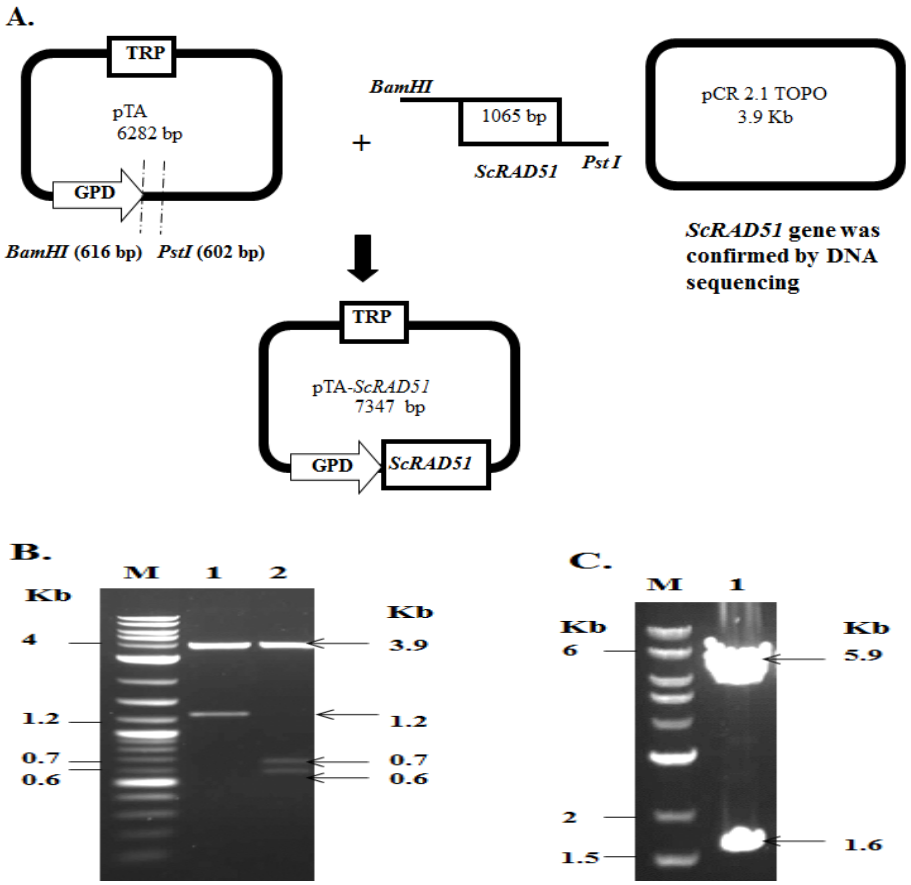

Figure A7: Cloning of *ScRAD51* gene in pTA plasmid. A). Strategy to clone *ScRAD51* in pTA background at *BamHI* and *PstI* sites. B). The pCR 2.1 TOPO:*ScRAD51* construct was confirmed by *BamHI* digestion which released the insert of size 1203 bp as indicated in lane 1. This showed that the orientation of the insert was opposite. Further, the plasmid sample was digested with *PstI* and *EcoRI*. Fragments of desired size 0.6, 0.7 and 3.9 Kb in lane 2 proved that the construct has *ScRAD51* insert. C). Confirmation for pTA:*ScRAD51* clone: The plasmid sample was digested with *EcoRI* which showed fragments of desired sizes 5.9 and 1.6 Kb in lane 1. Standard DNA marker was represented as M in Figure B and C and the bands are indicated by arrowhead in left panel.

A.1.8 Cloning of *KANMX6* cassette into pSD158 vector (pSD158:*KANMX6*)

A 1536 base pair fragment containing *KANMX6* gene was excised from pFA6a-KANMX6 plasmid (Longtine MS et al., 1998) using *NotI* restriction enzyme. This fragment was cloned at the *NotI* site of pSD158 plasmid (Diede SJ et al., 1999), at the upstream of 5'ADH4 sequence (Figure A8.B). In this plasmid ADH4 gene was insertionally inactivated by introduction of *ADE2* gene. The 5'ADH4 and 3'ADH4 encompass 477 and 1527 base pair respectively. The cassette comprised of KANMX6-ADH4-ADE2-ADH4 (Figure A8.A) was released from the modified pSD158 plasmid by *SalI* digestion (Figure A8.B) which was subsequently used for gene targeting assay.

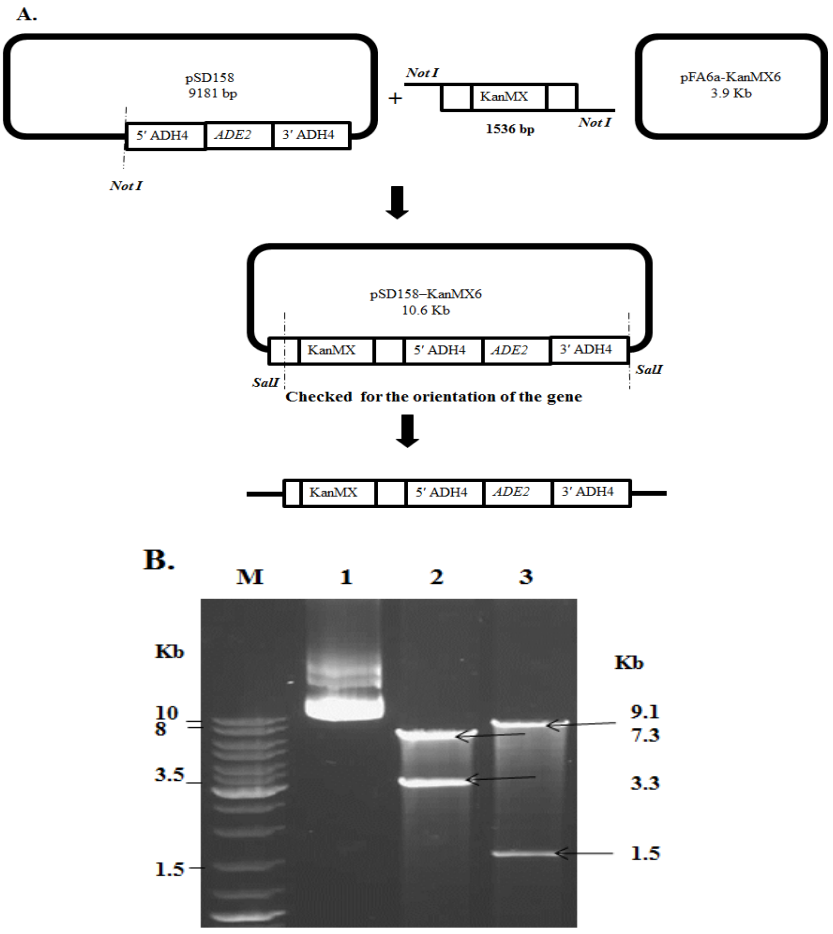

Figure A8: Construction of *KANMX6* clone into pSD158 plasmid. A). Strategy to clone *KanMX6* gene into pSD158 at *NotI* site B). Lane 1 exhibited uncut pSD158-*KanMX6* plasmid. Lane 2 displayed the DNA fragments of sizes 3.3 Kb and 7.3 Kb obtained after *SalI* digestion of pSD158-*KanMX6* plasmid. Lane 3 displayed the *NotI* digested pSD158-*KanMX6* DNA which released the insert of size 1.5 Kb. Arrowheads in the right panel indicated the sizes of DNA fragments released after the digestions. Standard DNA marker was represented as M in Figure B and the bands are indicated by arrowhead in left panel.

APPENDIX - 2

A. 2. 1 Synopsis of the Ph.D thesis:

DNA damage is inevitable. Failure to repair DNA double strand break leads to apoptosis and cancer in case of the multicellular organisms, whereas it causes cell death in unicellular organisms. DNA double strand breaks (DSB) are repaired either by homologous recombination (HR) pathway or by non-homologous end joining (NHEJ) mechanism. When a DSB is formed, HR repair depends on searching of extensive homologous stretches of DNA. Non-homologous end joining of broken DNA ends depends on little or no homology. Both pathways compete with each other for repair. In DNA repair, Mre11 (Meiotic recombination 1 protein) is recruited very early within minutes to the broken ends. If Ku70/80 heterodimeric protein binds to DSB, the repair takes place by NHEJ. But if the ends are resected by nucleases such as Sae2, it forms long 3' single stranded DNA wherein HR mediated repair occurs. During this repair process, Rad51 is recruited which forms nucleoprotein filament (Mimitou EP and Symington LS, 2009). During HR, Rad51 plays an important role in searching of homologous template and promotes strand invasion. Interestingly, the usage frequency of HR over NHEJ is different in different organisms. In higher eukaryotes, like in humans, NHEJ is the predominant mechanism, whereas in lower eukaryotes, like in *Saccharomyces cerevisiae*, HR is the major pathway. Even prokaryotes, such as *E. coli* choose HR as the predominant pathway. Interestingly, two lower eukaryotes, very closely related apicomplexan parasites; *Plasmodium falciparum* and *Toxoplasma gondii* show strikingly opposite choice of the repair pathway. *P. falciparum* apparently lacks NHEJ (based on the genome information) and employs HR for DSB repair and *T. gondii* appears to use NHEJ as predominant pathway. *T. gondii* shares structural similarities with other closely related protozoan

parasites like *Plasmodium, Cryptosporidium* etc As a result, *T. gondii* has become a model system to do functional genomics. However efficient gene targeting is a challenge in this parasite. It has been observed that in this parasite gene targeting efficiency is very low as they demonstrate high degree of non homologous end joining (Donald RG et al, 1998).

In this work, three important questions have been addressed. Firstly, we asked why *Toxoplasma gondii*, despite being a lower eukaryote chooses NHEJ over HR. Further, we also wanted to investigate what determines the pathway choice in this organism and finally, how does the cell enable the chosen pathway once the decision is made.

To address why *Toxoplasma gondii* chooses NHEJ, we identified and cloned *MRE11*, *KU80* and *RAD51* genes of *T. gondii* and in order to understand the interplay between these three proteins during DSB repair. Understanding the functional role of TgMre11 would give insights on pathway choice as it binds to the broken DNA ends much before the decision is made between HR and NHEJ. To understand NHEJ mechanism, we choose the central player TgKu80 for characterization. Further to understand HR pathway and the targeted gene disruption in this parasite, we choose the key player TgRad51 for our study. We have cloned, purified and characterized TgRad51 biochemically as well as genetically. Previously, *Plasmodium falciparum* Rad51 protein (PfRad51) was identified, biochemically characterized and its role in homologous repair was elucidated (Bhattacharyya M K et al, 2003, 2005). Though, TgRad51 and PfRad51 have 82% identity in their catalytic domain, they differ markedly in ssDNA dependent ATP hydrolysis activity. Here we report the mechanistic insights for differential repair choices between these two closely related

lower eukaryotes. We hypothesize that compromised ATPase activity of *TgRAD51* leads to inefficient gene targeting and poor gene conversion efficiency in *T. gondii*.

Since, *T. gondii* is not easily amenable to genetic manipulation, the first objective was to develop a surrogate yeast assay system to study DNA double strand break repair proteins of *Toxoplasma gondii.* Primary sequence analysis revealed that 44% sequence identity is shared in the nuclease domain of TgMre11 with ScMre11 proteins. For the repair of the MMS induced breaks, a functional Mre11 protein is essential. The inability of the cells to survive after MMS treatment in the cell harboring *TgMRE11* clearly showed that TgMre11 was not functional in yeast. Since both the nuclease and DNA damage response (DDR) domains should be functional for the complementation to happen (Bhattacharyya MK et al., 2008), we generated a chimera by fusing the more conserved nuclease domain of *TgMRE11* with the DDR domain of *ScMRE11*. Under MMS treated conditions, growth was not observed in the cell expressing chimeric *MRE11* which clearly stated that ScMre11 was not functionally complemented by TgMre11. To understand the functional role of TgKu80, we performed NHEJ assay in yeast. Here, the DSB generated can only be repaired by NHEJ as there is no homologous template for HR to occur. The NHEJ efficiency of TgKu80 was poor in yeast model system which indicated that there is no functional complementation between TgKu80 and ScKu80. By semi-quantitative RT-PCR, we found that *TgMRE11*, chimeric *MRE11* and *TgKU80* were getting expressed in yeast but they are not functional. To investigate the functional role of TgRad51 in HR pathway, we first looked at its functional complementation in yeast surrogate system. When TgRad51 was expressed in yeast, functional complementation was observed by MMS sensitivity assay. From this part of the study, we concluded that the

yeast model system can be used for the study of TgRad51 but cannot be used for the study of TgMre11, chimeric Mre11 and TgKu80.

Since, gene targeting experiment requires extensive homology searching which is aided by ATPase activity of Rad51 protein, we questioned whether the ATPase activity of TgRad51 is weak and thus it results in inefficient gene targeting. So we cloned, expressed and purified TgRad51 protein and tried to look at its ATP hydrolysis activity. TgRad51 protein was purified to near homogeneity. MALDI and MS-MS analysis matched five peptides to the TgRad51 protein. Single peak was observed, which indicated that there was no other contaminating protein. Full length DNA and protein sequence were submitted to Genbank and the accession number (JQ771675) was assigned. Biochemical characterization of TgRad51 protein revealed weak ssDNA dependent ATPase activity.

In order to investigate whether the weak ATP hydrolysis activity of TgRad51 affects homology search during HR mediated repair, we investigated for the efficiency of TgRad51 mediated gene conversion. Our studies clearly revealed that the weak ATPase activity of TgRad51 protein was responsible for poor homology searching and hence lowered gene conversion efficiency. We observed that TgRad51 has lower gene conversion rate when compared to ScRad51. Western blot clearly showed that both the proteins were expressed at comparable level. Gene targeting studies in yeast demonstrated that TgRad51 displayed preference for random gene integration. *TgRAD51* exhibited about three times reduction in gene targeting efficiency when compared to that of *ScRAD51*. However, with increase in homologous flanking ends, there was an increase in targeted gene integration similar to the trend observed with *ScRAD51*. We also monitored the gene targeting efficiency of TgRad51 in both *KU80* wild type and *ku80* null background. As expected, there

was no significant change in gene targeting efficiency in *ku80* null background. This is because, the absence of Ku80 abrogates NHEJ pathway but does not increase the efficiency of Rad51 recombinase.

Our findings explore the underlying reasons of inefficient HR mediated DNA repair and gene targeting in *Toxoplasma gondii*. We have cloned, expressed and purified recombinant TgRad51 and have shown that it possesses ATP hydrolysis activity, which is the lowest among all the eukaryotic Rad51 proteins studied so far. Three independent experiments- repair choice, targeted gene knock-in, targeted gene knock-out were carried out in yeast to genetically characterize TgRad51. Our studies revealed that the Rad51 dependent gene conversion pathway is compromised in cells harboring TgRad51. Since the HR machinery is weak in *T. gondii*, targeted gene disruption or tagging of endogenous genes are very less efficient in this parasite. Works from Bzik laboratory and Carruthers laboratory (Fox BA et al., 2009; Huynh MH et al., 2009) have demonstrated enhanced gene targeting in *ku80* null *T. gondii*. However, it does not increase the integration efficiency at the correct locus. This notion is supported by the finding that targeted repair of Δ*hxgprt* became independent of TgKu80 when enough flanking homology (910 bp) was provided (Fox BA et al., 2009). Results from our gene targeting experiments suggest that Rad51 mediated homology search must be less inefficient in this parasite. As a result, TgRad51 preferred random integration. But, the increase in homology favored more of targeted gene integration. The absence of Ku80 in yeast assay system did not increase the efficiency of TgRad51 recombinase during gene targeting. Findings from this study significantly contribute towards the development of a transgenic strain of *T. gondii*, wherein TgRad51 can be replaced by Rad51 of other organisms like *Leishmania*. LmRad51 ATPase activity is 85 fold more than TgRad51. The flanking homologous

sequences can be increased up to 1 Kb to facilitate more of the targeted integration. This strain with improved gene targeting efficiency can also be used to study the functions of the genes of *Plasmodium falciparum* and other apicomplexan parasites. This strain can thus would facilitate genetic dissection of protozoan parasite biology which can lead to new treatments for significant parasitic diseases.

REFERENCES

1. Bhattacharyya MK, Kumar N, 2003: Identification and molecular characterization of DNA damaging agent induced expression of *Plasmodium falciparum* recombination protein PfRad51. International Journal for Parasitology, 33: 1385-1392.
2. Bhattacharyya MK, Bhattacharyya S, Jayabalasingham B, Kumar N, 2005: Characterization of kinetics of DNA strand-exchange and ATP hydrolysis activities of recombinant PfRad51, a *Plasmodium falciparum* recombinase. Molecular and Biochemical Parasitology, 139: 33-39.
3. Bhattacharyya MK, Mathews KM, Lustig AJ, 2008: Mre11 nuclease and C-terminal tail-mediated DDR functions are required for initiating yeast telomere healing. Chromosoma, 117:357-366.
4. Donald RG, Roos DS, 1998. Gene knock-outs and allelic replacements in *Toxoplasma gondii*: HXGPRT as a selectable marker for hit-and-run mutagenesis. *Mol Biochem Parasitol*. 91(2):295-305.
5. Fox B A, Ristuccia J G, Gigley J P and Bzik D J, 2009: Efficient gene replacements in *Toxoplasma gondii* strains deficient for non-homologous end joining. Eukaryotic cell, 8(4): 520-529.
6. Huynh MH and Carruthers VB, 2009: Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. Eukaryotic cell, 8(4): 530-539.
7. Mimitou EP, Symington LS, 2009: DNA end resection: many nucleases make light work. DNA Repair, 8(9):983-95.

APPENDIX -3

A. 3. 1 Published research paper:

Sita Swati Achanta, Shalu M Varunan, Sunanda Bhattacharyya, Mrinal Kanti Bhattacharyya, 2012: Characterization of Rad51 from apicomplexan parasite *Toxoplasma gondii*: an implication for inefficient gene targeting. *PLoS ONE*,7(7): e41925.

OPEN ACCESS Freely available online

PLoS one

Characterization of Rad51 from Apicomplexan Parasite *Toxoplasma gondii*: An Implication for Inefficient Gene Targeting

Sita Swati Achanta[1], Shalu M. Varunan[1], Sunanda Bhattacharyya[2], Mrinal Kanti Bhattacharyya[1]*

1 Department of Biochemistry, School of Life Sciences, University of Hyderabad, Hyderabad, Andhra Pradesh, India, **2** Department of Biotechnology, School of Life Sciences, University of Hyderabad, Hyderabad, Andhra Pradesh, India

Abstract

Repairing double strand breaks (DSBs) is absolutely essential for the survival of obligate intracellular parasite *Toxoplasma gondii*. Thus, DSB repair mechanisms could be excellent targets for chemotherapeutic interventions. Recent genetic and bioinformatics analyses confirm the presence of both homologous recombination (HR) as well as non homologous end joining (NHEJ) proteins in this lower eukaryote. In order to get mechanistic insights into the HR mediated DSB repair pathway in this parasite, we have characterized the key protein involved in homologous recombination, namely TgRad51, at the biochemical and genetic levels. We have purified recombinant TgRad51 protein to 99% homogeneity and have characterized it biochemically. The ATP hydrolysis activity of TgRad51 shows a higher K_M and much lower k_{cat} compared to bacterial RecA or Rad51 from other related protozoan parasites. Taking yeast as a surrogate model system we have shown that TgRad51 is less efficient in gene conversion mechanism. Further, we have found that TgRad51 mediated gene integration is more prone towards random genetic loci rather than targeted locus. We hypothesize that compromised ATPase activity of TgRad51 is responsible for inefficient gene targeting and poor gene conversion efficiency in this protozoan parasite. With increase in homologous flanking regions almost three fold increments in targeted gene integration is observed, which is similar to the trend found with ScRad51. Our findings not only help us in understanding the reason behind inefficient gene targeting in *T. gondii* but also could be exploited to facilitate high throughput knockout as well as epitope tagging of *Toxoplasma* genes.

Citation: Achanta SS, Varunan SM, Bhattacharyya S, Bhattacharyya MK (2012) Characterization of Rad51 from Apicomplexan Parasite *Toxoplasma gondii*: An Implication for Inefficient Gene Targeting. PLoS ONE 7(7): e41925. doi:10.1371/journal.pone.0041925

Editor: Gordon Langsley, Institut national de la santé et de la recherche médicale - Institut Cochin, France

Received March 22, 2012; **Accepted** June 26, 2012; **Published** July 30, 2012

Funding: This work is supported by a grant from the Indian funding agency, Department of Biotechnology, India, to MKB (grant BT/PR 11174/MED/29/98/2008). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: mkbsl@uohyd.ernet.in

Introduction

A broken chromosome if unrepaired leads to loss of genetic information and cell death. It is thus very important for a cell to identify and repair the damaged region. There are two major pathways by which double stranded DNA breaks (DSB) can be repaired. In homologous recombination (HR) mediated repair it relies on searching of extensive homologous stretches of DNA, whereas non homologous end joining (NHEJ) requires little or no homology. These two pathways compete with each other for repairing a DSB. It is a complex phenomenon and how cell decides which pathway to choose is not clear yet. Also different organisms show distinct usage frequency of HR over NHEJ during DSB repair. If a DSB is occupied by Ku70/80 hetero dimer, NHEJ follows. However, as soon as a DSB is formed, if the ends are resected by 5′-3′ exonuclease to form a long 3′ ss DNA, HR sets into action. This is because the long ssDNA overhang inhibits the loading of Ku70/80 and recruits the recombination proteins [1]. Depending upon the nature and orientation of the homologous sequences with respect to the break, a DSB can be repaired by any of the three different HR pathways: the single strand annealing (SSA), break induced replication (BIR) and the gene conversion (GC), also known as synthesis dependent strand annealing (SDSA) [2,3,4]. Presence of homology at both the DSB ends leads to GC. When homology is present only at one DSB end, BIR follows and when a DSB is flanked by direct repeats, SSA is the mechanism of choice. In the case of SSA, a minimum of 30 base pair homologous sequence on either side of the DSB is sufficient and such repair leads to the deletion of intervening sequences. All of the above mentioned pathways are driven by several proteins that belong to Rad52 epistatis group [5]. Strand invasion is the central step of HR mechanism and it requires Rad51 protein. However, Rad51 is essential only for GC, not for BIR or SSA [6,7]. Rad51 protein has ATP dependent DNA binding activity, it multimerizes on single stranded DNA to form helical filament similar to that formed by bacterial RecA protein [8]. In an ATP hydrolysis dependent manner its motor activity searches for the homologous sequences between a single strand DNA and double stranded DNA and catalyses the strand exchange reaction. During this process it interacts with replication protein A (RPA), Rad52, Rad54 and Rad55 [9,10,11].

Toxoplasma gondii belongs to the eukaryotic phylum Apicomplexa, which infects about one third world population and under immune compromised condition it can cause serious illness. This

protozoan parasite shares a number of structural similarities with disease causing parasites namely *Plasmodium*, *Cryptosporidium* etc. Thus it has become a model system to do functional genomics. However efficient gene targeting is a challenge in this parasite. It has been observed that in this parasite gene targeting efficiency is very low as they demonstrate high degree of non homologous end joining [12]. It is interesting to note that *T. gondii* is the only protozoan parasite that harbors non homologous end joining mediated DNA break repair mechanism. A divergent eukaryotic parasite *Trypanosoma brucei* possesses Ku70/80, however TbKu70/80 function in telomere maintenance [13] and Ku dependent NHEJ was not observed in *T. brucei* cell extract [14,15]. The Ku70/80 heterodimer binds at the DSB along with DNA ligase IV-Xrcc4 complex and DNA-PKc to catalyze the break repair [16]. In a ku80 knock out background there is a 300–400 fold increase in targeted gene disruption in *T. gondii* [17]. The key protein, TgRad51, involved in targeted gene disruption has not been characterized yet. We have cloned, purified and characterized TgRad51 biochemically as well as genetically. It is observed that *T. gondii* and *Plasmodium falciparum* show strikingly opposite choice in DNA repair pathways. While *Plasmodium falciparum* depends solely on HR and apparently lacks NHEJ, *T. gondii* prefers NHEJ. Here we report the mechanistic insights for differential repair choices between these two closely related lower eukaryotes. PfRad51 has been identified as a DNA repair protein and has been speculated to play major role during mitotic recombination [18,19]. PfRad51 has also been characterized biochemically [20] and despite having 82% identity in the catalytic domain, the kinetics of ssDNA dependent ATP hydrolysis activity differs markedly between PfRad51 and TgRad51. We hypothesize that compromised ATPase activity of TgRAD51 leads to inefficient gene targeting and poor gene conversion efficiency in *T. gondii*. However, with increase in homologous flanking ends, we observe an increase in targeted gene integration similar to the trend observed with ScRad51.

Materials and Methods

Yeast Strains

The yeast strains used in this work are tabulated in Table 1. The yeast expression vectors harboring *TgRAD51* and *ScRAD51* were transformed into the *Δrad51* (LS402) strain to create MVS26 and NRY2 strains respectively. The empty vector pTA [21] was transformed in *Δrad51* strain to generate NRY1. NA14 and NA14Δrad51 strains were used [4] for gene conversion study. We transformed the above mentioned yeast expression vectors with *TgRAD51* or *ScRAD51* to NA14Δrad51 strain to generate SSY1 and SSY2 strains respectively. To knockout Ku80 gene from a series of isogenic strains *KANMX* gene flanked by up-and down-sequences of *KU80* ORF was amplified from an already existing yeast *ku80Δ* strain. The primer pair OSB127 (AGT CTA TTA GCG GAA GTA CC) and OSB128 (GAA CGT CCT CTA CCC ACG) resulted in 200 bp and 220 bp flanking sequence of *KU80* on each side of *KANMX* gene. This linear *KANMX* cassette of 2220 bp was used to knockout *KU80* from the *rad51Δ* (LS402) strain. The yeast expression vector pTA was transformed into the *Δrad51Δku80* (SSY3) to generate SSY4 strain. The yeast expression vectors harboring *ScRAD51* and *TgRAD51* were transformed into the *Δrad51Δku80* (SSY3) strain to create SSY5 and SSY6 strains respectively.

Construction of the pET: TgRad51 Expression Vector

The open reading frame encoding *Toxoplasma gondii* Rad51 was amplified from the cDNA library (provided by Prof. Vern Carruther) of *T. gondii* RH strain using hot start Kapa Hifi DNA polymerase (Kapa Biosystems) as described by the manufacturer. The forward primer OMKB65 (GGATCCATGAGCGCCGTCTCTCTTCAG) and reverse primer OMKB82 (GGATCC TCAGTTGTCTTCGTAGTCGCC) are complementary to the 5′ and 3′ ends of the coding sequence of TgRad51 gene and the underlined oligo nucleotide represents the *BamHI* sites incorporated in both the primers. The PCR product of size 1065 bp was first cloned into TOPO2.1 TA vector (Invitrogen). The resultant plasmid was digested with *BamHI* to release the insert containing TgRad51 gene and was ligated with pET28a (Novagen) to construct N-terminal His_6 tagged recombinant protein. The insert was sequenced and submitted to the Gene Bank (JQ771675). All primers used in this study were purchased from Sigma Aldrich.

Plasmids

The *TgRAD51* gene was subcloned into *BamHI* site of the 2 μ expression plasmid, pTA under the control of GPD promoter. In a similar way, *ScRAD51* was PCR amplified from the genomic DNA isolated from W303a strain using the forward primer OMKB90 (GGATCCTGTCTCAAGTTCAAGAAC) and the reverse primer OMKB88 (CTGCAGCTACTCGTCTTCTTCTC), the underlined oligo nucleotides represented *BamHI* and *PstI* sites incorporated in the forward and reverse primer respectively. The PCR product of size 1203 bp was cloned into the yeast expression vector pTA under GPD promoter. All the cloning was confirmed by sequencing.

Protein Purification

The expression vector pET28a:TgRad51 having N-terminal His_6 tag was transformed into *Escherichia coli* host strain Rosetta (DE3). The whole transformation mixture was added to 10 ml LB media containing chloramphenicol and kanamycin and incubated at 37°C for overnight. Next morning 5% of the primary culture was added to 100 ml fresh LB media (containing chloramphenicol and kanamycin) and tested for the IPTG induced expression of the TgRAD51 protein. After OD_{600} reached 0.8, protein was induced by IPTG (Sigma) and grown on Luria broth for 4 hours at 37°C. About 1.0 gm cell pellet was suspended in 3 ml lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 30% Glycerol, 20 mM β-mercaptoethanol) containing 1 mg/ml lysozyme and protease inhibitor (PMSF). It was then lysed on ice by sonication using (SONICS Vibra^m Cell^{TM}) giving ten 40 second bursts at 200 W with intermittent cooling. After lysis the cell debris was spun down by centrifugation for 45 minutes at 10,000 g at 4°C. The supernatant was applied to 0.75 ml 50% Ni-NTA agarose (Qiagen) and the loading flow through was collected. The resin was next washed with 16 volume of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM Imidazole, 30% glycerol, 20 mM β-mercapitoethanol, 2% Tween 20). Further the column was washed with 5 volume of wash buffer 2 (50 mM NaH_2PO_4, 300 mM NaCl, 50 mM imidazole, 30% glycerol, 20 mM β-mercapitoethanol, 2% Tween 20). Finally the required protein was eluted first with buffer containing 250 mM imidazole (5 column volume) and then with 400 mM imidazole (50 mM NaH_2PO_4, 300 mM NaCl, 400 mM Imidazole, 30% glycerol, 20 mM β-mercapitoethanol, 2% Tween 20). The aliquots from the indicated steps in the elution profile were separated using 10% SDS-PAGE and visualized by Coomassie brilliant blue G-250 (Bio-Rad). The protein fractions eluted with 400 mM imidazole were pooled and dialysed against the dialysis buffer containing 20 mM Tris HCl (pH = 8), 1 mM dithiothretol (DTT), 5% glycerol. The concentration of the purified recombinant TgRad51 protein was

Table1. Yeast strains used in this study.

Strains	Genotype
W303a	*MATa 15ade2-1, ura3-1, 112 his 3-11, trp1, leu2-3*
NRY1	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] RAD51:: LEU2 pTA*
NRY2	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] RAD51::LEU2 pTA/ScRAD51*
MVS26	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] RAD51::LEU2 pTA/TgRAD51*
NA14	*MATa-inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100*
NA14Δrad51	*MATa-inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100 RAD51::LEU2*
SSY1	*MATa-inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100 RAD51::LEU2 pTA/ScRAD51*
SSY2	*MATa-inc ura3-HOcs lys2::ura3-HOcs-inc ade3::GALHO ade2-1 leu2-3,112 his3-11,15 trp1-1 can1-100 RAD51::LEU2 pTA/TgRAD51*
SSY3	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] RAD51::LEU2 KU80::KANMX*
SSY4	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] RAD51::LEU2 KU80::KANMX pTA*
SSY5	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] RAD51::LEU2 KU80::KANMX pTA/ScRAD51*
SSY6	*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+] RAD51::LEU2 KU80::KANMX pTA/TgRAD51*
SLY3	*MATa SBA1:: KANr*
Yku80Δ	*MATα leu2-3 112 trp1-1 can1-100, ura3-1 ade2-1 his 3-11,15 adh4::Ade2 rad5+ KU80::KANMX*
NKY8	*MATa SPC29-CFP:: KAN mRFP-TETR URA3::TETO GFP-LACI::LEU2 CHL1::HIS3*

doi:10.1371/journal.pone.0041925.t001

determined by UV absorbance at 280 nm by using the extinction coefficient of 21025 M^{-1} cm^{-1} as calculated from the amino acid sequence by using ExPaSy Protparam tool. The purified protein was further confirmed by MALDI-TOF analysis and MS sequencing.

ATPase Assay

The rate of ATP hydrolysis of TgRad51 was determined using Enz Chek Phosphate assay kit (Molecular Probe, E-6646). A typical reaction mixture was composed of 2 μM TgRad51 along with 30 fold molar excess of ϕxssDNA (New England Biolab) (60 μM) in presence of ATP containing reaction buffer (25 mM Tris-HCl pH 8, 5% glycerol, 1 mM DTT and 1 mM $MgCl_2$). The reaction mixture was incubated at 37°C for 35 minutes and at every 5 minutes time interval a measured reaction volume was added (at 22°C) to purine nucleoside phosphorylase (PNP) reaction mix (manufacturer description) for 30 minutes. The ATP hydrolysis results in the formation of inorganic phosphate which was then reacted with MESG (2-amino-6-mercapto 7-methyl purine riboside) generating 2-amino-6-mercapto 7-methyl purine having an absorbance at 360 nm. This reaction is catalyzed by PNP. At a typical ATP concentration, the amount of inorganic phosphate released was plotted against different time interval and from the slope of the straight line, rate of the reaction was calculated. For determining the kinetic parameters the reaction was done at various concentrations of ATP (20 μM - 600 μM). The rate of the reaction was plotted using Graph pad Prism software against each concentration of ATP to determine the value of K_M and k_{cat}.

Western Blotting

The expression of histidine tagged TgRad51 protein in bacterial cell was detected by using anti-His antibody (Santa Cruz Biotechnology Inc., CA). In order to detect endogenous level of TgRad51 and ScRad51 from NRY2, MVS26, SSY1 and SSY2, the protein was isolated according to the protocol as described in [21]. TgRad51 level was detected using anti PfRad51 antibody [18] in 1:4000 dilution. The anti ScRad51 antibody (Santa Cruz Biotechnology Inc., CA) and anti Actin antibody (Abcam) were used in 1:5000 dilution. Horseradish peroxide-conjugated rabbit IgG (Promega) was used as the secondary antibody for ScRad51 at 1:10,000 dilution and HRP conjugated mouse IgG (Promega) was used as a secondary antibody for PfRad51 and actin at the same dilution. Proteins were visualized by an enhanced chemiluminescence system developed by the manufacturer (pierce). All the bands were normalized against Actin.

Gene Conversion Assay

In NA14 strain, there is a cassette integrated in chromosome V where two consecutive *URA3* genes are separated by *KANMX6*. The first *ura3* gene is having a HO endonuclease recognition site in it. This strain expresses HO endonuclease when grown on galactose medium. Initially single colonies from each of the three strains NA14Δrad51, SSY1 and SSY2 were grown to logarithmic phase in the medium containing glycerol as a carbon source. About 1000 cells of each strain were then plated on two kinds of medium; one of them containing glucose and the other containing galactose as carbon sources. Colonies were counted on both the conditions after incubation for 3 days at 30°C. Galactose induced expression of HO creates a double stranded break at one of the *URA* locus and the repair efficiency was calculated by measuring the ratio of the number of colonies grown on the medium containing galactose, compared to that grown on medium containing glucose (no HO induction). Since such DSB repair could be achieved either by GC or SSA, in order to determine the efficiency of repair choice, we had replica plated the cells grown on galactose medium on the plates containing G418 sulphate. This assay was done thrice and the mean value was plotted using Graph Pad Prism5.

Gene Targeting Assay

A 1536 base pair fragment containing *KANMX6* gene was excised from pFA6a-KANMX6 plasmid [22] using *NotI* restriction enzyme. This fragment was cloned at the *NotI* site of pSD158 plasmid [23], at the upstream of 5′ADH4 sequence. In this plasmid ADH4 gene was insertionally inactivated by introduction of *ADE2* gene. The 5′ADH4 and 3′ADH4 encompass 477 and 1527 base pair respectively. The cassette comprised of KANMX6-ADH4-ADE2-ADH4 was released from the modified pSD158 plasmid by *SalI* digestion which was subsequently used for gene targeting assay. This cassette was transformed to *Δrad51* (NRY1), NRY2 and MVS26, where efficient gene targeting at the ADH4 locus by homologous recombination resulted in expression of *ADE2* gene but did not confer resistivity to G418 sulphate. The transformed colonies were subsequently selected on SC-ade-trp plates. The colonies grown in individual strains were replica plated on G418 sulphate containing SC-trp plate. Colonies sensitive to G418 sulphate represents targeted integration of *ADE2* gene, while G418 sulphate resistant colonies correspond to random integration. Finally we calculated the % gene targeting efficiency by using the following formulae:

$$\%\,\mathrm{GeneTargetingEfficiency} = (\mathrm{NumberofG418sulphatesensitivecolonies})/(\mathrm{TotalnumberofADE2^{+}colonies}) \times 100.$$

In each case the value obtained for *Δrad51* strain was subtracted. Each experiment was repeated three times. In another gene targeting assay, we monitored the gene targeting efficiency with increasing stretches of homologous flanking sequences (200 bp, 500 bp or 1000 bp). For that purpose we started with SLY3 strain where the *SBA1* gene was already insertionally inactivated by a *KANMX6* cassette [21]. We had amplified that *KANMX6* cassette by three different sets of primers positioned at 200 bp/500 bp/1000 bp upstream and 200 bp/500 bp/1000 bp down stream of the *KANMX6* insertion. The primer pair OSB5 (TGCTACCCGCCTTCGAGTG) and OSB6 (CACATA-CAGTTCCATTACTTGAC) resulted in 200 bp flanking sequence on each side of *KANMX6* cassette. Similarly OMKB193 (CTCAGAAGAATTTCGTAAATCGG) and OMKB194 (GGA-GATGGTACCGGTTAAGCG) produced 500 bp flanking sequences of *SBA1* on either side of *KANMX* and OMKB191 (TCACACGTCCGTCATGTCTAC) and OMKB192 (GTCCTGCAGGAGACTTATTAGC) amplified 1 Kb flanking regions on either side of *KANMX6* cassette. These three different PCR products were individually transformed to *Δrad51* (NRY1), NRY2 and MVS26 strain to knock out *sba1* by homologous recombination and selected on SC-trp containing G418 sulphate. The number of G418 sulphate resistant colonies was measured. We compared the efficiency of targeted integration in the *rad51Δ* and *rad51Δku80Δ* strains. The gene targeting efficiency was monitored in these strains background with increasing length of homologous flanking sequences (500 bp or 1000 bp). In NKY8 strain, the *CHL1* gene was already insertionally inactivated by a *HIS* cassette. We had amplified the *HIS* cassette by two sets of primers positioned at 500 bp/1000 bp upstream and 500 bp/1000 bp downstream of the *HIS* insertion. The primer pair OMKB 211 (CAG CTC TCT AGC CAA CAG CAG) and OMKB 212 (CTT GCG TAT TAT CTA TAG CGG C) resulted in 500 bp flanking sequence on each side of *HIS3* marker. Similarly OMKB 213 (CAC TCG TTG ACT AGT TCA GAG G) and OMKB 214 (GAC GAA CTT CAT GTG ACG GCT G) produced 1 Kb flanking sequences on each side of the *HIS3* marker. These two PCR products were individually transformed into NRY1, NRY2 and MVS26 strain to knock out *CHL1* gene by homologous recombination and the transformants were selected on SC-trp-his plate. The numbers of colonies were measured. Similar experiments were done with SSY4, SSY5 and SSY6 strain to replace *CHL1* gene with *HIS3* marker by homologous recombination and the transformants were selected on Sc-Trp-His plates. The numbers of His+ colonies were counted. Integration at the targeted loci was determined by PCR. Each experiments were repeated at least three times and was plotted using Graph Pad Prism with error bars.

Results

Identification of *Toxoplasma gondii* Rad51 Open Reading Frame

A BLAST search of ToxoDB database (www.toxodb.org) have identified three ORFs corresponding to putative RAD51 orthologs from ME49 strain (TGME49_072900), VEG strain (TGVEG_020310) and GT1 (TGGT1_112080) strain. Based on these sequences we have designed PCR primers and amplified TgRad51 ORF from RH strain using cDNA library as template (a kind gift from Dr. Vern Carruthers). Sequence comparison of the amplified ORF indicated that it is an ortholog of Rad51. We found little sequence polymorphism of RAD51 gene among the four strains of *T. gondii* (Figure S1). The deduced amino acid sequence of TgRad51 protein contains 354 amino acids with a predicted molecular mass of 38,796 and isoelectric point (pI) value of 5.77. Further sequence comparison of TgRad51 protein with other known eukaryotic Rad51 orthologs revealed a high degree of overall identity (52 to 72%) and even higher degree of conservation (65 to 82%) in the core catalytic domain. Computer alignment using Clustal method (Meg align, DNA star) revealed a near perfect alignment and identified the two nucleotide binding Walker motifs (Figure 1). Phylogenetic analysis revealed that TgRad51 is closer to another apicomplexan parasite PfRad51.

Purification of Recombinant TgRad51 Protein

The expression system Rosetta (DE3)/pET:TgRad51 was used for the purification of TgRad51 protein. The expression vector pET:TgRad51 was transformed into Rosetta (DE3) cells, which has the pRARE plasmid that supplies the tRNA for six codons that are rarely used in *E. coli* (AUA, AGG, AGA, CUA, CCC and GGA). Figure 2A (lane 2) shows that TgRAD51 protein corresponds to about 25% of the total proteins in the crude cell extract after IPTG induction. The induced recombinant protein was predominantly detected in the soluble fraction (lane 3). Figure 2A shows the purification profile of recombinant TgRad51 protein through Ni-NTA column. Glycerol (30%) was included in the wash buffer in order to remove the unwanted hydrophobic interactions between non specific proteins attached with TgRad51. The elution at 250 mM imidazole results in about 90% purified protein with minor protein contaminants (lane 6). However, 400 mM imidazole containing elution buffer resulted in greater than 99% purified protein as judged by SDS-polyacrylamide gel electrophoresis (lane 7). Finally the purified protein was dialyzed thoroughly against Tris-HCl (pH 8) to remove excess imidazole.

Mass spectroscopic analysis confirmed the presence of TgRad51 protein and yielded a molecular mass of 39,398 Da in agreement with the molecular mass of 38,796 Da predicted by ExPasy Protparam tool. MS-MS analysis yielded sequence of 5 peptides

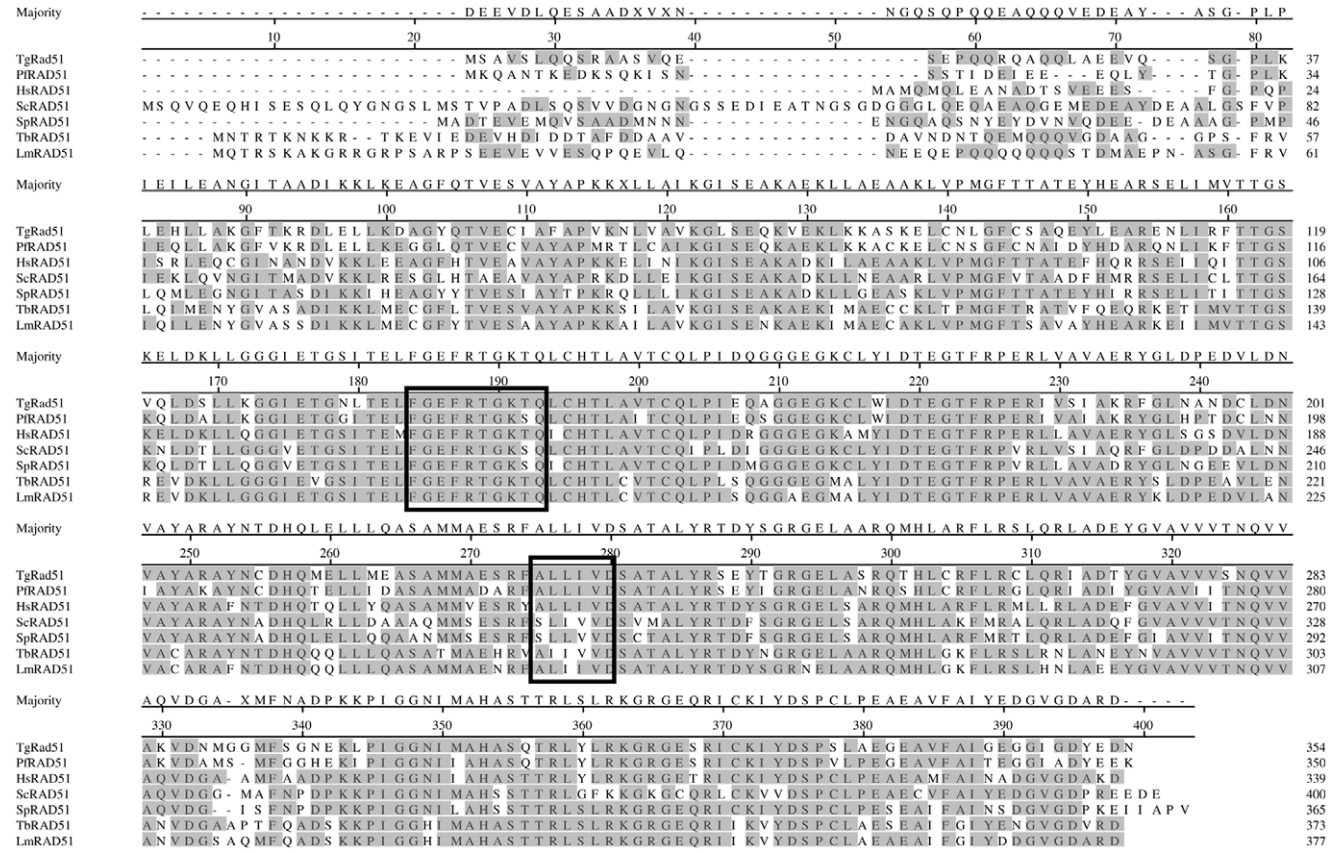

Figure 1. Alignment of TgRad51 with *Plasmodium falciparum* (PfRad51), human (HsRad51), *Saccharomyces cerevisiae* (ScRad51), *Schizosaccharomyces pombe* (SpRad51), *Trypanosoma bricei* (TbRad51) and *Leishmania major* (LmRad51) using Clustal method (Meg align, DNA star). Shaded areas represent identical amino acids. The two DNA binding Walker motifs are boxed.
doi:10.1371/journal.pone.0041925.g001

corresponding to the predicted amino acid sequence of TgRad51 (Figure S2). The western blot analysis with anti histidine antibody confirmed the presence of purified TgRad51 (Figure 2B).

ssDNA Dependent ATP Hydrolysis of TgRAD51

In order to investigate whether the low gene targeting efficiency in *T. gondii* is due to a compromised Rad51 protein activity, we determined the ATP hydrolysis activity of purified TgRad51 protein. We used EnzChek Phosphate Assay kit (Molecular Probes) to determine the rate of ATP hydrolysis of TgRad51. In this assay the inorganic phosphate when combines with the substrate MESG (2-amino-6-mercapto-7-methylpurine riboside) in presence of the enzyme PNP (purine nucleoside phosphorylase) produces ribose 1-phosphate and 2-amino-6-mercapto-7-methylpurine. Formation of the product changes the absorbance of the substrate from 330 nm to 360 nm which can be measured spectroscopically. Figure 2C shows the steep increment of the absorbance at 360 nm with increasing concentration of the inorganic phosphate. When we incubated TgRad51 with ATP at 37°C in presence of ϕxssDNA, ATP is hydrolyzed to release phosphate. At different time interval aliquots of this reaction mixture were withdrawn and added to a second reaction mixture containing MESG along with PNP. Accordingly we monitored the kinetics of ATP hydrolysis of TgRAD51 in presence of ssDNA. Figure 2D shows that TgRAD51 is incapable of ATP hydrolysis in absence of ssDNA. However in presence of 30 fold molar excess of ssDNA, ATP hydrolysis took place. At 300 μM ATP concentration, the rate of ATP hydrolysis as calculated from the slope of the curve was found to be 0.027 μM/min. The ATPase activity of TgRad51 was measured at different ATP concentrations and used to plot Michaelis Menten curve (Figure 2E). The data was also used to generate Lineweaver Burk plot. The k_{cat} value ($0.034\ min^{-1}$) and K_m value for TgRad51 (81.53 μM) were calculated using Graph Pad Prism. Our results indicate that the ATP hydrolysis activity of TgRad51 protein is indeed very weak.

Gene Conversion Efficiency of TgRAD51 is Poor Compared to that of ScRAD51

The ATP hydrolysis activity of Rad51 protein is essential for its motor activity, which is instrumental for the genome-wide search for a homologous template during the pre-synapsis stage of homologous recombination. In order to investigate how the weak ATP hydrolysis activity of TgRad51 affects homology search, we have assayed gene conversion (GC) efficiency of TgRad51. Repair of a DSB by GC relies on Rad51 mediated homology search. To this end we have used yeast as a surrogate model system where GC could be assayed accurately and quantitatively. We used a yeast strain NA14 having two copies of *URA3* gene in chromosome V with a *KANMX6* selectable marker in between (Figure 3A). One of the *ura3* genes is made nonfunctional by incorporating HO endonuclease site in it. The functional *URA3* gene is present in the same chromosome about 3 kb apart. The strain harbors *HO* endonuclease from GAL1 promoter, which is activated by growing the cells in presence of galactose as a carbon source. When a DSB is generated in *ura3* gene by induction of *HO* endonuclease it can

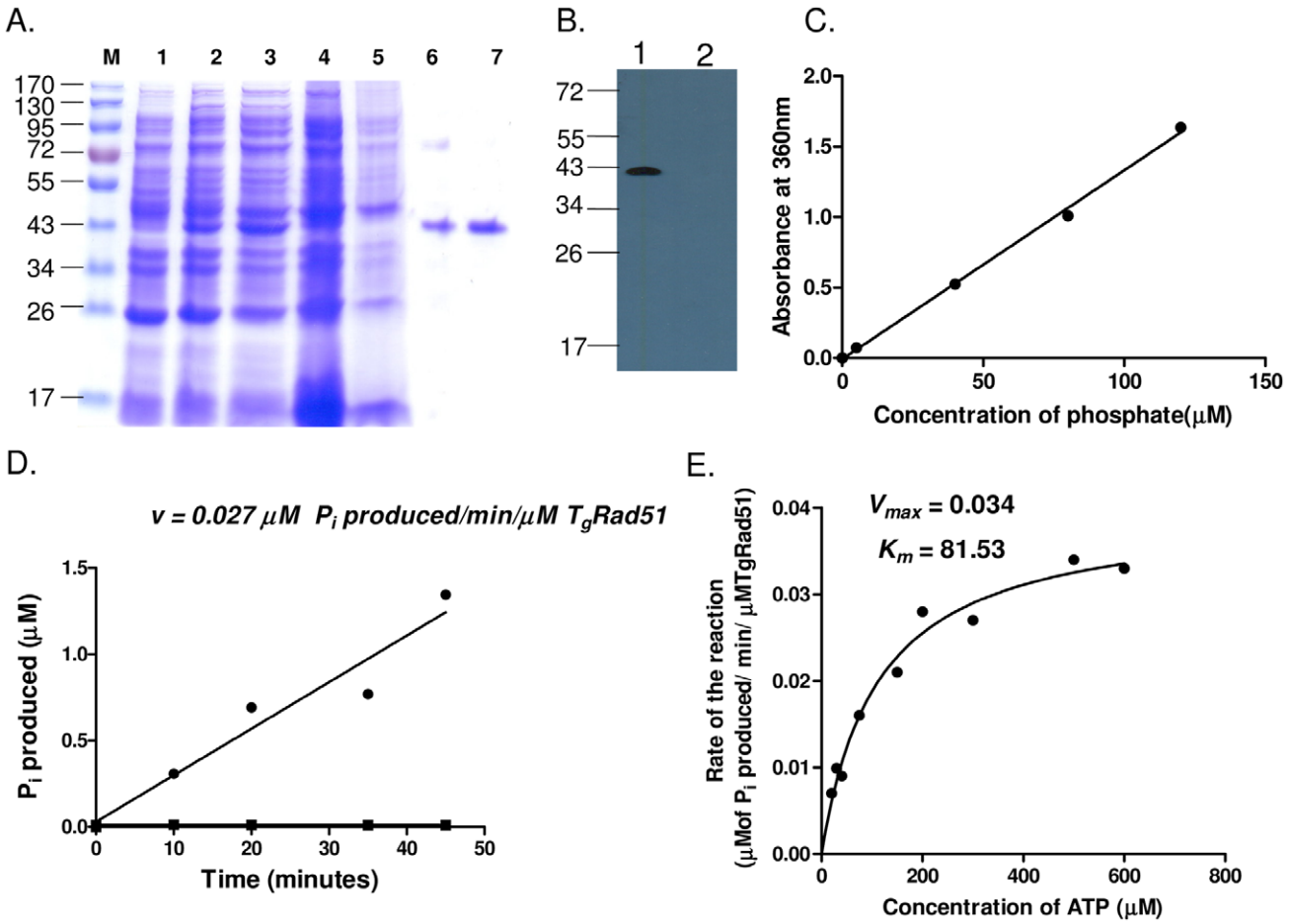

Figure 2. Purification and ssDNA dependent ATPase activity of recombinant TgRAD51 protein. (A) M corresponds to molecular weight standards, the sizes of marker proteins in kDa are indicated, Lane 1 and 2 are uninduced and induced cell free extracts respectively, Lane 3 corresponds to the proteins in soluble fraction, Lane 4 being the loading flow through and Lane 5 is the washing flow through. Lane 6 and 7 correspond to the eluted fractions with increasing imidazole concentrations i.e. 250 mM and 400 mM respectively. B) Western blot analysis shows the presence of Histidine tagged TgRAD51 in lane 1 from the cells bearing pET-TgRAD51 and lane 2 shows the absence of the protein in the cell bearing the empty vector pET. C) The standard curve showing linear relationship between inorganic phosphate added and increase in absorbance at 360 nm in response to the addition to MESG as substrate and the enzyme PNP (EnzChek phosphate assay kit). D) ssDNA dependent ATP hydrolysis of TgRAD51 using EnZChek phosphate assay kit. At 300 μM ATP concentration, with 2 μM TgRAD51 and 60 μM ϕxssDNA, the rate of the reaction as calculated from the slope of the curve is 0.027 μM P_i produced/min (●) where as in absence of ssDNA (■) no ATP hydrolysis occurs. E) Michaelis Menten curve has been plotted with ATP concentrations in the range of 20 μM to 600 μM and kinetic parameters are derived from the curve.
doi:10.1371/journal.pone.0041925.g002

be repaired either by GC using the homologous template (*URA3*) present in the same chromosome 3 kb apart, or it can be repaired by single strand annealing (SSA), which results in deletion of a stretch of intervening DNA sequence. We created two strains SSY1 and SSY2 by transforming the expression vector pTA*ScRAD51* and pTA*TgRAD51* respectively to *NA14Δrad51* strain. In order to assay the efficiency of DNA repair activity of *TgRAD51* we created a single DSB in the *ura3* gene by HO induction and counted the number of cell survival on galactose containing plate and compared to the number of cells on glucose containing plate. Our result shows that the repair efficiency of TgRAD51 (83%) is comparable to that of the wild type (96%), while *NA14Δrad51* being 63% (Figure 3B). In order to monitor the contribution of GC among the survivors, we looked for G418 sulphate resistant cells. If the DSB is repaired by GC, the survivor colonies will be G418 resistant. On the other hand if the break is repaired by SSA, the survivor colonies will be G418 sensitive due to the loss of *KANMX6* cassette. We found that the majority of the survivors (77%) of SSY2 strain (TgRad51) are due to repair by SSA, which does not require the presence of Rad51 protein, gene conversion efficiency being only 23% of the total. On the contrary SSY1 strain (ScRad51) exhibited almost equal preferences for GC and SSA pathways, i.e. approximately 42% of its survival by GC and 58% by SSA (Figure 3B). Hence the GC efficiency of SSY1 (ScRad51) strain is about two fold more than that of SSY2 (TgRad51) strain. The key protein involved in homologous recombination mediated gene conversion is the Rad51 and our result shows that TgRad51 possess inefficient gene conversion activity compared to ScRad51. To rule out the possibility that inefficient gene conversion is not due to poor expression of TgRAD51 in heterologous system, we have done Western blot analysis of SSY1 and SSY2 strains and probed with anti ScRAD51 and PfRAD51 antibodies respectively (anti-PfRad51 antibody cross-reacts with TgRad51 protein). Our data (Figure 3C) shows that TgRad51 expression is comparable to that of ScRad51, actin being the loading control. We conclude that low ATPase activity of TgRad51 is responsible for poor gene conversion efficiency.

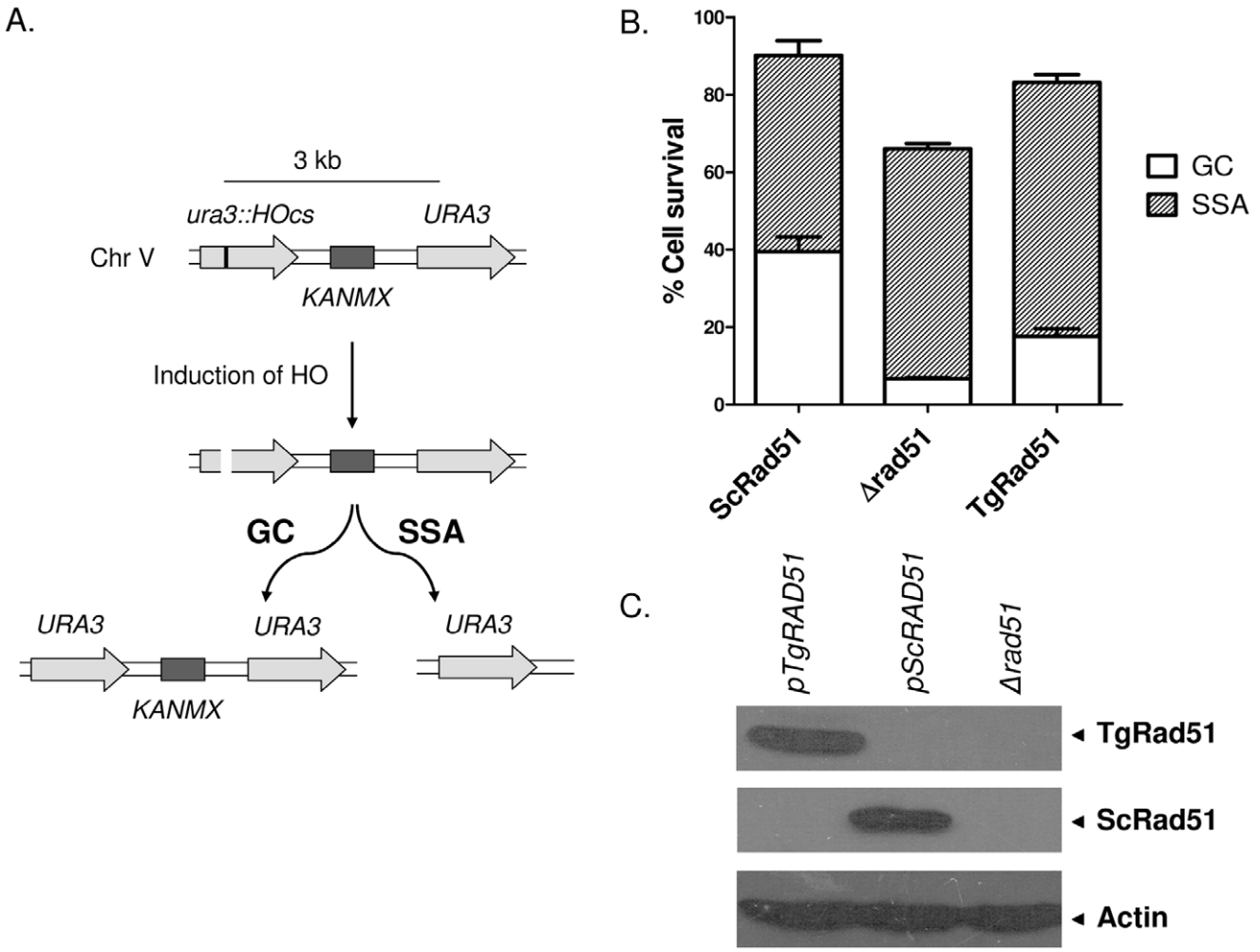

Figure 3. Gene conversion efficiency of *TgRAD51* is poor compared to that of *ScRAD51*. (A) Schematic diagram of DSB repair choice experiment. *URA3* and *KANMX* represent wild type alleles. A HO endonuclease site is incorporated in *ura3::HOcs* mutant allele. Once the HO induced DSB is repaired by gene conversion (GC), the mutant *ura3::HOcs* allele is converted into wild type *URA3* allele. A single strand annealing (SSA) event leads to deletion of the intervening sequence (containing *KANMX* gene) and merging of *ura3::HOcs* and *URA3* alleles to create the wild type *URA3* allele. (B) Bar diagram showing the percentage of cells survived after induction of DSB. The relevant genotypes are marked on the X-axis. The white bars represent fraction of the cells survived by repairing the DSB using GC mechanism, where as the hatched bars denote the fraction of the survivors that employed SSA mechanism. Each bars represent mean value ± SD from 4 different experiments. (C) Western blots showing the abundance of ScRad51 and TgRad51 proteins. Different lanes are marked with the respective genotypes. Actin is the loading control.
doi:10.1371/journal.pone.0041925.g003

Gene Targeting Efficiency of TgRAD51 Increases with Increase in Stretch of Homologous Sequences

A linear piece of DNA bearing homology to a part of the genome can either be integrated at the correct chromosomal locus by means of homologous recombination, or it could be integrated at any random site on the chromosomes via non-homologous recombination. In *T. gondii* targeted integration is extremely less efficient. We wanted to investigate whether such less efficient gene targeting is due to the weak ATP hydrolysis activity of TgRad51, since gene targeting at the correct locus demands search of two homologous needles in the genomic hay stack. In order to appreciate the contribution of TgRad51 protein alone, from other *Toxoplasma* proteins presumably involved in HR pathway, we have used a surrogate yeast model, where the yeast *RAD51* gene is deleted and is replaced by *TgRAD51* gene. Thus, in this system, the main recombination protein (Rad51) is from *T. gondii* and all other auxiliary proteins are from yeast origin. We have engineered this system in such a way that it would lead us to investigate the frequency of targeted vs. random integration choices in the cell harboring TgRAD51. The integration construct contains two selectable markers: *ADE2* and *KANMX6*; while the *ADE2* marker scores for all the integration events (targeted or random), the *KANMX6* marker is able to differentiate between random versus targeted integration. In case of targeted integration at the *ADH4* locus the *KANMX6* marker will be lost as a result of homologous recombination. On the other hand, during NHEJ mediated random integration the *KANMX6* cassette will be retained and such transformants will be G418 resistant. Thus random integration of the cassette would result in $ADE2^{+}$ and $G418^{R}$ colonies (Figure 4Aii), whereas the targeted integration would lead to $ADE2^{+}$ and $G418^{S}$ colonies (Figure 4Ai).

We transformed the KANMX-ADH4-ADE2-ADH4 cassette in three strains: NRY1 (Δ*rad51* containing an empty vector pTA); NRY2 strain (Δ*rad51* containing an expression vector pTA: *ScRAD51*); and MVS26 strain (Δ*rad51* containing an expression vector pTA: *TgRAD51*). The experiments with three

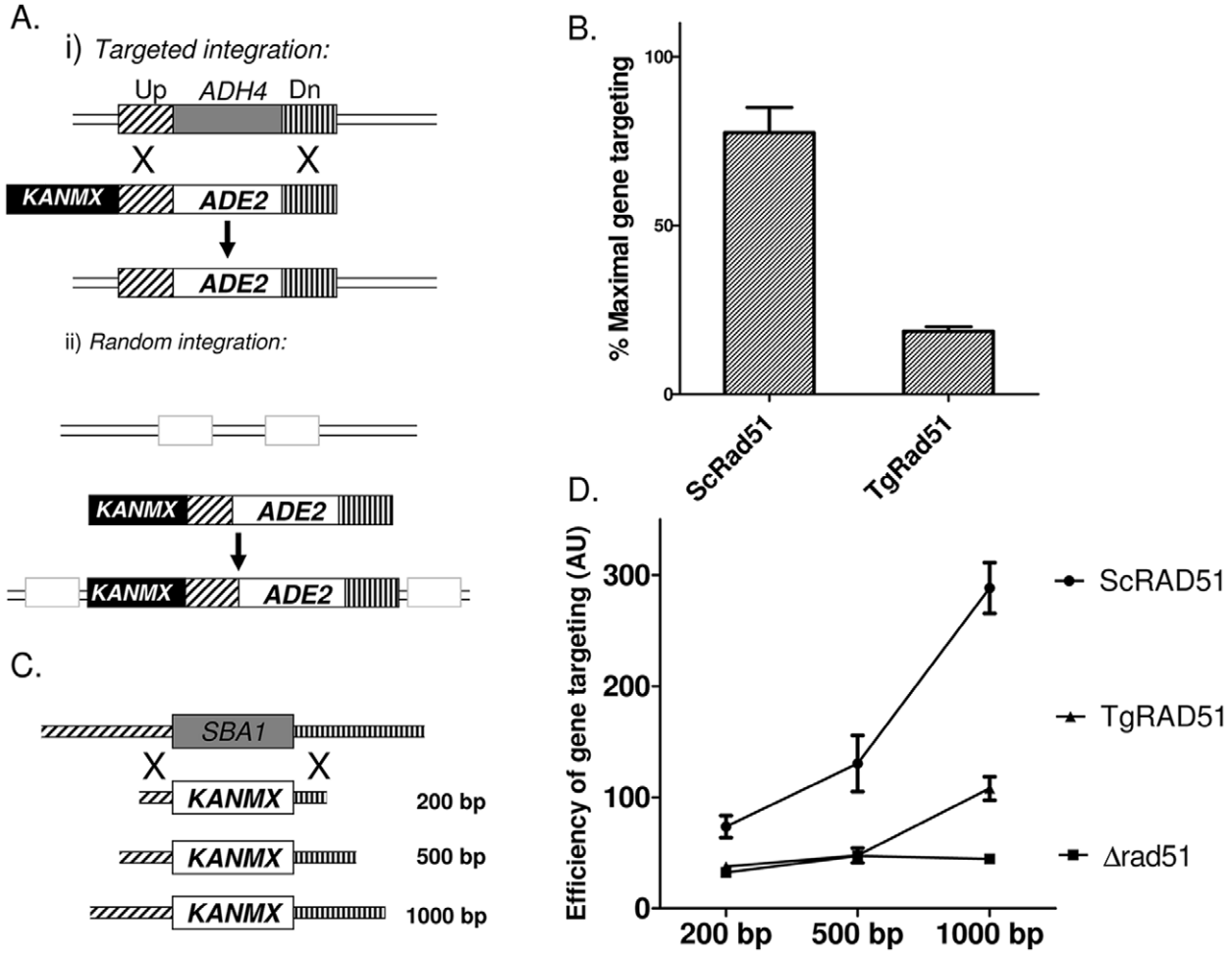

Figure 4. Gene targeting efficiency of TgRAD51 increases with increase in stretch of homologous sequences. (A) Schematic diagram showing molecular events leading to targeted integration of ADE2 gene at ADH4 locus versus random integrations. KANMX is a second selectable marker retained only in case of random integration via non-homologous recombination mechanism. (B) Bar diagram showing efficiency of gene targeting at the ADH4 locus in cells harboring ScRad51 or TgRad51 as the sole recombinase. The mean value from four independent experiments is plotted with standard deviations. (C) Schematic diagram showing knockout strategy for *SBA1* gene. Varying lengths of flanking homologous sequences on either side are indicated. (D) Efficiency of gene knockout with increasing flanking homology. The lengths of the flanking homologous stretches are indicated on the X-axis. These experiments are done at least 3 times and the mean values with standard deviations are plotted.
doi:10.1371/journal.pone.0041925.g004

repeats showed that NRY2 cells (*ScRAD51*) prefer to go for targeted integration at the ADH4 locus with about 85% frequency, where as *TgRAD51* showed a preference for random integration having only 24% frequency for targeted integration (Figure 4B). We hypothesize that poor ATP hydrolysis activity of TgRad51 results in inefficient motor activity of the TgRad51 protein, as a result of which it shows three times reduction in gene targeting efficiency compared to that of *ScRAD51*. A prediction of this hypothesis would be that the increment of flanking homologous sequences would facilitate the homology search and thus would compensate for the weak ATP hydrolysis of Rad51. So, the next question we asked was whether the increase in homologous stretch of DNA has any effect on the frequency of TgRad51 protein mediated targeted integration. To test this hypothesis we designed three constructs to knockout *SBA1* gene. These constructs contains a selectable marker *KANMX6* flanked by upstream and downstream sequences (200 bp/500 bp/1000 bp on either side) of SBA1 gene (Figure 4C). These three constructs were used to disrupt the *SBA1* gene in the cells bearing *ScRAD51* (NRY2) and *TgRAD51* (MVS26). As expected, our result showed that with the increase in homologous sequence there was a gradual increase in the frequencies of the targeted gene disruption (Figure 4D). NRY1 (*Δrad51)* cells showed very less number of targeted disruptions as measured by the survival on G418 sulphate plates throughout varied degree of homology. With increase in flanking homologous sequences, NRY2 cells (*ScRAD51*) showed a steep rise in the gene targeting efficiency (Figure 4D). MVS26 (*TgRAD51*) cells behaved much like *Δrad51* strain up to 500 base pair flanking homology. As the homologous stretches were increased to 1000 base pairs on either side of *KANMX6*, the number of $G418^R$ colonies also increased sharply for MVS26 strain. The increase in homologous flanking region causes same fold increase (approximately 2.5 fold) in the gene targeting efficiency for both *ScRAD51* and *TgRAD51*.

Gene Targeting Efficiency of TgRad51 is Independent of Ku80 Function

In order to investigate whether efficiency of Rad51 mediated gene targeting is affected by the absence of Ku80 protein, we have

compared gene targeting at the *CHL1* locus in presence or in absence of Ku80. To this end we have tested six strains: MVS26 (*TgRAD51*); SSY6 (*TgRAD51* Δ*ku80*); NRY2 (*ScRAD51*); SSY5 (*ScRAD51* Δ*ku80*); NRY1 (Δ*rad51*); and SSY4 (Δ*rad51* Δ*ku80*). To disrupt *CHL1* gene we have generated two constructs. These constructs contain a selectable marker (*HIS3*) flanked by upstream (500 bp/1000 bp) and downstream (500 bp/1000 bp) sequences of *CHL1* ORF (Figure 5A). In accordance to our findings with gene targeting at the *SBA1* locus, we observed that the larger the flanking homology the greater is the gene targeting efficiency at the *CHL1* locus. When compared between the *KU80* proficient and *KU80* deficient cells, we find no significant change in gene targeting efficiency. This observation holds true for both ScRad51 and TgRad51 mediated gene targeting (Figure 5B). Thus, we conclude that gene targeting efficiency is independent of Ku80 function.

Discussion

Several important findings of this study collectively help us in understanding the DSB repair pathways in *T. gondii* as a whole and the underlying reasons for inefficient gene targeting in particular. We have cloned, expressed and purified recombinant TgRad51 and have shown that it possesses ATP hydrolysis activity, which is the lowest among all the eukaryotic Rad51 proteins studied so far. Using yeast as a surrogate system we have characterized *TgRAD51* genetically. To this end we have performed three independent experiments: repair choice experiment, targeted gene knock-in experiment and targeted knock-out experiment. In the repair choice experiment the induced DSB created within *URA3* gene is repaired by any of the two HR mechanisms: gene conversion (GC) or single strand annealing (SSA). While SSA does not depend on Rad51, GC totally depends on Rad51 mediated homology search. Our finding that only GC is compromised in cells harboring

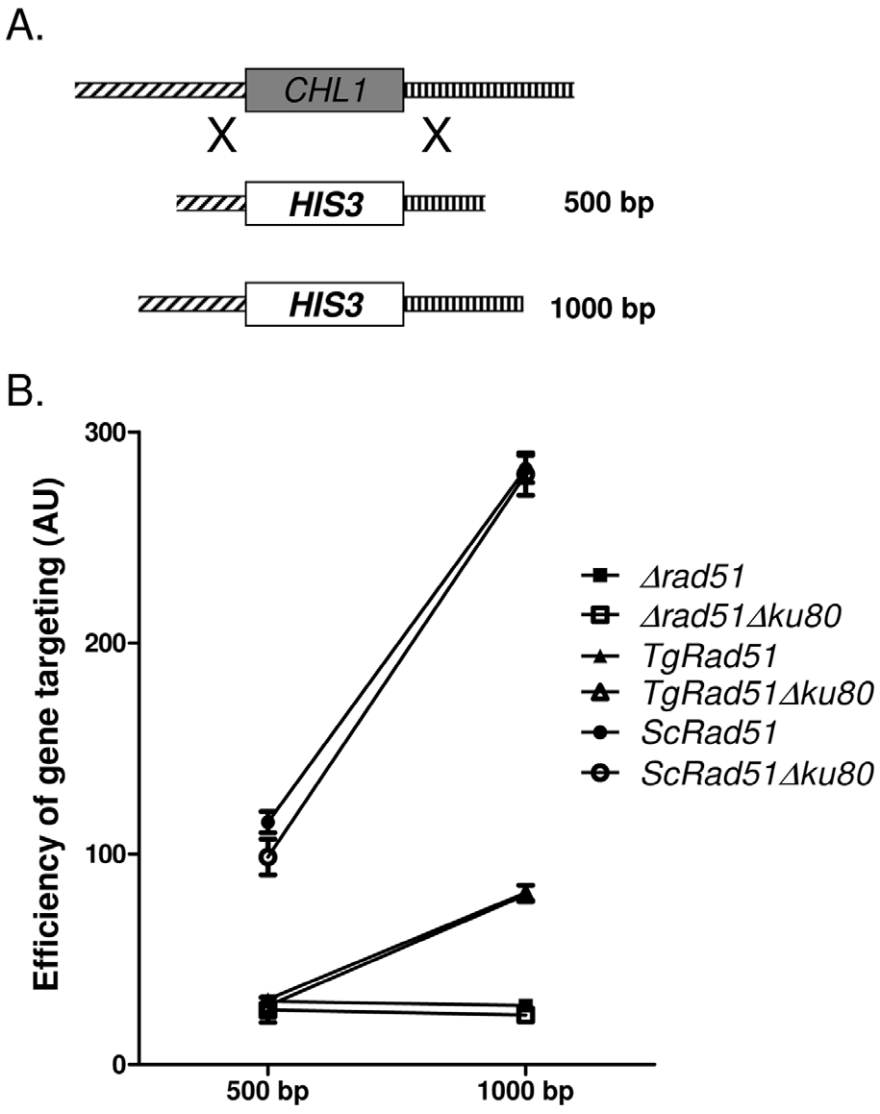

Figure 5. Gene targeting efficiency of TgRad51 is independent of Ku80 function. (A) Schematic diagram showing knockout strategy for *CHL1* gene. Varying lengths of flanking homologous sequences on either side are indicated. (B) Efficiency of gene knockout with increasing flanking homology in *KU80* proficient (closed- circle, triangle or square) and KU80 deficient (open- circle, triangle or square) cells. The lengths of the flanking homologous stretches are indicated on the X-axis. These experiments are done at least 3 times and the mean values with standard deviations are plotted.
doi:10.1371/journal.pone.0041925.g005

TgRad51 is thus consistent. Two other assays (namely gene knock-in and gene knock-out) that also depend on Rad51 mediated homology search are also found to be affected by TgRad51.

Double strand breaks (DSBs) can cause damage to the genomic integrity of a cell. Repair of such DSBs are essential for cell survival. A DSB can be repaired either by homologous recombination (HR) or by non-homologous recombination (NHEJ). Prokaryotes and lower eukaryotes prefer high-fidelity repair mechanism such as HR, where as higher eukaryotes show preference towards mutagenic NHEJ pathway. An attractive hypothesis links the repair choice to gene density. Due to the low gene density in higher mammals mutations acquired during NHEJ mediated repair of random DSBs, are less likely to affect gene function, where as in organisms with high gene density, the likelihood of having random DSBs within the genes are higher and concomitantly repair of such breaks by NHEJ might affect gene function. For example there are about 7 genes per mega base of mouse or human genome, thus the probability of having random breaks in the intergenic regions are higher than within the genes. On the other hand prokaryotes and lower eukaryotes have very high gene density. The gene densities (number of genes per mega base) of *E. coli*, *S. cerevisiae*, *S. pombe*, *Leishmania major* and *Plasmodium falciparum* are 952, 465, 396, 297 and 231 respectively. Thus it is not surprising that all the aforementioned organisms prefer HR over NHEJ. The recombinase RecA (Rad51 ortholog) from most of these organisms have robust ATP hydrolysis activity (the k_{cat} values for EcRecA, ScRad51, and LmRad51 are 18 min^{-1}, 2.9 min^{-1}, 0.7 min^{-1}) [24,25,26] compared to that of mammals (k_{cat} of hRad51 is 0.16 min^{-1}) [27]. Thus, a strong ATPase activity of Rad51 is an indicative of a robust HR system. This notion is supported by studies with mutant hRad51 (K133R) incapable of ATP hydrolysis, where gene targeting is severely compromised. However, this mutant does not have any effect on strand exchange between homologous templates. Moreover, the DNA repair ability of the mutant cell is not affected [28]. *T. gondii* despite being a lower eukaryote shows preference to NHEJ pathway. The low gene density of *Toxoplasma* genome (100 genes per mega base) and our finding that TgRad51 too have a weak ATP hydrolysis activity (k_{cat} is 0.034 min^{-1}) fits very well with the choice of DSB repair pathway in this organism. Although, it appears that HR might not be a general mechanism for repair of DSB in *T. gondii*, whether HR activity is required to repair specific types of DSB ends remains as an open question. It is likely that the primary function of HR in *T. gondii* could be during sexual reproduction.

TgRad51 is the first member of the recombination machinery of *T. gondii* to be characterized. The existence of putative orthologues of Dmc1, Rad54, Rad50 and Mre11 in *Toxoplasma* genome suggests that this parasite does possess a functional recombinosome. Interestingly, the apparent lack of Rad52 orthologue in *T. gondii*, *Plasmodium falciparum* and *Cryptosporidium parvum* is suggestive of a Rad52 independent recombination mechanism in these apicomplexan parasites.

Gene disruption and epitope tagging are the two most important tools to study gene function in the post genomics era. Both of these techniques rely on homologous recombination mechanism. Since the HR machinery is weak in *T. gondii*, targeted gene disruption or tagging of endogenous genes are very less efficient in this parasite. Works from Bzik laboratory and Carruthers laboratory have demonstrated enhanced gene targeting in *ku80* null *T. gondii* [17,29]. However, as expected there is no change in gene targeting efficiency in *ku80* null parasite lines [17]. In our assay we used yeast as surrogate organism to investigate whether the absence of Ku80 protein has any positive effect on the efficiency of gene targeting. Since it is possible to obtain individual clonal transformants on solid medium using yeast system, it allowed us to determine the locus of integration (targeted versus random) in each of the individual colonies. We found that the presence or absence of Ku80 protein does not have any significant effect on gene targeting efficiency per se. This is because, the absence of Ku80 abrogates NHEJ pathway but does not increase the efficiency of Rad51 recombinase. An organism that harbors these two competing pathways (HR and NHEJ) may results in either random integration of the transfected DNA *via* NHEJ mechanism or integration at the right locus *via* HR mechanism. Only one of the two types of recombinant clones (where integration took place at the right locus) is desirable. Abrogating NHEJ pathway (by knocking out *KU80* or other molecular players of NHEJ) would definitely minimize the number of those recombinant clones where random integration has taken place and thereby facilitate the screening of the desired clone. However, it would not increase the integration efficiency at the correct locus. This notion is supported by the finding that targeted repair of Δ*hxgprt* became independent of TgKu80 when enough flanking homology (910 bp) was provided [17]. Genome wide homology search by ssDNA bound Rad51 is the rate determining step in homology directed recombination. Thus, longer homologous stretches are likely to have great positive impact on such process. Our finding that increased flanking homology facilitate targeted gene knock out by TgRad51 is in corroboration with this hypothesis. Similarly, since ATP hydrolysis activity of Rad51 is required for such homology search a higher ATPase activity also results in more efficient gene targeting. It will be interesting to see whether a transgenic *T. gondii* harboring Rad51 from *L. major* or *S. cerevisiae* could significantly increase the gene targeting efficiency in this protozoan parasite. Future studies might test this possibility either by expressing LmRad51/ScRad51 (having robust ATPase activity) in *T. gondii* from an episomal plasmid or by delivering purified Rad51 proteins *via* nano particle mediated delivery systems. Positive results from such studies will have a great impact on functional studies of Toxoplasma biology.

Supporting Information

Figure S1 Phylogenetic analysis of eukaryotic Rad51 proteins using Clustal method (Meg align, DNA star). T. gondii Rad51 is highlighted.
(TIF)

Figure S2 Amino acid sequence of TgRad51 protein from RH strain. The underlined peptide sequences were generated from MS-MS analysis.
(TIF)

Acknowledgments

We thank Professor Martin Kupiec, Tel Aviv University for providing the strains NA14 and NA14Δrad51 and Professor Vern Curruther, University of Michigan, Ann Arbor, USA, for providing *T. gondii* cDNA library.

Author Contributions

Conceived and designed the experiments: MKB. Performed the experiments: SSA SMV SB. Analyzed the data: MKB SB. Contributed reagents/materials/analysis tools: MKB SB. Wrote the paper: MKB SB.

References

1. Mimitou EP, Symington LS (2009) DNA end resection: many nucleases make light work. DNA Repair 8(9): 983–95.
2. Haber JE (2000) Partners and pathways repairing a double strand break. Trends in Genetics 16(6): 259–264.
3. Jain S, Sugawara N, Lydeard J, Vaze M, Tanguy Le Gac N, et al. (2009) A recombination execution checkpoint regulates the choice of homologous recombination pathway during DNA double-strand break repair. Genes Dev 23(3): 291–303.
4. Agmon N, Pur S, Liefshitz B, Kupiec M (2009) Analysis of repair mechanism choice during homologous recombination. Nucleic Acids Res 37(15): 5081–92.
5. Symington LS (2002) Role of Rad52 epistatis group genes in homologous recombination and double strand break repair. Microbiol Mol Biol Rev 66: 630–670.
6. Paques F, Haber JE (1999) Multiple pathways of recombination induced by double strand break in Saccharomyces cerevisiae. Microbiol Mol Biol Rev 63: 349–404.
7. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double strand break repair model for recombination. Cell 33(1): 25–35.
8. Ogawa T, Yu X, Shinohara A, Egelman EH (1993) Similarity of the yeast RAD51 filament to the bacterial RecA. Science 259(5103): 1896–9.
9. Jiang H, Xie Y, Houston P, Stemke-Hale K, Mortensen UH, et al. (1996) Direct association between the yeast Rad51 and Rad54 recombination proteins. J Biol Sci 271: 33181–33186.
10. Sung P (1994) Catalysis of ATP dependent homologous DNA pairing and strand exchange by yeast Rad51 protein. Science 265: 1241–1243.
11. Hays SL, Firmenich AA, Berg P (1995) Complex formation in yeast double strand break repair: participation of Rad51, Rad52, Rad55 and Rad57 proteins. Proc National Sci Academy 92(15): 6925–6929.
12. Donald RG, Roos DS (1998) Gene knock-outs and allelic replacements in Toxoplasma gondii: HXGPRT as a selectable marker for hit-and-run mutagenesis. Mol Biochem Parasitol 91(2): 295–305.
13. Janzen CJ, Lander F, Dreesen O, Cross GA (2004) Telomere length regulation and transcriptional silencing in KU80-deficient Trypanosoma brucei, Nucleic Acids Res 32: 6575–6584.
14. Burton P, McBride DJ, Wilkes JM, Barry JD, McCulloch R (2007) Ku heterodimer-independent end joining in Trypanosoma brucei cell extracts relies upon sequence micro homology. Eukaryot Cell 61773–1781.
15. Glover L, McCulloch R, Horn D (2008) Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. Nucleic Acids Res 362608–2618.
16. Walker JR, Corpina RA, Goldberg J (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 412607–614.
17. Fox BA, Ristuccia JG, Gigley JP, Bzik DJ (2009) Efficient gene replacements in Toxoplasma gondii strains deficient for non homologous end joining. Eukaryotic Cell 8(4): 520–529.
18. Bhattacharyya MK, Kumar N (2003) Identification and molecular characterization of DNA damaging agent induced expression of Plasmodium falciparum recombination protein PfRad51. Int J Parasitol 33(12): 1385–92.
19. Bhattacharyya MK, Norris DE, Kumar N (2004) Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation. Infect Genet Evol 4(2): 91–8.
20. Bhattacharyya MK, Bhattacharyya nee Deb S, Jayabalasingham B, Kumar N (2005) Characterization of kinetics of DNA strand-exchange and ATP hydrolysis activities of recombinant PfRad51, a Plasmodium falciparum recombinase. Mol Biochem Parasitol 139(1): 33–9.
21. Laskar S, Bhattacharyya MK, Shankar R, Bhattacharyya S (2011) Hsp90 controls SIR2 mediated gene silencing. PloS One 6(8): e23406.
22. Longtine MS, McKenzie A 3rd, Demarini DJ, Shah NG, Wach A, et al. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast 14(10): 953–61.
23. Diede SJ, Gottschling DE (1999) Telomerase-mediated telomere addition in vivo requires DNA primase and DNA polymerases alpha and delta. Cell 99(7): 723–33.
24. Roca AI, Cox MM (1990) The RecA protein structure and function. Crit Rev Biochem Mol Biol 25: 415–456.
25. Bianco PR, Tracy RB, Kowalczykowski SC (1998) DNA strand exchange proteins: a biochemical and physical comparison. Front Biosci 3: D570–603.
26. McKean PG, Keen JK, Smith DF, Benson FE (2001) Identification and characterisation of a RAD51 gene from Leishmania major. Mol Biochem Parasitiol 115: 209–16.
27. Baumann P, Benson FE, West SC (1996) Human Rad51 Protein Promotes ATP-Dependent Homologous Pairing and Strand Transfer Reactions In Vitro. Cell 87(4): 757–66.
28. Morrison C, Shinohara A, Sonoda E, Yamaguchi-Iwai Y, Takata M, et al. (1999) The essential functions of human Rad51 are independent of ATP hydrolysis. Mol Cell Biol 19(10): 6891–7.
29. Huynh MH, Carruthers VB (2009) Tagging of endogenous genes in a Toxoplasma gondii strain lacking Ku80. Eukaryot Cell 8(4): 530–9.