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 His6 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 \textit{ScRAD51} gene in pTA plasmid. A). Strategy to clone \textit{ScRAD51} in pTA background at \textit{BamHI} and \textit{PstI} sites. B). The pCR 2.1 TOPO:\textit{ScRAD51} construct was confirmed by \textit{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 \textit{PstI} and \textit{EcoRI}. Fragments of desired size 0.6, 0.7 and 3.9 Kb in lane 2 proved that the construct has \textit{ScRAD51} insert. C). Confirmation for pTA:\textit{ScRAD51} clone: The plasmid sample was digested with \textit{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 \textit{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 \textit{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 \textit{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 \textit{ku80} null \textit{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 \textit{Δhxgprt} 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 \textit{T. gondii}, wherein TgRad51 can be replaced by Rad51 of other organisms like \textit{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 facilitate genetic dissection of protozoan parasite biology which can lead to new treatments for significant parasitic diseases.
REFERENCES


APPENDIX -3

A. 3. 1 Published research paper: