Details of the research work duly signed by the applicant, for which the Sun Pharma Science Foundation Research Award is claimed, including references and illustrations (not to exceed 6000 words).

Replicative and Post-Replicative Processes to maintain genomic Integrity

The research work for which the award is claimed involves studies:

A. that provide deep mechanistic insight regarding enzymes involved in replicative and post-replicative processes to maintain genomic integrity in bacteria.
B. conducted to identify novel inhibitors of enzymes involved in synthesis or proofreading of the genome in viruses.

A. DNA synthesis by DNA-dependent-DNA polymerases

DNA-dependent DNA polymerases (dPols) are the primary enzymes responsible for duplication of the genome. In addition, erroneous incorporation of incorrect nucleotides by dPols gives rise to mutations that may aid evolution. We study different dPols from various organisms to understand the chemical mechanism utilized to achieve their role in replication and/or evolution. Also, many organisms possess specialized DNA polymerases that rescue replication stalled at damaged nucleotides (DNA lesions) and we aim to unearth the structural mechanism utilized by these enzymes to achieve translesion DNA synthesis.

1. Chemical Mechanism of DNA synthesis by DNA polymerases

In all living organisms, deoxyribonucleic acid (DNA) is synthesized by DNA polymerases and these enzymes catalyze template-directed synthesis of DNA. DNA polymerases employ semiconservative mode of replication using primer-template duplex DNA and deoxynucleotide triphosphates (dNTPs) as precursors for DNA synthesis. The primer provides a 3'-hydroxyl group that can be extended by the polymerase and the identity of the incoming dNTP is determined by the template residues. Mg^{2+} ions also play an important role in the polymerization reaction. DNA polymerases extend the primer in the 5'-3' direction. The formation of a phosphodiester bond between the α-phosphate of the incoming dNTP and the 3'- hydroxyl group of the terminal primer nucleotide is the primary chemical reaction catalyzed by the DNA polymerase enzyme.

We have conducted time-resolved crystallography on DNA polymerase IV (PolIV) from *Escherichia coli* to elucidate the steps involved in the formation of a phosphodiester bond during DNA synthesis by a DNA polymerase. Crystals of PolIV in complex with DNA and incoming

dTTP were grown and incubated with Mg^{2+} for different time points before freezing. Data was collected from the frozen, the structures were determined and the electron density maps were examined. The maps clearly showed the formation of the phosphodiester bond between the 3' –OH and the α-phosphate of the incoming dTTP and the breakage of the bond between the α-phosphate. The time-resolved crystallography experiment showed that DNA synthesis by DNA polymerases follows a stepwise dissociative SN_2 reaction. Recent studies on DNA polymerases had suggested that DNA synthesis by DNA polymerases involves not two but three Mg2+ ions. However our analysis suggested that this might be an artefact of excessive Mg2+ ion concentration and that two octahedrally co-ordinated Mg2+ ions are sufficient to drive the DNA synthesis reaction.

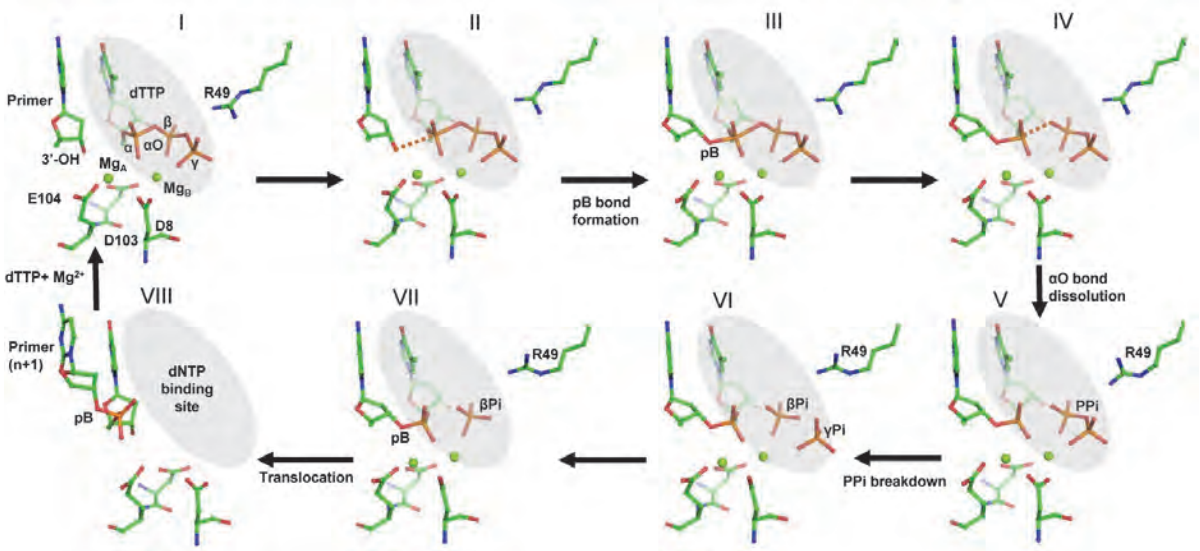


Figure 1. ***Sequence of events during incorporation of dTTP opposite dA by DNA polymerase IV****. The different steps associated with synthesis of a phosphodiester bond are displayed. The terminal primer nucleotide (dC), the incoming nucleotide (dTTP) and active site residues are shown in stick representation and colored according to element. The* Mg^{2+} *ions are shown in the form of green spheres.*

More importantly out study showed that hydrolysis of the PPi byproduct is an intrinsic and critical step of the DNA synthesis reaction. Biochemical assays with appropriately modified dNTPs suggest that this step may be conserved in different DNA polymerases, including the RNA-dependent-DNA polymerase responsible for duplication of the RNA genome of retroviruses such as HIV.

These studies dispel a long-standing belief that DNA synthesis is an example of a coupled reaction wherein the byproduct PPi moiety is hydrolysed by accompanying pyrophosphatase enzymes to render a large negative free energy to the overall reaction. Our study shows that the hydrolysis of PPi occurs after the formation of the phosphodiester bond and ensures that the DNA synthesis reaction is energetically favorable without the need for additional enzymes (Figure 1). Overall, the study brings to light the mechanism of the fundamental reaction responsible for genome duplication and the insight obtained from this study may aid the development of improved PCR-based diagnostic kits and novel therapeutic strategies against retroviruses. The manuscript resulting from the study was accorded breakthrough status by the journal Nucleic Acids Research and the editor wrote a commentary highlighting the importance of this work.
Publication: Kottur, J., and **Nair, D. T.** (2018) Pyrophosphate hydrolysis is an intrinsic and critical step of the DNA synthesis reaction. ***Nucleic Acids Res.*** 46:5875-5885.

2. *A polar filter in DNA polymerases prevents ribonucleotide incorporation*

Ribonucleotide incorporation during DNA synthesis by DNA polymerases leads to genomic instability and cellular lethality. To prevent the addition of rNTPs, the majority of DNA polymerases possess a steric filter in the form of a bulky aromatic residue that stacks against the ribose sugar. MsDpo4 from *Mycobacterium smegmatis* naturally lacks this steric filter and hence can incorporate ribonucleotides. In comparison to MsDpo4, DNA polymerase IV from *Escherichia coli* (PolIV) exhibits stringent selection of deoxyribonucleotides during DNA synthesis. We observed that engineering the bulky aromatic residue in MsDpo4 did not completely abrogate its ability to incorporate ribonucleotides. A rigorous comparison of MsDpo4 and PolIV using structural, biochemical tools and growth assays led to the discovery of an additional polar filter utilized by DNA polymerases to prevent ribonucleotide incorporation (Figure 1). In the case of PolIV, this polar filer is represented by Thr43 and it forms interactions with the incoming nucleotide. The equivalent residue in MsDpo4 is Cys47. The mutation C47T in MsDpo4 led to a significant decrease in the ability of MsDpo4 to incorporate ribonucleotides. The protein bearing the bulky aromatic residue and the C47T mutation did not show any ribonucleotide incorporation for all possible Watson-Crick base pairs. Conversely, the presence of the T43C mutation in PolIV led to an increase in ribonucleotide incorporation.

Based on our studies it is clear that the interaction of the incoming nucleotide with the polar filter draws the incoming nucleotide closer to the surface of the fingers subdomain. As a result, the 2' –OH on ribonucleotide triphosphates will sterically clash with the surface of the enzyme and will not be able to bind to the enzyme in a conformation capable of productive catalysis. As a result, incorporation of rNTPs will be minimal. Overall our studies show that the polar and steric filter acts in concert to avert adventitious rNTP addition by DNA polymerase. (Johnson *et al.*, 2019, Nucleic Acids Res., 47:10693). An analysis of sequences of human DNA polymerases in different databases such as NIH-NCI GDC data portal, NCBI ClinVar and BIH-DB SNP Short Genetic Variation, revealed a correlation between mutations at the polar filter site and disease in human dPols δ and ζ.

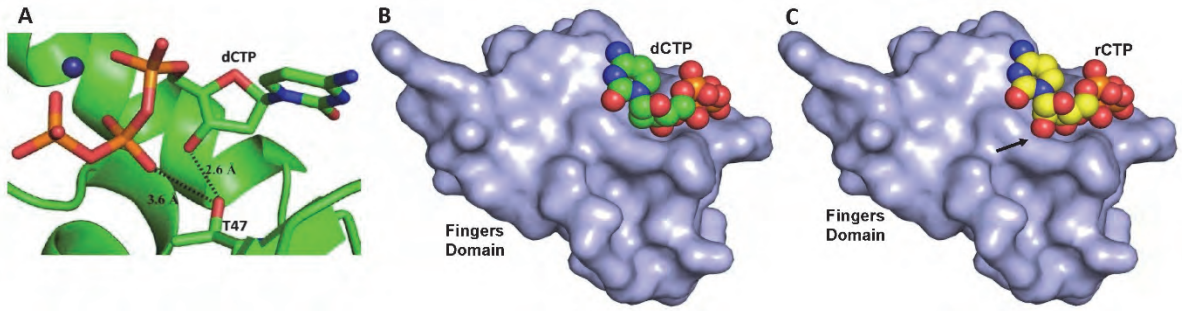


Figure 2. ***A polar filter aids sugar selectivity.*** *(A) The crystal structure of the MsDpo4-C47T$_{DNA(dG):dCTP}$ complex shows that the 3' -OH and the β-phosphate of the incoming nucleotide form interactions with the engineered polar filter (T47). (B) The surface of the fingers domain from this structure along with the space-filling representation of the incoming dCTP is displayed and shows that the dCTP is accommodated in the catalytic site cavity without any steric clashes. (C) The surface of the fingers domain from this structure along with the space-filling representation of incoming rCTP modelled in the catalytic site is displayed. The 2'-OH of the modelled rCTP forms steric clashes with the surface of the fingers domain (highlighted by an arrow). The engineered residue (Thr47) therefore represents a polar filter that aids sugar selectivity.*

Publication: Johnson, M. K., Kottur, J., and **Nair, D. T.** (2019) A polar filter in DNA polymerases prevents ribonucleotide incorporation. ***Nucleic Acids Res.*** 47:10693-10705.

3. The activity of the PfPrex from Plasmodium falciparum vis-à-vis oxidized dNTPs

The DNA polymerase module of the Pfprex enzyme (PfpPol) is responsible for duplication of the genome of the apicoplast organelle in the malaria parasite. We show that PfpPol can misincorporate oxidized nucleotides such as 8oxodGTP opposite dA. This event gives rise to transversion mutations that are known to lead to adverse physiological outcomes. The apicoplast

genome is particularly vulnerable to the harmful effects of 8oxodGTP due to very high AT content (~ 87%). We observe that the proofreading activity of PfpPol has the unique ability to remove the oxidized nucleotide from the primer terminus (Figure 1). Due to this property, the proofreading domain of PfpPol is able to prevent mutagenesis of the AT-rich apicoplast genome and neutralize the deleterious genotoxic effects of ROS generated in the apicoplast due to normal metabolic processes (Sharma et al, 2020, Sci. Rep. 10:11157). The proofreading activity of the Pfprex enzyme may, therefore, represent an attractive target for therapeutic intervention. Also, a survey of DNA repair pathways shows that the observed property of Pfprex constitutes a novel form of dynamic error correction wherein the repair of promutagenic damaged nucleotides is concomitant with DNA replication.

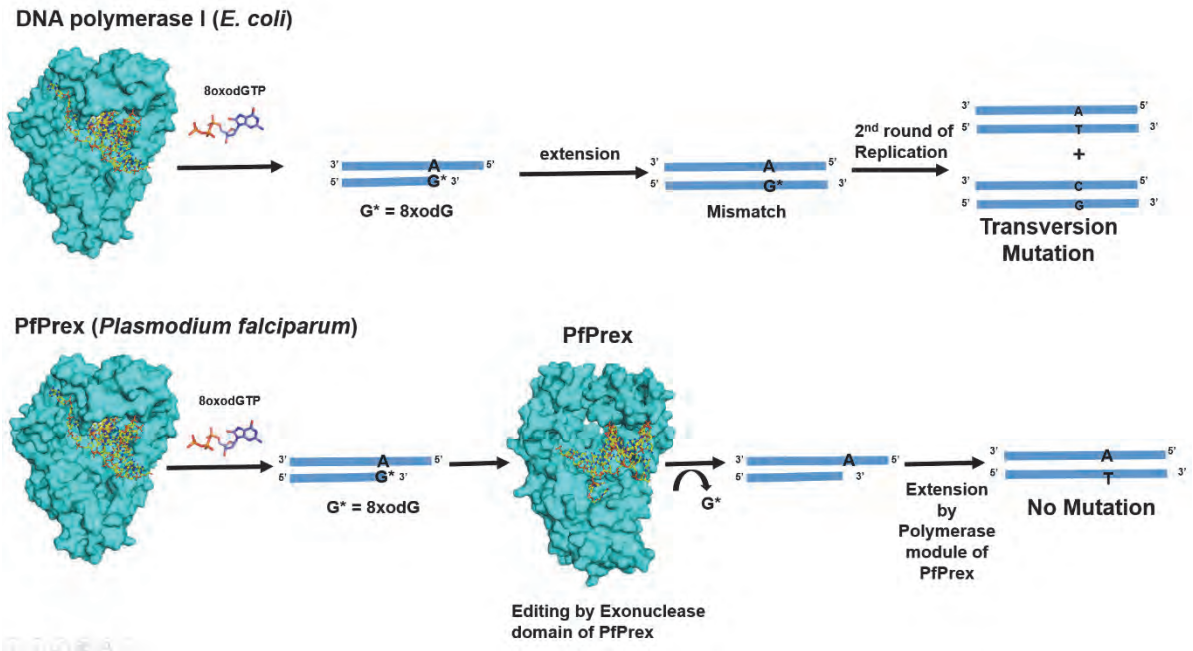


Figure 2. ***Proofreading exonuclease of Pfprex from Plasmodium falciparum ca remove misincorporated oxidized dNTPs.*** *The incorporation of oxidized nucleotide 8oxodGTP in the genome by DNA polymerase I (E. coli) will lead to G:A mismatches leading to transversion mutations. However, the 3' to 5' exonuclease of Pfprex has the unique ability to excise out misincorporated 8xoxdGTP and thus prevent mutagenesis of the AT rich circular genome present in the apicoplast organelle*

Publication: Sharma, M., Narayanan, N. and **Nair, D. T.** (2020) The proofreading activity of Pfprex from Plasmodium falciparum can prevent mutagenesis of the apicoplast genome by oxidized nucleotides. **Sci. Rep.** 10:11157. (Impact Factor: 4.379)

4. Role of the proofreading activity of PolX from Staphylococcus aureus in replication fidelity

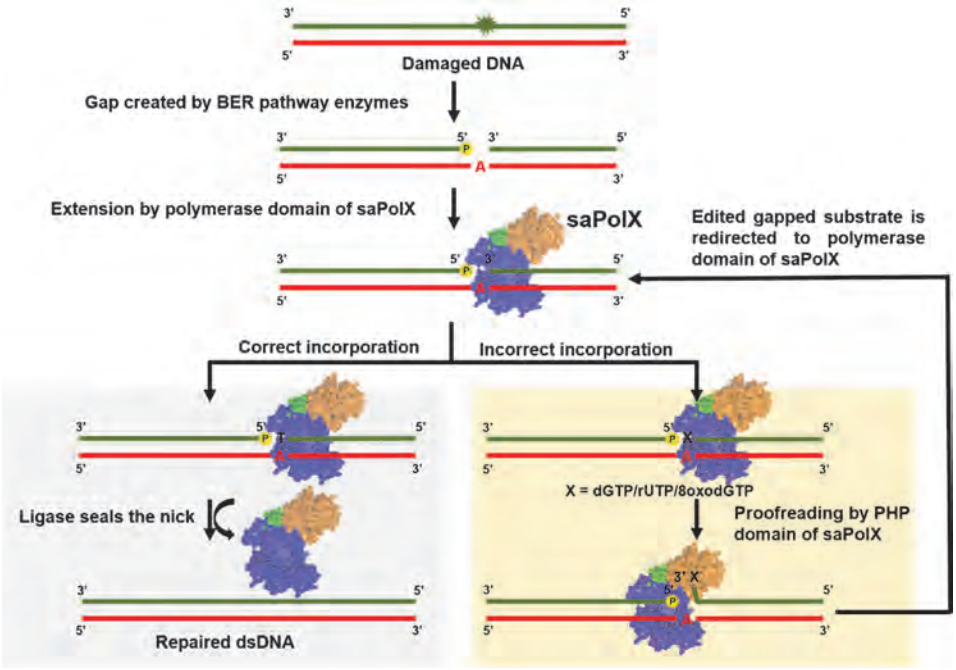


Figure 1: ***Model for the high-fidelity DNA synthesis by saPolX.*** *The core polymerase domain, PHP domain and linker connecting these two domains are displayed in blue, orange and green colors, respectively. The schematic shows the interplay between the 5'-3' polymerization and 3'-5' exonuclease activities residing in polymerase and PHP domains, respectively. saPolX incorporates either correctly paired nucleotide or mismatched nucleotide on gapped DNA substrate generated during BER pathway. The correctly filled DNA substrate is sealed by the ligase enzyme in the subsequent step. The DNA substrates with misincorporated deoxyribo-, ribo- and oxidized nucleotides are directed towards the PHP domain. The PHP domain removes these bases and redirects the gapped DNA substrate to the polymerase domain to repair it correctly.*

The X family is one of the eight families of DNA polymerases (dPols) and members of this family are known to participate in the later stages of Base Excision Repair. Many prokaryotic members of this family possess a Polymerase and Histidinol Phosphatase (PHP) domain at their C-termini. The PHP domain has been shown to possess 3'-5' exonuclease activity and may represent the proofreading function in these dPols. PolX from *Staphylococcus aureus* also possesses the PHP domain at the C-terminus, and we have shown that this domain has an intrinsic Mn^{2+} dependent 3'-5' exonuclease capable of removing misincorporated dNMPs from the primer. The misincorporation of oxidized nucleotides such as 8oxodGTP and rNTPs are known to be pro-mutagenic and can lead to genomic instability. Here, we show that the PHP domain aids DNA replication by the removal of misincorporated oxidized nucleotides and rNMPs (Nagpal et al, 2020, Sci. Rep. 11:4178). Overall, our study shows that the proofreading activity of the PHP

domain plays a critical role in maintaining genomic integrity and stability (Figure 1). The exonuclease activity of this enzyme can, therefore, be the target of therapeutic intervention to combat infection by methicillin-resistant-*Staphylococcus-aureus*.

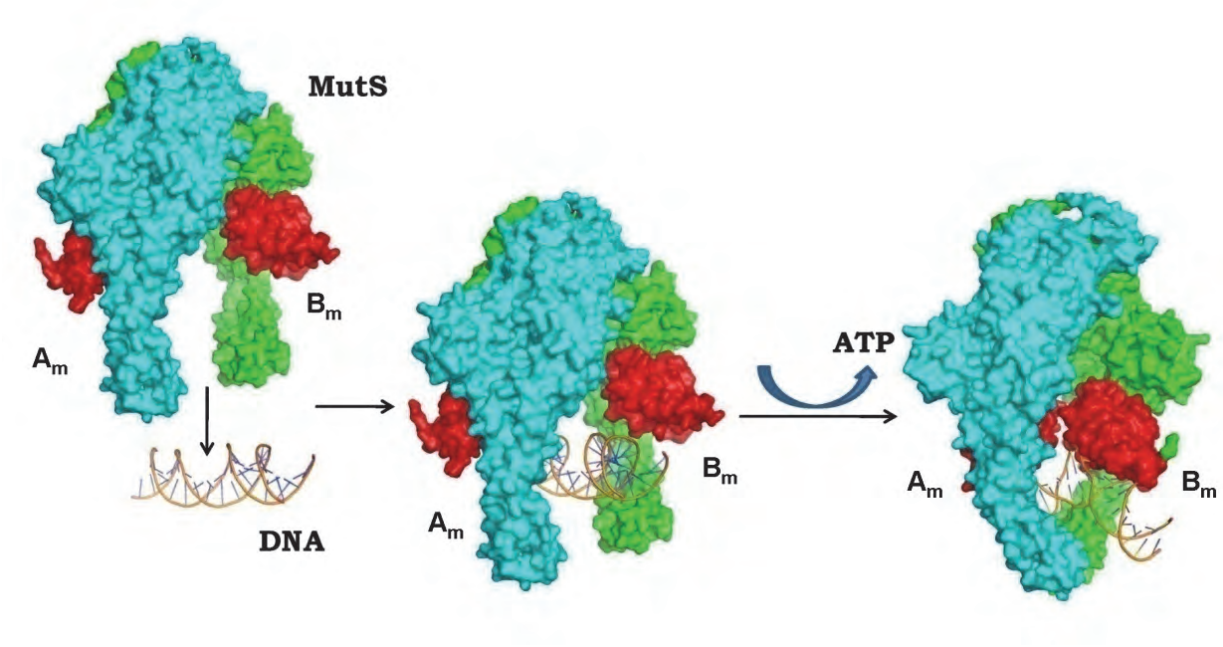
Publication: Nagpal, S., and **Nair, D. T.** (2021) The PHP domain of PolX from Staphylococcus aureus aids high fidelity DNA synthesis through the removal of misincorporated deoxyribo-, ribo- and oxidized nucleotides. ***Sci Rep.*** 11:4178. (Impact Factor: 4.379)

B. DNA Mismatch Repair

The Mismatch Repair (MMR) Pathway serves to maintain genomic integrity by correcting errors that appear during replication. Although MMR in *E. coli* is reasonably well-characterized, it is known that the majority of bacteria and all eukaryotes do not follow a similar pathway. Using MMR in *Neisseria gonorrhoeae (Ngo)* as a model system, we aim to elucidate the mechanism of MMR in organisms that do not follow the *E. coli* paradigm. In *Ngo*, the specific proteins associated with MMR are represented by NgoS (ortholog of MutS) and NgoL (ortholog of MutL). Previoulsy. We had determined the structure of the MutL-CTD dimer and shown that it forms an inverted dimer[29]. This study also helped correct the erroneous dimer interface of MutL-CTD that was present in literature[29].

MutS represents the primary mismatch sensor and forms a dimer clamp that encircles DNA and bends it to scan for mismatches. Two monomers of MutS associate to form an oval disc-shaped asymmetric dimer with a central channel into which DNA is loaded. MutS shows the presence of four regions, the N-terminal domain (NTD), the central domain, clamp region and the C-terminal domain (CTD). The mechanism by which the MutS dimer encircles DNA was not known, and the origin of force required to bend DNA was unclear.

We show that in the absence of DNA and presence of ADP or AMPPNP, NgoS forms a symmetric dimer wherein the two monomers are associated only through the CTD, and there is no interaction between the clamp regions. Consequently, a large gap exists between the clamp regions through which DNA can enter the central channel. The mismatch scanning monomer (B_m) then moves by nearly 50 Å to associate with the other monomer (A_m) so that the clamp regions come in contact with each other and the dimer encircles DNA. Due to the movement of B_m, the N-terminal domains of both monomers press onto DNA to bend it. The mechanism of toroid formation evinces that the force required to bend DNA arises primarily due to the movement of the monomer B_m and hence, the MutS dimer acts like a pair of pliers to bend DNA (Figure 2).



Mechanism of assembly of the MutS-DNA complex: *The different stages in the assembly of MutS-DNA complex are displayed in. The two monomers are shown in surface representation, labeled A_m and B_m and are colored cyan and green, respectively. The N-terminal domain of each monomer is coloured red.*

Also, our study shows that the ATP binding and not hydrolysis is critical for formation of the MutS-DNA complex and that the ATP molecule bound to A_m is expelled on DNA binding. Overall, this study provides mechanistic insight regarding the primary event in DNA mismatch repair i.e. the assembly of the MutS-DNA complex. The insight gained from this study can be exploited to develop small molecule inhibitors of the MMR pathway. It is predicted that inhibition of MMR will ultimately increase the frequency at which deleterious mutations appear in the genome and attenuate the ability to bacterial pathogens to proliferate and cause disease.

Publication: Nirwal, S., Kulkarni, D. S., Sharma, A., Rao, D. N. and **Nair, D. T**. (2018) Mechanism of formation of a toroid around DNA by the Mismatch Sensor protein. ***Nucleic Acids Res.*** 46:256-266. (Impact Factor: 16.971)

C. Viral Genome Replication

We aim to provide deep mechanistic insight regarding the replication of the genomes of viruses such as Japanese Encephalitis Virus, Chikungunya and SARS-CoV-2. We will exploit the information derived from these studies to develop novel therapeutic strategies. Towards this end, we have previously characterized a novel pre-initiation state that is critical for accurate initiation

of genome replication in Flaviviruses (Surana et al, 2014, Nucleic Acids Res., 42:2758). Recently, we have composed a comprehensive review of the antivirals available that target the RNA-dependent-RNA polymerase (RdRP) enzyme present in viruses that possess single stranded RNA genomes. These viruses are classified in Group IV according to the Baltimore system of virus classification and many disease causing viruses such as Dengue virus, Japanese Encephalitis virus, Poliovirus etc. are part of this Group. SARS-CoV-2 which is responsible for the ongoing COVID19 pandemic is part of the Group IV of viruses.

SARS-CoV-2 belongs to the Coronaviridae family. Like other members of this family, the virus possesses a positive-sense single-stranded RNA genome. The genome encodes for the nsp12 protein, which houses the RNA-dependent-RNA polymerase (RdRP) activity responsible for the replication of the viral genome. Using computational tools, we predict that vitamin B12 (methylcobalamin) may bind to the active site of the nsp12 protein (Figure 3). A model of the nsp12 in complex with substrate RNA and incoming NTP showed that Vitamin B12 binding site overlaps with that of the incoming nucleotide. A comparison of the calculated energies of binding suggested that the vitamin may bind to the active site of nsp12 with affinity comparable to that of the natural substrates. It is, therefore, possible that methylcobalamin binding may prevent association of nsp12 with RNA and NTP and thus inhibit the RdRP activity. Overall, our computational studies suggest that methylcobalamin form of vitamin B12 may serve as an effective inhibitor of the nsp12 protein (Narayanan and Nair, 2020, IUBMB Life, in press). The effect of methylcobalamin on replication of live SARS-CoV-2 virus in cells is currently under testing and preliminary results are encouraging. The efforts to purify the nsp12 protein are also ongoing and the effect of methylcobalamin and other potential inhibitors on RNA synthesis and binding will be tested *in vitro*.

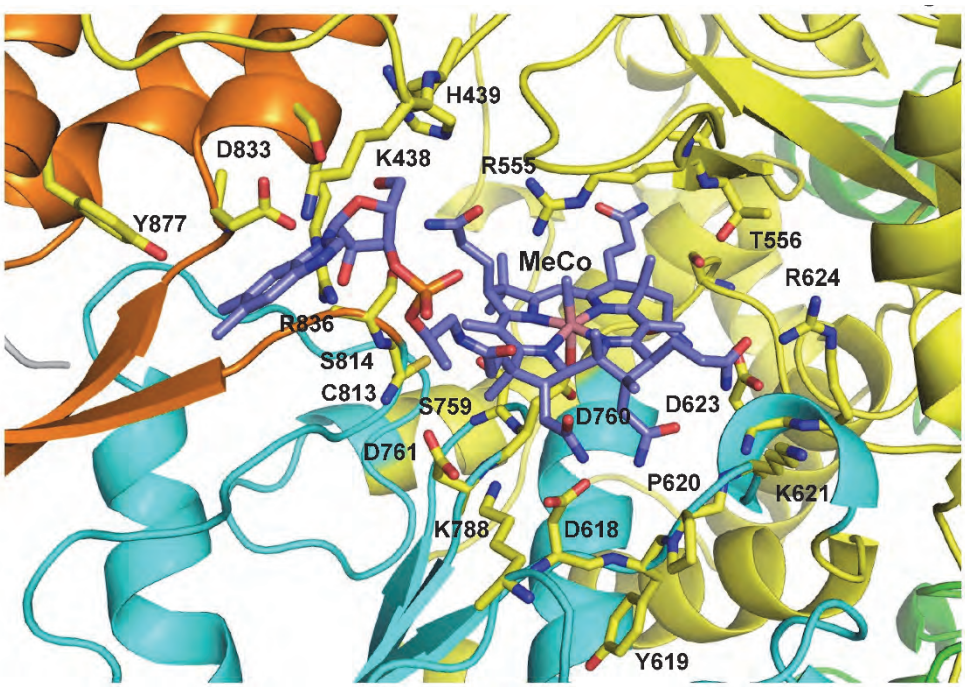


Figure 3: ***Computational model of nsp12 in complex with methylcobalamin****. The Vitamin B12 (MeCo) molecule is shown in stick representation and coloured according to element. The N-terminal extension region and the palm, fingers and thumb domains are shown in green, cyan, yellow and orange, respectively. The interacting residues are coloured according to element and shown in stick representation.*

Publication: Narayanan, N. and **Nair, D. T.** (2020) Vitamin B12 May Inhibit RNA-Dependent-RNA Polymerase Activity of nsp12 from the SARS-CoV-2 Virus. **IUBMB Life** 72:2112-2120. (Impact Factor: 3.885)

The SARS-CoV-2 genome encodes for the nsp14 protein, which houses the exoribonuclease (ExoN) activity responsible for the proofreading during replication of the viral genome. This protein forms a complex with the nsp10 protein and this complex is functionally active. The nsp14/10 complex activity is responsible for neutralizing the effect of chain terminating drugs such as remdesivir. The homology model of the SARS-CoV-2-nsp14/nsp10 complex available at the SWISS-MOD website was used to carry out *in silico* screening to identify molecules among natural products, or FDA approved drugs that can potentially inhibit the activity of nsp12. This exercise showed that the HIV-1 protease inhibitor ritonavir may bind to the exoribonuclease active site of the nsp14 protein (Figure 2). A model of the nsp14/10 in complex with substrate RNA showed that ritonavir binding site overlaps with that of the substrate RNA. A comparison of the calculated energies of binding suggested that ritonavir may bind to the active site of nsp14 with affinity comparable to that of the natural substrate. It is, therefore, possible that

ritonavir binding may prevent association of nsp14 with RNA and thus inhibit the exoribonuclease activity. Overall, our computational studies suggest that ritonavir may serve as an effective inhibitor of the nsp14 protein and thus potentiate the activity of chain terminating drugs such as remdesivir.

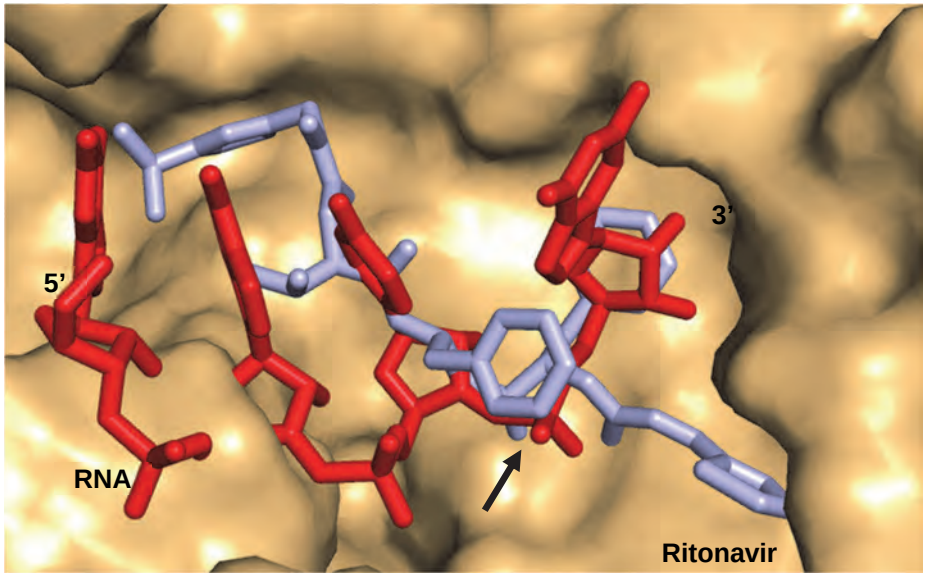


Figure 4: ***Binding site of ritonavir overlaps with that of substrate RNA****. Superimposition of the models of SARS-CoV-2-nsp10-nsp14:ritonavir (red) and that of SARS-CoV-2-nsp10-nsp14:ritonavir (cyan) are displayed. The surface of protein molecule is displayed in light orange. The site of cleavage on RNA is marked by an arrow. The comparison suggests that the presence of ritonavir may prevent binding of the natural RNA substrate.*

Publication: Narayanan, N. and **Nair, D. T.** (2021) Ritonavir may inhibit exoribonuclease activity of nsp14 from the SARS-CoV-2 virus and potentiate the activity of chain terminating drugs. ***Int. J. Biol Macromol***. 168:272-278. (Impact Factor: 6.953)

Prof. Deepak T. Nair

Regional Centre for Biotechnology