ABSTRACT

Compared to other expression systems, bacterial system (*E.coli)* remains the best candidate expression system due to their cost-effectiveness, well-studied genetics, high yield and short doubling time. The application of recombinant protein is widely present in the structural and functional characterization & biochemical studies of protein that requires a large amount of protein. In this work, we have studied the heterologous expression of eukaryotic (*Homo sapiens*) transmembrane protein, tetraspanin CD151, which is considered as one of the most challenging protein to express in the bacterial system and prokaryotic protein (*Helicobacter pylori*) Inosine-5'-monophosphate dehydrogenase, which is an equally challenging protein to express in soluble and active form.

Tetraspanin CD151 is involved in the various crucial biological process. It plays a significant role in cancer by modulating different stages such as tumour growth, metastasis & drug sensitivity in various types of cancers (skin and lung) and viral infectious diseases. Targeting CD151 could be a promising therapy for cancer and viral infections, but this requires structural information to understand the mechanism of the protein. While considering the challenges in expressing & purifying transmembrane protein in *E.coli*, we have introduced a construct wise expression approach for the recombinant expression of *h*CD151 in the bacterial system. The full-length *h*CD151 protein has been constructed into four different constructs and cloned into pMALc-5X-His. Construct-1 contains the full-length protein (28kDa), Construct-2 contains intracellular(IC) and large extracellular loops (LEL), and the 3rd and 4th transmembrane domain (TMDs) up to the C-terminal, the Construct-3 contains only the LEL, and the Construct-4 contains the LEL along with the 4th TMD helices up to the C-terminal end. Full-length *h*CD151 protein expression was not observed; however, in the other three constructs, the expression was observed. The reduction of TMD helices tailed by the presence of the extracellular region that fused with the MBP tag allowed the overexpression of the constructs. The recombinant expression of *h*CD151 constructs was confirmed by western blot and mass fingerprint analysis. The comparison of the CD curve of MBP and the *h*CD151constructs clearly showed the presence of helical content in the recombinant *h*CD151 constructs.

Based on the recently reported full-length crystal structure of *h*CD81 [5TCX] complexed with cholesterol molecule, we have predicted and characterized the structure of *h*CD151. The docking

study followed by the MD simulation showed an intramembrane pocket for the cholesterol moiety in *h*CD151 as seen in *h*CD81. The simulation results showed the apo-protein is stabilized by the hydrogen bond between Tyr^{23} (TMD-1) and Gln^{234} (TMD-4) and the cholesterol bound structure is stabilized by salt bridge formation via Lys^{45} (TMD-2) - Glu^{140} (LEL) & Glu^{218} (LEL) – Arg^{221} (TMD-4). These salt bridge formation and the hydrogen bonding between the apo and complex *h*CD151 structure suggest that the *h*CD151 protein also might undergo the open and closed conformations based on the cholesterol molecule presence as in *h*CD81. The crystallization trials were performed with the purified *h*CD151 constructs fused with MBP protein. The initial crystal hit was observed in the screening condition, 0.2 Ammonium Sulfate and PEG-8000 at 16° (Crystal Screen HT C7). In this condition, microcrystals of 10-20µm were observed with poor diffraction quality. The conditions have to be optimized to get good quality crystal in a desirable size to proceed for the structural characterization of *h*CD151.

Helicobacter pylori (*H. pylori*) bacteria is the major cause of several gastric disorders and recognized as a type I carcinogen by WHO. Due to the resistance developed by *H. pylori* strains, currently used antibiotic based treatments demonstrate high failure rates. Recently, bacterial IMPDH enzyme involved in *denovo* guanine biosynthesis pathway has been studied as a potential target to treat bacterial infections. Differences in the structural and kinetic parameters of the eukaryotic and prokaryotic IMPDH makes it possible to target bacterial IMPDH selectively and hence, the IMPDH protein from the *H.pylori* can be a valid target for the infection.

To understand the functional and structural characterisation of *Hp*IMPDH, we have recombinantly expressed the *Hp*IMPDH protein and confirmed by western blot and peptide mass fingerprint. The presence of secondary structure elements in the *Hp*IMPDH is characterized by circular dichroism. The ability of the recombinant *Hp*IMPDH to convert the NAD^+ to NADH shows the bioactivity of the recombinantly expressed *Hp*IMPDH. The kinetic profile of the recombinant protein showed the K_M value of 18.36±1.53 µM and 76.37±1.09 µM respectively, for IMP and NAD^+. To understand the structural characterization of the *Hp*IMPDH, crystallization trials were carried out. Crystal hits were observed in (a).0.2 M Sodium Fluoride, 20% w/v Polyethylene glycol 3,350 (PEG/ Ion screen ^{TM} HR2 -126) at 16°C and (b). 0.2 M L-Proline, 0.1 M HEPES pH 7.5 10% w/v Polyethylene glycol 3,350 (INDEX) at 16°C). These crystals were tried for the X-ray diffraction experiment in RRCAT synchrotron facility, and no diffraction

pattern was observed. Hence, these conditions have to be further optimised for the better quality crystal.

From our inhibitory studies, we identified, a new indole-based scaffold against *Hp*IMPDH and they showed non-competitive inhibition against IMP and NAD^+, whereas the already reported benzimidazole compound C91 (IC_{50} 0.18µM) was found to be an uncompetitive inhibitor. Further, we have identified indole-based diazo compounds to be the most potent inhibitors amongst the database with IC_{50} of 0.8 ±0.02µM and 1±0.03µM, and this can be a good lead for the further development of the selective and potent inhibitors of *Hp*IMPDH.

Mutational studies were performed in the NAD^+ binding domain (Asn^{252}Ala and Tyr^{439}Ala) and the flap region (Asp^{396}Gly, Arg^{397}Gly, Tyr^{398}Gly, Arg^{397}Lys & Tyr^{398}Phe) of *Hp*IMPDH for a better understanding of the catalytic reaction. The enzymatic characterization of Asn^{252}Ala, Asp^{396}Gly and Tyr^{439}Ala showed the binding pattern of the substrates and the cofactor had been affected in the mutant enzymes as compared to wildtype. The Asp^{396}Gly showed Asp^{396} could be one of the critical factors in maintaining the open and closed equilibrium of the *Hp*IMPDH protein. The NADH product inhibition in the absence of RY dyad (Arg^{397}Gly, Tyr^{398}Gly, Arg^{397}Lys & Tyr^{398}Phe) revealed both the residues are equally crucial in the NADH release and their reported role of water activation as that of Euk-IMPDH. These studies would aid for the better understanding of the catalytic reaction of *Hp*IMDH and designing of higher affinity inhibitory molecules with the host specificity.