Abstract for the general public

Nonsense mediated mRNA decay (NMD) is a quality-control checkpoint that detects and eliminates aberrant messenger RNAs (mRNAs) with premature termination codons (PTCs). The PTCs are often more deleterious compared to missense mutations, as they result in the dysfunctional or complete loss of protein function or expression. In addition, the PTCs arise from single nucleotide mutations that convert a canonical triplet nucleotide codon into one of three stop codons, e.g., TAG, TGA, or TAA. The abundance of mRNA containing the PTCs are reduced through the NMD, but in some cases, the truncated proteins may have dominant negative functions. As such, PTCs represent a unique constellation of diseases which afflict over 30 million people worldwide, accounting for 10-15% of all genetic diseases. Majority of the gene mutations found in genetic disorders, including cancer, result in premature termination codons, and the rapid degradation of their mRNAs by NMD.

The nonsense-mediated mRNA decay pathway plays an important role in the degradation of mRNA containing PTCs, but the molecular mechanism and structural rearrangements during this process has not been fully delineated. However, the working model of NMD with exon junction complex (EJC) and eukaryotic release factors (eRFs) complexes is proposed. The NMD machinery involves the UPFs proteins (UPF1, UPF2, UPF3a/b), and also includes the suppressors with morphological effects on genitalia proteins (SMG1, SMG5, SMG6, SMG7, SMG8, and SMG9), these proteins are bound to mRNAs via the exon junction complex. The EJC complex recruits the evolutionarily conserved UPF proteins that play an essential role in NMD. During the pioneer round of translation some EJC components are displaced by ribosome, and this positional information by EJC is preserved until the mRNA is translated. At some point, the PTC is recognized and the translation pauses upstream of the EJC. Eukaryotic release factors physically binds and recruit UPF1, the RNA helicase. Thereafter, SMG1 (phosphoinositide 3-kinase-related kinase) is recruited, which leads to the formation of the SURF complex (SMG1, UPF1, eRF). Along with the UPF proteins this SURF complex, promotes the phosphorylation of UPF1 by SMG1. In contrast, the dephosphorylation of UPF1 requires a multiprotein complex, that is composed of SMG5, SMG6, SMG7, and protein phosphatase 2A. This suggests that NMD is a process in which multiprotein comes into action and binds to the aberrant mRNA, therefore protein-protein, protein-RNA, or RNA-RNA binding networks become an interesting area to be investigated. Protein-protein/RNA interactions are fundamental to the formation of intricate interaction networks and the assembly of multisubunit protein complexes that represent the functional workhorses of the cell.

Genetic disease and cancer represents a significant burden to the society, and any new information, mechanistic details, and molecular properties about different components from NMD machinery could help to bring new therapeutic strategies that are beneficial to the society. The current project originates from an unmet need to understand the NMD mechanisms, and therefore protein-protein or protein-RNA or RNA-RNA interactions will be studied using different *in silico* (molecular modeling and molecular dynamics) and mass spectrometry (MS)-based structural techniques (or the proteomics techniques). The overall cancer patient survival status relating to the components from NMD, EJC, and eRFs, suggest that there is substantial decrease in the survival in the patients from the altered group. Understanding the interacting partners or the components from the eRFs, NMD, and EJC complexes as well as their involvement in different important pathways suggest that these factors are involved in several critical and important functions of the human body. In addition to the nonsense-mediated mRNA decay pathway, components from eRFs, NMD, and EJC are widely involved in the regulation of telomere maintenance, regulation of chromosome organization, and RNA transport/localization/catabolic process. Therefore, it becomes crucial to understand the properties and effects of cancer mutations on the functioning of different components from the NMD machinery. The outcome of this project will bring a mechanistic overview with characterising molecular properties of the eRFs, NMD, and EJC complexes. Reaching these aims, the project will develop data science techniques aimed at the high-throughput analysis of protein complexes, their structures and methods to put these into quickly use these features for the interpretation of cancer mutation landscapes. Particularly in the context of RNA-protein interactions this is lacking, adding a novel dimension to structural informatics.