SUMMARY

Due to the central role played by ubiquitin in proteasomal degradation of proteins as well as in several other functions like DNA repair, regulation of transcription, protein synthesis, cell cycle, endocytosis and lysosomal degradation, it has been the subject of multiple studies ever since its discovery. And the extraordinary level of sequence conservation displayed by it makes it all the more interesting to study. Nearly all of ubiquitin's residues, except three, are conserved from yeast to humans (Gavilanes et al., 1982; Watson et al., 1978; Schlesinger et al., 1975; Schlesinger and Goldsteiner, 1975). This extraordinary sequence similarity between yeast and human ubiquitins means that inferences derived from the study of yeast ubiquitin can be extrapolated to human ubiquitin as well.

There have been several structural studies previously done on ubiquitin. The structure of ubiquitin was solved using X ray diffraction crystallography (Vijay-Kumar et al., 1987). The results showed that ubiquitin has a β grasp fold, consisting of a β sheet surrounding, or "grasping" an α helix. Ubiquitin also consists of two bulges. One near its N terminal and the other near its C terminal.

The structural studies previously carried out on ubiquitin include systematically replacing the surface residues on ubiquitin with alanine (Sloper-Mould et al., 2001), to identify those surface residues that are essential for life in yeast, mutating charged surface residues to determine the role of salt bridges or ionic bonds in the stability of ubiquitin, and replacing non-polar core residues with polar ones and vice versa, to determine the role of non-polar and polar residues in determining the structural stability of ubiquitin (Beal et al., 1996; Haririnia et al., 2008; Roscoe et al., 2013; Lee et al., 2014). Our laboratory has focused on two different categories of mutations in ubiquitin in order to better understand the role of different residues in it. One category consists of the UbEP42 mutation and its derivatives (Prabha et al., 2010; Doshi et al., 2014; Doshi et al., 2017; Sharma et al., 2021), and the other category consists of the substitutions of ubiquitin's G1 β bulge residues and their derivatives (Mishra et al., 2009; Mishra et al., 2011; Sharma et al., 2015).

The UbEP42 mutation was generated through random mutagenesis using error prone PCR (Prabha et al., 2010). UbEP42 mutation consists of 7 base substitutions, 3 of which do not result in amino acid substitutions as they resulted in synonymous codons and are therefore of no consequence. The other 4 base substitutions resulted in S20F, A46S, L50P and I61T replacements in the amino acid sequence. Ubiquitin plays a pivotal role in helping the cell survive heat stress, by degrading denatured and misfolded proteins. The UbEP42 mutation does

not just fail to complement the stress hypersensitive phenotype of SUB60 cells, but actually reduces the survival of SUB60 compared to the untransformed SUB60 cells under heat stress. This suggests that the ubiquitin carrying UbEP42 mutation is not just defective in many of its functions but competes with wild type ubiquitin, impeding its functioning as well (Doshi et al., 2014). Ubiquitin carrying UbEP42 mutation may therefore be used as a competitive inhibitor to manage diseases resulting from overexpression of ubiquitin in cells.

Unlike the UbEP42 mutation that was randomly generated through error prone PCR, the Q2N, E64G and S65D substitutions of the G1 β bulge were planned substitutions that were generated through site directed mutagenesis. The amino acids to substitute the naturally occurring ones were carefully selected based on their occurrence at the same position in G1 β bulges in other proteins (Chan et al., 1993), so that the substitutions would not disturb the structure of the β bulge. In fact, the Q2, E64 and S65 residues present in the β bulge of ubiquitin are found less frequently in the G1 β bulges of other proteins. Instead, the residues N, G and D, are found more commonly in the respective positions in the G1 β bulges of other proteins. Despite this, the Q2, E64 and S65 residues are highly conserved in ubiquitin, suggesting that they play some vital structural role. Indeed, previous studies in our lab showed that all three β bulge substitutions increase sensitivity of yeast to cycloheximide, which is a translational inhibitor acting on the elongation phase of protein synthesis (Mishra et al., 2009; Mishra et al., 2011).

The generation of UbEP42 mutation and the β bulge mutations represent two different approaches to studying protein structure. The generation of β bulge mutations is an example of informed changes to the amino acid sequence of a protein, based on information already known about the structure and functionality ofthat region of the protein. In other words, it is planned. It was known that the G1 β bulge of ubiquitin displayed unusual features, like residues that are not usually found in it, and yet displayed remarkable conservation. It was also known that other residues occur far more frequently in G1 β bulges in other proteins, meaning that using them to replace the β bulge residues of ubiquitin would not affect the structural integrity of β bulge. It was based on this knowledge, that substitutions of β bulge residues were carried out using site directed mutagenesis.

However, this planned approach based on already available information about the protein's structure, limits the possibilities of studying protein structure to those regions of the protein that are already predetermined for either their functional and/ or structural significance.

Residues that might play a vital structural or functional role, but are not yet identified and characterized remain unstudied. To address this limitation of planned approach, *in vitro* evolution using error prone PCR is commonly used to study protein structure, and this is the approach through which the individual significances of the S20, A46, L50 and I61 residues of ubiquitin have been discovered through the randomly generated UbEP42 mutation. Random mutagenesis therefore expands the field of study of protein structure by helping us discover roles of various residues that were previously overlooked.

Ubiquitination is involved in several cell functions. The functions affected by the β bulge and UbEP42 substitutions are briefly described here. Generation time is the time that yeast takes to double its population, and is directly affected by many defects in the ubiquitin system. Ubiquitin proteasome system also plays a significant role in cell cycle. A cell's transition through the stages of the cell cycle is made possible by degradation of certain regulatory proteins involved in the various stages by the proteasome. CDK1 (Cyclin Dependent Kinase 1) a serine/threonine kinase, is one of the regulatory proteins acting on the cell cycle. In *Saccharomyces cerevisiae*, CDK1 is encoded by the cdc28 gene, and is known to be the only cyclin dependent kinase to regulate cell cycle. While proteasome degrades cytosolic proteins, the lysosome is responsible for the degradation of membrane proteins. And previous studies indicate that ubiquitin is involved not only in proteasomal degradation, but in lysosomal degradation as well. Uracil permease is one of the membrane proteins that are degraded in lysosomes, and whose lysosomal degradation is dependent on the ubiquitin system.

The protein composition of the cell changes over time through the proteasomal degradation of proteins, caused by their K48 linked polyubiquitination. One of the possible mechanisms underlying the effect of UbEP42 derived double mutations on ubiquitin function is the prevention of the formation of K48 linked polyubiquitin chains by these mutations. K-63 linked polyubiquitination is another form of ubiquitination that could be affected by β bulge and UbEP42 mutations. Previous studies have shown the involvement of ubiquitin system in the regulation of protein synthesis. L28, a ribosomal protein which increases the resistance of yeast to certain translational inhibitors, has been found to undergo K-63 linked polyubiquitination (Spence et al., 2000). Many other ribosomal proteins are also ubiquitinated. Previous studies in our lab have shown that the UbEP42 mutant cells are more sensitive than the wild type cells to cycloheximide, which is a translational inhibitor acting on the elongation phase of protein synthesis (Doshi et al., 2014).

The N-end rule decides the rate of degradation of a protein, based on the identity of the N-terminal residue of the protein (Varshavsky, 1998). In other words, the half-life of a protein is determined by the identity of the N-terminal amino acid of that protein. The N end rule applies to both eukaryotes and prokaryotes organisms. In eukaryotes, the N-terminal residue of the protein is recognized and targeted by ubiquitin ligases, which mark the protein for degradation by ubiquitinating it. The efficiency of this degradation is determined by the identity of the N terminal residue of the protein, as per the N end rule. Ubiquitin fusion degradation (UFD) pathway involves a translational fusion of ubiquitin fused to a reporter protein such as β-galactosidase with Pro as its N terminal residue, which is degraded entirely by proteasome, as ubiquitin C-terminal hydrolases cannot cleave between the two proteins (Bachmair et al., 1986; Baker and Board, 1991). The possible effects of the mutations generated in our laboratory on N-end rule pathway and UFD pathway were investigated.

Overall, the effects of β bulge substitutions seem to be far milder compared to the effects of UbEP42 substitutions. The β bulge substitutions do not affect the growth characteristics of yeast or its survival under heat stress. Nor do they seem to affect endocytosis and lysosomal degradation of membrane proteins, or the regulation of cell cycle. And while Q2N and E64G substitutions have a mild impact on the K48 linked polyubiquitination, that effect might not be too consequential to yeast, since the growth characteristics of yeast and its survival under heat stress have not been affected. The β bulge substitutions do, however, affect the regulation of protein synthesis, as seen from the increased sensitivity of β bulge mutants to translational inhibitors cycloheximide and G418. The fact that β bulge mutants have increased sensitivity to cycloheximide and G418, both of which act on the elongation phase of protein synthesis, but they do not show any change in sensitivity to hygromycin at the concentration of the antibiotic used.

In contrast to the mild effect exerted by the β bulge mutations, the UbEP42 derived double mutations exert detrimental effects on various functional aspects of the ubiquitin system. The most benign of all six double mutants is A46S-I61T, but does not have any another detrimental effect on yeast. The next most benign double mutation is S20F-A46S, which increases generation time of yeast when induced by copper sulphate. S20F-L50P, the third most benign double mutation, affects growth, generation time, viability and survival under heat stress, is lethal under overexpression, and displays heightened sensitivity towards cycloheximide. Moreover, previous structural studies in our laboratory have showed that while

β bulge mutations do not have an impact on the structural integrity of ubiquitin, the UbEP42 and its derivative mutations do impact structural integrity to some extent.

The other three double mutations are highly detrimental to yeast. S20F-I61T and A46S-L50P affect growth, generation time, viability and survival under heat stress, UFD pathway, Cdc28 cyclin dependent protein kinase levels, lysosomal degradation and K48 linked polyubiquitination, and display heightened sensitivity towards cycloheximide. In addition, A46S-L50P is lethal when overexpressed. L50P-I61T, the most detrimental of all six double mutations, displays all these detrimental effects of A46S-L50P. UbS20F-I61T, UbA46S-L50P, UbL50P-I61T mutants failed to sort CPS into endosomes, while in UbS20F-A46S, UbS20F-L50P, UbA46S-I61T mutants the sorting was as efficient as UbWt control.

The combinations in which the four substitutions of the UbEP42 mutations are most and least detrimental shows that the L50P and I61T substitutions are far more detrimental than the S20F and A46S substitutions (Sharma et al., 2021). For example, the three most detrimental double mutations all contain at least one of these two substitutions, and the most detrimental of them all, namely L50P-I61T, contains both of them. In contrast, one of the most benign double mutations, namely S20F-A46S, contains neither of the two substitutions. L50P substitution could be so detrimental because it involves a very substantial structural change, as leucine and proline have very different side chains. I61T could be detrimental as it involves substituting a hydrophobic residue buried in the core of ubiquitin, with a residue that is not only very structurally different, but also hydrophilic. Both S20 and A46 are surface residues, which might explain why their substitutions are relatively benign. Interestingly, A46S substitution involves substituting a surface hydrophobic residue with a hydrophilic residue, which is not known to decrease protein stability.

In summary, the β bulge mutations and the UbEP42 and its derivative mutations give us key insights into the role of these residues in ubiquitin's structure and functioning. A defective ubiquitin system has been implicated in numerous diseases, including Alzheimer's, Parkinson's and cancer, making it a prime area of research to explore novel methods of treatment. The high level of detrimental effects that the L50P-I61T double mutation has might make it useful as a competitive inhibitor to manage diseases resulting from overexpression of ubiquitin. This work, like many previous studies, also demonstrates the usefulness of ubiquitin as a model to study protein function, owing to its small size, and monomeric, single domain, and globular nature. the insights gained in this study about ubiquitin's structure and function

ubiquitin will be useful in developing drugs targeting the ubiquitination system. Table 4 is a summary of the functional characterisation of UbEP42 double mutants.

Table 4: Summary of the results of the functional characterisation of UbEP42 double mutants. Bold font signifies a change with respect to cells expressing ubiquitin wild type (UbWt).

Mutant	**20–46**	**20–50**	**20–61**	**46–50**	**46–61**	**50–61**
UFD pathway	Shows complementation	Shows complementation	**No complementation**	**No complementation**	Shows complementation	**No complementation**
Cdc28	Shows complementation	Shows complementation	**No complementation**	**No complementation**	Shows complementation	**No complementation**
Uracil permease	Shows complementation	Shows complementation	**No complementation**	**No complementation**	Shows less complementation	**No complementation**
K48 poly-ubiquitination	Shows complementation	**No complementation**	**No complementation**	**No complementation**	Shows complementation	**No complementation**
Endosomal sorting of CPS	Not affected	**Failure of sorting**	**Failure of sorting**	**Failure of sorting**	Not affected	**Failure of sorting**