Chapter 3

CLONING AND CHARACTERIZATION OF GENES INVOLVED IN ALTERNATIVE PATHWAY OF GLUTATHIONE DEGRADATION
Chapter 3

3.1 Introduction

As described in the general introduction (section 1.5), genetic investigations on the alternative pathway for GSH degradation were initiated by previous researchers in the laboratory (Kumar et al. 2003b; Kumar C., Ph.D. Thesis, 2005). The genetic investigation into the degradation pathway revealed 4 dug complementing groups (dug = defective in utilization of GSH). Of these, the dug4 complementing group was found to be identical to hgt1 (yeast GSH transporter), while the DUG1 gene was cloned by library complementation and was revealed to be a gene coding for a hypothetical ORF, YFR044c with a M20 peptidase domain. The genes for the dug2 and dug3 complementation groups however were not cloned.

In the present chapter, I describe the cloning and initial characterization of the DUG2 and DUG3 genes.

Results

3.2 Isolation of a DNA fragment complementing the dug2 mutant.

To identify the gene corresponding to the dug2 mutant group, a dug2 mutant was transformed with a LEU2 based multicopy yeast genomic library (Fig3.1a). Transformants were directly selected on SD-LEU+0.2mM GSH. Colonies that were able to grow on GSH were selected. Of the total of 60,000 transformants screened, 55 were able to grow on GSH. The complementation of the mutant by MET15 gene product would also lead to growth, as the strain would be able to utilize inorganic sulphate. To eliminate these clones, the transformants were examined for the growth on SD-LEU plates without any added organic sulphur source. Out of the 55 transformants, 44 grew on sulphur-free medium and were likely to be fragments containing the MET15 gene and were not subjected to further analysis. Plasmids were extracted from two of the remaining eleven yeast transformants that most likely harbored gene other than MET15, and these were transformed into E. coli.

Plasmids were isolated from these E. coli transformants and restriction fragment digestion carried out to indicate unique plasmids. Reconfirmation was carried out by re-transformation into yeast met15Δecm38Δdug2 and complementation of the growth defect on GSH was confirmed. Two purified plasmids that were showing different restriction
(met15Δ ecm38Δ dug2) or (met15Δ ecm38Δ dug3)

Transformation
(Yeast genomic library: LEU2)

~60000 transformants screened for growth on GSH as organic sulphur source

growth on GSH containing Medium

growth on GSH containing Medium but can not utilize inorganic sulphur

Plasmids isolated from 2 yeast transformants and transformed in E.coli

2 Plasmids with different Restriction Enzyme patterns
Re-transformed in the met15Δ ecm38Δ dugx strain

Functional complementation by growth and utilization of GSH confirmed

Sequencing of the flanks of the inserts of the 2 Plasmids

Sequence of insert assembled and compared with published genome sequence

Identification of ORFs

Fig 3.1a Flow diagram describing the strategy for isolation of a dug complementing DNA fragment
fragment patterns (pDUG2-C1 and pDUG2-C2) and were able to complement the GSH growth defect were investigated further. The plasmid with the smallest insert, pDUG2-C1 was subjected to sequence analysis by sequencing the two ends of the inserts. Analysis of the sequences and comparison with the published genome sequence revealed an insert of 6.235kb that was also supported by restriction fragment analysis (data not shown). The insert contained two partial ORFs, \textit{YBR280c} and \textit{YBR283c}, and two complete ORFs \textit{YBR281c} and \textit{YBR282w} (Fig 3.1b). Two ORFs \textit{YBR282w} and \textit{YBR283c} coded for protein of known function and while other two ORFs, \textit{YBR280c} and \textit{YBR281c} were described as hypothetical proteins of unknown function as seen in the \textit{Saccharomyces cerevisiae} genome database (SGD) (Table 3.1).

3.3 The ORF \textit{YBR281c} corresponds to the \textit{DUG2} gene.

Among the 4 ORFs (2 partial and 2 complete) (Fig 3.1b) present on the fragment complementing the \textit{dug2} mutant, the ORF \textit{YBR281c} seemed to be the most likely candidate corresponding to \textit{DUG2}, since sequence analyses of Ybr281cp revealed a M20 peptidase domain in the C-terminal part of the protein. To investigate if indeed \textit{YBR281c} was involved in the utilization of GSH as a sole source of sulphur, a \textit{met15\textsuperscript{Δ}ybr281c\textsuperscript{Δ}} strain was procured from Euroscarf and examined for its growth on SD medium having either GSH or methionine as sole source of sulphur. We observed that \textit{met15\textsuperscript{Δ}ybr281c\textsuperscript{Δ}} failed to grow on GSH containing plate (Fig 3.2b) and broth (Fig 3.9) while it grew on methionine containing plates (Fig 3.2a) and broth (Fig 3.9). This suggested that \textit{YBR281c} is essential for the utilization of GSH as a sole source of sulphur in \textit{met15\textsuperscript{Δ}} strain background, and indicated that \textit{dug2} mutants carried a mutation in \textit{YBR281c} locus.

3.4 Sequence analysis of the Dug2p reveals an N-terminal WD40 repeat region and C-terminal M20 peptidase domain.

To further confirm if \textit{YBR281c} corresponds to \textit{DUG2}, we cloned the \textit{YBR281c} ORF in a centromeric plasmid by PCR amplification under the TEF promoter to yield p416TEF-YBR281c (p416TEF-DUG2). The \textit{dug2} mutant was transformed with this plasmid as well as vector alone. The p416TEF-YBR281c clone restored the growth defect of \textit{dug2} mutants for utilization of GSH as a sole source of organic sulphur while
Fig 3.1b Schematic diagram showing the ORFs present in the sequenced clone (pDUG2-C1) that complements the GSH utilization defect in dug2 mutant. The ORFs present are shown in colour blocks. The region spanned by the insert is shown by black coloured arrows.

<table>
<thead>
<tr>
<th>ORFs present</th>
<th>Description</th>
<th>Function</th>
<th>Protein class/Conserved domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>YBR280c</td>
<td>Hypothetical</td>
<td>Unknown</td>
<td>Putative F-Box protein; interacts with Skp1</td>
</tr>
<tr>
<td>YBR281c</td>
<td>Hypothetical</td>
<td>Unknown</td>
<td>WD40 domain Metallopeptidase</td>
</tr>
<tr>
<td>YBR282w</td>
<td>YmL27</td>
<td>Structural constituent of ribosome</td>
<td>None</td>
</tr>
<tr>
<td>YBR283c</td>
<td>Sec61p Homologue</td>
<td>Protein transporter activity</td>
<td>SecY; Preprotein translocase</td>
</tr>
</tbody>
</table>

Table 3.1 List of ORFs present in the yeast library clone (pDUG2-C1) that complements the GSH utilization defect of dug2 mutant
**Fig 3.2** Growth of *met15A ybr281cA* and *met15A* strains on SD plates having (a) methionine and (b) GSH as the sole exogenous source of sulphur. GSH and methionine were used at concentration of 0.2 mM.

**Fig 3.3** Complementation of growth defect of *dug2* mutant on GSH containing plates by *(DUG2) YBR281c* *(p416TEF-YBR281c)*. SD-ura plates having (a) methionine and (b) GSH as the sole source of sulphur. GSH and methionine were used at concentration of 0.2 mM.
the cells transformed with vector alone showed no growth on plates (Fig 3.3). Moreover, the over-expression of Dug2p (Ybr281cp) in ABC733 (met15Δ) strain didn’t lead to any altered growth phenotype of cells (Fig 3.10).

Sequence analysis of Dug2p predicted a protein of 878 amino acids. BLAST analysis of Dug2p protein revealed that the protein has 2 distinct domains. At the N-terminal domain (1-422 aa) is a putative WD40 repeat region, while the C-terminal region (441-878 aa) shows homology to the M20A metallopeptidase family (Fig 3.4a). The WD40 sequence-containing repeat is a protein interaction motif that is normally present as a tandemly repeated unit of variable number. The sequence repeat unit is normally characterized by the presence of Gly-His (GH) dipeptide, 11-24 residues from it’s N-terminus or start of the repeat and ends with a Trp-Asp (WD) dipeptide at it’s C-terminus and it is about 44-60 amino acids long, containing conserved core of amino acid residues with separating variable length region. The WD40 repeat region usually contains 4-10 repeat units. A minimum of 4 repeat units is required for folding into the higher order β-propeller structure with several blades. Each blade is composed of four-stranded anti-parallel β-sheet. Each WD40 sequence repeat forms the first three strands of one blade and last strand in the next blade thereby forming a closed-ring propeller structure. This presents a structural element to which proteins can bind stably or reversibly (Smith et al. 1999).

The proteins containing the WD40 repeat regions participate in numerous functions that include cytoskeletal assembly, vesicular trafficking, signal transduction, and control of transcription initiation complex, mRNA splicing among others. The WD40 repeat region of these proteins is a modular interaction domain that is often a component of larger proteins, and may function to form large multivalent complexes or act as a binding site for 2 or more proteins to enable transient interactions among these and other proteins to occur (Smith et al. 1999). *Saccharomyces cerevisiae* has at least 23 proteins that have WD repeats and a few more that have possible WD repeats as seen from the SCOP database (Murzin et al., 1995). Interestingly the protein Dug2p (Ybr281cp) is not included in this list possibly owing to the very weak consensus seen with the classical WD40 repeat regions. However this region is listed as a WD40 repeat region in the conserved domain database (Marchler-Bauer et al., 2004) (Fig 3.4a) with significant e-
PSSMs producing significant alignments:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length of the region</th>
<th>E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gln</td>
<td>CDD</td>
<td>29257 cd00200.2 (WD40)</td>
</tr>
<tr>
<td>gln</td>
<td>CDD</td>
<td>29257 cd00200.2 (WD40)</td>
</tr>
<tr>
<td>gln</td>
<td>CDD</td>
<td>29257 cd00200.2 (WD40)</td>
</tr>
<tr>
<td>gln</td>
<td>CDD</td>
<td>41591 pfam01546 (Peptidase_M20)</td>
</tr>
</tbody>
</table>

Fig 3.4 (a) Dug2p consists of two distinct domains, N-terminal WD40 repeat domain and C-terminal peptidase of class M20. (using CDD, NCBI).
(b) Alignment study using PSSMs-NCBI shows presence of WD40 domain with significant similarity in Dug2p.
value (Fig 3.4b). Furthermore, the multiple sequence alignment with other fungal homologues using Clustal X indicates several conserved \( W \) residues in this region (Fig 3.4c).

The Dug2p (Ybr281cp) C-terminal M20 peptidase domain (441-878aa) shows a significant homology to Dug1p (Yfr044cp), which belonged to the M20 peptidase family. Both the proteins belong to the metallohydrolases of the MEROPS MH clan, family M20 and sub-family A of \( \text{Zn}^{2+} \) metallopeptidases (Rawlings et al., 2006).

The \textit{Lactobacillus delbrueckii} PepV protein is also a member of the above mentioned family, and its crystal structure has been solved and the key amino acid residues involved in Zinc binding, catalysis and substrate binding have been very well characterized. The amino acid residues D119, E154 and H439 of LdPepV are involved in Zinc 1 binding whereas the amino acid residues D117 and H87 are believed to bind to Zinc 2 atom. E153 is the key residue involved in catalysis (Jozic et al. 2002).

The sequence alignment of the Dug2p (Ybr281cp) (421-878aa), Dug1p (Yfr044cp) (1-481aa) and \textit{Lactobacillus delbrueckii} PepV (1-470aa) protein revealed that the key Zinc 1 and Zinc 2 binding residues present in LdPepV protein (as mentioned above) were highly conserved across the three proteins (Fig 3.4d). The corresponding conserved residues in Ybr281cp are D553, E587 and H853, which may be involved in Zinc 1 binding and apart from these, H520 and D607 may be involved in Zinc 2 binding. The residue E153 of LdPepV involved in catalysis was also conserved in Dug1p and Dug2p.

### 3.5 Isolation of a DNA fragment complementing the \textit{dug3} mutant.

To identify the gene corresponding to the \textit{dug3} mutant group, we transformed a \textit{dug3} mutant with a LEU2 based multicopy yeast genomic library exactly as described for the \textit{dug2} mutant (Section 3.2). Transformants were directly selected on SD-LEU+0.2mM GSH. Colonies that were able to grow on GSH were selected. Out of a total of approximately 20,000 transformants, 62 were growing on GSH. As complementation of mutants by a fragment containing the \textit{MET15} gene product would also lead to growth, as the strain would be able to utilize inorganic sulphite, these needed to be eliminated. Therefore all the transformants were plated on SD-LEU plates without added organic
Fig 3.4 (c) Multiple sequence alignment of WD40 repeat domain of Dug2p (Ybr281c) protein of *Saccharomyces cerevisiae* (ScYBR281c, 1-396 amino acids) with homologues from other organisms: *Candida glabrata* (CgYBR281c homolog, 1-396 amino acids), *Ashbya gossypii* (AgYBR281c homolog, 1-396 amino acids), *Kluyveromyces lactis* (KlYBR281c homolog, 1-396 amino acids). The multiple sequence alignment was carried out using CLUSTAL X (Thomson et al., 1997) and reformatted using CLOURE-C (Kohli and Bachhawat, 2003). The conserved W residues of WD40 repeat motif is highlighted.
Fig 3.4 (d) Multiple sequence alignment of M20 peptidase domain of Ybr281cp (Dug2p) protein of *Saccharomyces cerevisiae* (ScYBR281c, 421-878 amino acids) along with YFR044cp (Dug1p) protein of *Saccharomyces cerevisiae* (ScYFR044c, 1-481 amino acids) and PepV protein of *Lactobacillus delbrueckii* (Ld PepV, 1-470 amino acids). The multiple sequence alignment was carried out using CLUSTAL X (Thomson et al., 1997), manually edited and was reformatted using CLOURE-C (Kohli and Bachhawat, 2003). The conserved amino acid residues important for Zn^{2+} binding are highlighted by black arrows and catalytically important residue by Red color (Jozic et.al, 2002).
Fig 3.5 Schematic diagram showing the ORFs present in the sequenced clone (pDUG3-C1) that complements the GSH growth defect in *dug3* mutant. The ORFs present are shown in colour blocks. The region spanned by the insert is shown by black coloured arrows.

<table>
<thead>
<tr>
<th>ORFs present</th>
<th>Description</th>
<th>Function</th>
<th>Protein class/Conserved domains</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>YNL192w</em></td>
<td>CHSI Chitin synthase I</td>
<td>Chitin synthase activity required in cytokinesis</td>
<td>Chitin synthase_1 and Chitin synthase_2</td>
</tr>
<tr>
<td><em>YNL191w</em></td>
<td>Hypothetical ORF</td>
<td>unknown</td>
<td>Glutamine-amidotransferase class II</td>
</tr>
<tr>
<td><em>YNL190w</em></td>
<td>Hypothetical ORF</td>
<td>unknown</td>
<td>none</td>
</tr>
<tr>
<td><em>YNL189w</em></td>
<td>SRPI Karyopherin alpha homolog.</td>
<td>forms a dimer with karyopherin beta Kap95p, to mediate import of nuclear proteins</td>
<td>Importin beta binding domain (IBB) Armadillo/beta-eatenin-like repeats (ARM)</td>
</tr>
</tbody>
</table>

Table 3.2 List of complete ORFs present in the yeast library clone (pDUG3-C1) that complements the GSH utilization defect of *dug3* mutant
sulphur source. Among the 62 transformants, 49 were growing on sulphur free medium and were likely to be MET15 and were not picked for further analysis. Plasmids were extracted from two of the thirteen yeast transformants that appeared to have plasmids other than the plasmid harboring MET15 ORF and these were transformed into E. coli.

Plasmids were isolated from these E. coli transformants and restriction fragment digestion carried out to indicate unique plasmids. Reconfirmation was carried out by re-transformation into yeast met15Δecm38Δdug3 and complementation of the growth defect on GSH. The two purified plasmids that were showing different restriction fragment patterns (pDUG3-C1 and pDUG3-C2) and were able to complement GSH growth defect, were selected. The plasmid with smallest insert, pDUG3-C1, was subjected to sequence analysis by sequencing the two ends of the inserts. Sequencing and comparison of the sequence with the published genome sequence of Saccharomyces cerevisiae (in SGD) revealed an insert of 6.287kb (Fig 3.5) that was supported by restriction enzyme fragment analyses (data not shown). The insert contained two partial ORFs, YNL189w and YNL192w, and two complete ORFs, YNL190w and YNL191w. Of these four ORFs, two ORFs YNL190w and YNL191w code for protein of unknown function, while the other two ORFs, YNL192w and YNL189w were proteins of known function in SGD (Table 3.2).

3.6 The ORF YNL191w corresponds to the DUG3 gene.

Among the four ORFs in the dug3 complementing fragment, two ORFs (YNL189w and YNL192w) were partial in the library clone. Sequence analyses of YNL191w revealed a protein with a glutamine-amido transferase type II domain. The sequence analysis of YNL190w revealed homologues with only hypothetical function. To investigate if either YNL191w or YNL190w was involved in the utilization of GSH as a sole source of sulphur, met15Δynl191wΔ and met15Δynl190wΔ strain were procured from Euroscarf and examined for their growth on SD medium having either GSH or methionine as sole source of sulphur. We observed that only met15Δynl191wΔ strain failed to grow on GSH containing plate (Fig 3.6b) and broth (Fig 3.9) while it grew on methionine containing plates (Fig 3.6a) and broth (Fig 3.9). This indicated that it is essential for the utilization of GSH as a sole source of sulphur in met15Δ strain background (Fig 3.6), and indicated that dug3 mutants possibly carried a mutation in the YNL191w locus.
To further confirm the result we cloned the YNL191w ORF in a centromeric plasmid under the TEF promoter to yield, p416TEF-YNL191w (p416TEF-DUG3). The dug3 mutant was transformed by this plasmid. The p416TEF-YNL191w clone restored the growth defect of dug3 mutant, associated with the inability for utilization of GSH as a sole source of sulphur, while the cells transformed with vector alone showed no growth (Fig 3.7). No growth defect or toxic effect in WT cells was observed due to over-expression of Ynl191wp (Fig 3.10).

3.7 Sequence analysis of the Dug3p reveals a Glutamine amidotransferase type II domain.

Sequence analysis of Dug3p revealed a protein of 357 amino acids that contains a putative glutamine amidotransferase (GATase) class II (YafJ_type) domain at its N-terminal region (residues 1-294) (Fig 3.8a and b). YafJ is a glutamine amidotransferase-like protein of unknown function found in prokaryotes, eukaryotes and archaea. YafJ has an Ntn hydrolase domain (described later) similar to those of the class II glutamine amidotransferases.

Dug3p shows considerable sequence similarity across its length spanning 45-254 amino acid with GATase Class II (accession no. cd00352), this class comprises of similar proteins of known biological function (Fig 3.8 a & b).

GATase enzymes have two structural domains, an N-terminal glutaminase domain that hydrolyzes the amide of glutamine and C-terminal acceptor/synthase domain that accepts the released amide and transfer it to some substrate.

All the GATase enzymes studied till date fall under two families (class I and class II), depending upon the type of glutaminase domain. The Class I GATase glutaminase domain has been found to contain a signature motif of Cys-His-Glu, which forms the catalytic triad, and thus called as the triad amidotransferase (formerly G-type) (Tesmer et al. 1996).

Members of the GATase class II family (to which Dug3p belongs) are able to catalyze the removal of ammonia from glutamine and then transfer it to a substrate to form a new carbon-nitrogen bond. The family includes glucosamine-fructose-6-phosphate amido transferase (GFAT/GlmS), glutamine phosphoribosyl pyrophosphate
Fig 3.6 Growth of met15Δ ynl191Δ and met15Δ strains on SD plates having (a) methionine and (b) GSH as the sole exogenous source of sulphur. GSH and methionine were used at concentration of 0.2 mM.

Fig 3.7 Complementation of growth defect of dug3 mutant on GSH containing plates by (DUG3) YNL191w (p416TEF- YNL191w). SD-ura plates having (a) methionine and (b) GSH as the sole source of sulphur. GSH and methionine were used at concentration of 0.2 mM.
Fig 3.8a The N-terminus of Dug3p shows homology to Glutamine Amidotransferase-classII (Gln-AT-II) domain (using CDD, NCBI)

PSSMs producing significant alignments:

<table>
<thead>
<tr>
<th>PSSMs</th>
<th>Length of sequence</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gln</td>
<td>CDD</td>
<td>30155 cd01908 (YafJ, Gn_AT-II)</td>
</tr>
<tr>
<td>gnl</td>
<td>CDD</td>
<td>30142 cd00352 (GnAT-II)</td>
</tr>
</tbody>
</table>

Fig 3.8b The N-terminus of Dug3p shows homology to Glutamine Amidotransferase-classII (Gln-AT-II) domain (PSSMs alignment results, NCBI)
Fig 3.9 Growth of *met15Δ, met15Δyfr044cΔ, met15Δybr281cΔ* and *met15Δynl191wΔ* on methionine or GSH as the sole source of sulphur (0.2 mM). Cells were grown overnight in SD medium containing (Methionine), washed twice with SD medium and reinoculated at an OD$_{600}$ of about 0.1 in SD medium containing MET or GSH as the sole source of sulphur. The growth was monitored by recording absorbance at 600 nm at regular intervals.

Fig 3.10 Growth phenotype analysis to check the effect of over-expression of DUG2 and DUG3 in *S. cerevisiae* ABC733 (WT). The transformants were grown in SD-ura supplemented either with GSH or methionine as sole source of sulphur and growth was monitored by recording the absorbance at 600 nm at regular intervals.
transferase (PRPP) and asparagine synthase (Isupov et al., 1996). This family of proteins is a member of N-terminal nucleophile (Ntn) hydrolase superfamily (formerly F-type) since the precursor proteins are usually synthesized with an N-terminal 11-residue propeptide that is cleaved off autocatalytically. The precursor protein is thus referred to as a peptidase (the mature protein, in contrast, lacks peptide bond hydrolytic activity). The cysteine-12 residue, the nucleophile in the cleavage of the propeptide becomes the N-terminal active site residue (N-terminal catalytic cysteine) (Tso et al. 1982). In some members of this family, which appears to include Dug2p, the catalytic cysteine immediately follows the methionine start codon (Cys2), and is activated by cleavage of the N-terminal methionine by a methionine amino peptidase (MAP). As these precursors lack the autocatalytic peptidase activity they are referred to as non-peptidase homologues of the Ntn-hydrolase superfamily in MEROPS database (Rawlings et al., 2006).

To determine if all key residues found in this class of protein were conserved in YNL191w; multiple sequence alignment were carried out with well-characterized Gln-AT-I enzymes. The Dug3p (Ynl191wp) sequence was aligned with S. cerevisiae GFA1 (GFA1, Glucosamine-fructose 6-phosphate aminotransferase), S. cerevisiae ADE4 (ADE4/PRPPAT, Glutamine phosphoribosyl pyrophosphate amidotransferase) and E.coli PurF (amidophosphoribosyltransferase) (Fig 3.8c).

Studies on the crystal structure of E.coli PurF (PRPPAT/ADE4) enzyme had led to the identification of key residues, namely C2, R27, N102, G103, D128 and R74 (amino acid positions are referred by considering methionine of the pre-polypeptide as 1st amino acid instead of cysteine), which played important role in Gln-AT (class II) activity (Muchmore et al. 1998). Extensive studies had also been carried out earlier to understand the mechanism of catalysis by Glucose-6-phosphate synthase (GlcN-6-p-synthase/Glms/GFAT/GFA1) from various organisms. The crystal structure studies of GlcN-6-p synthase have led to the identification of various key residues namely, C2, R27, N100, G101, D124, R74, T77, H78, and W75 (amino acid positions are referred by considering methionine of the pre-polypeptide as 1st amino acid instead of cysteine), which played important role in Gln-AT (class II) activity (Milewski 2002).

Sequence alignment study of Dug3p (Ynl191wp) along with ScADE4, ScGFA1, and EcPurf revealed that although there is very little similarity across the entire stretch of
Fig 3.8c Multiple sequence alignment of GlnAT-II (Glutamine Amidotransferase-ClassII) domain of Ynll191wp (Dug3p) protein of Saccharomyces cerevisiae (ScYNL191w, 1-240 amino acids) along with its homologues; Gfal protein of Saccharomyces cerevisiae (ScGFA1, 1-210 amino acids), Ade4 protein of Saccharomyces cerevisiae (ScADE4, 1-240 amino acids), PurF protein of E.coli (EcPurF, 1-240 amino acids). The multiple sequence alignment was carried out using CLUSTAL X (Thomson et al., 1997) and was reformatted using CLOURE-C (Kohli and Bachhawat, 2003). The conserved residues of Gln-AT-II domain are highlighted by black arrows.
Gn_At domain, the above-mentioned key residues of E.coli PurF and Glc-6-p synthase are highly conserved across the four proteins. By comparing the sequence of all the four proteins it was observed that C2, R40, N121, G122, D151, R96 and T99 residues of Dug3p (Ynl191wp) came under the same highly conserved class of amino acid residues which are believed to play crucial role in Gn_AT activity of known enzymes (Fig 3.8c). The sequence analysis further suggests that Dug3p (Ynl191wp) may indeed be a Gln-AT class II enzyme. The mechanism of Gn_AT activity may be similar to any other Gn_AT Class II enzyme. C2 is possibly the active site residue that may function as a nucleophile and attack the carboxyamide of glutamate and generate a γ-glutamyl thioester intermediate. N121 and G122 thereafter may play crucial role in the formation of oxyanion hole for the stability of the cysteiny1-glutamine tetrahedral intermediate. D151 and R96 side chains may provide the specificity for the α-amino and α-carboxyl groups of glutamine (Muchmore et al. 1998).

3.8 Dug1p, Dug2p and Dug3p localize in the cytoplasm.

To obtain the information on the possible location of Dug1p, Dug2p and Dug3p proteins, we carried out studies using yeast strains in met15Δ background where the Dug1p, Dug2p and Dug3p proteins were previously tagged with GFP at C-terminal ends (Huh et. al, 2003) and were procured from Invitrogen. The tagged proteins did not affect the function of the Dug1p and Dug3p to any significant extent as seen by growth in GSH as a sulphur source, however a partial defect was seen in the Dug2p-GFP fusion strain. This strain was grown in methionine followed by re-inoculation in GSH containing broth. Using confocal microscopy, all the three proteins were observed to have strong and uniform GFP fluorescence throughout the cytoplasm, suggesting that all the three proteins were localized to the cytoplasm (Fig 3.11).

3.9 The alternative pathway of GSH degradation is fungal specific.

GSH is present in almost all eukaryotes and some prokaryotes and metabolized by the γ-glutamyl cycle in these systems. With the cloning of the DUG1, DUG2 and DUG3 genes it was now possible to determine if this alternative pathway was also present in all
Fig 3.11. Localization of Dug proteins by DUG-GFP visualization using Confocal microscopy.
organisms. We searched the databases for the presence of homologues of Duglp, Dug2p and Dug3p in other organisms.

Duglp homologues are present in bacteria, fungi and eukaryotes including plants and mammals e.g human. The human homologues include cytosolic non-specific dipeptidase (CNDP) and carnosinase etc.

Searching for homologues of Dug2p, we observed that only some fungi like Kluyveromyces lactis, Candida albicans, Candida glabrata, Magnaporthe grisea, Neurospora crassa and Aspergillus nidulans showed the presence of putative Dug2p homologous protein. These homologues were similar in length, containing homology to both N-terminus WD40 domain and C-terminus M20A peptidase domain of Dug2p. Surprisingly, no homologues were found in S. pombe. Higher eukaryotes showed the presence of protein having homology with only the M20A peptidase domain of Dug2p, suggesting that orthologues were absent in these organisms.

Similar searches with Dug3p showed that the presence of Dug3p homologues is restricted to fungal species Kluyveromyces lactis, Ashbya gossypii, Candida glabrata, Debaryomyces hansenii, Yarrowia lipolytica Candida albicans. In bacteria, only N-terminus Gln_AT classII domain of Dug3p picked up homologous proteins. These homologues did not appear to be true orthologues of Dug3p, as they did not align across the full length of the protein.

This analysis revealed that probably only some fungi have evolved a specialized mechanism of degradation of GSH independent of γ-GT, involving Duglp, Dug2p and Dug3p, but this is absent in bacteria, plants and higher eukaryotes.

3.10 Conclusion

In the present chapter, we described the results of our efforts to identify and characterize the DUG2 and DUG3 genes.

The yeast genomic library complementation assay of dug2 has revealed that the YBR281c ORF corresponds to DUG2 gene. The ORF was cloned and shown to functionally complement met15aybr281cA (dug2 mutant) growth defect on medium containing GSH as a sole sulphur source. Sequence analyses have shown that Dug2p harbors two domains: an N-terminal (1-422aa) putative WD40 repeat domain and C-
terminal (441-878aa) putative M20A peptidase domain that was homologous to Dug1p. The key catalysis and zinc binding residues found in this family were conserved in the Dug1p and Dug2p proteins.

Similar library complementation studies with *dug3* have revealed that *YNL191w* corresponds to *DUG3* gene. The ORF was cloned and shown to functionally complement the *met15Δynl191wΔ* (dug3 mutant) growth defect on medium containing GSH as a sole sulphur source. Sequence analyses of Dug3p have shown that the N-terminus (45-254aa) contains a putative Glutamine amidotransferase class II domain with the key catalytically important residues conserved.

Localization studies using Dug1pGFP, Dug2pGFP and Dug3GFP constructs have revealed that all the Dug proteins are localized in cytoplasm.

The search for homologous proteins have shown the presence of true orthologues of Dug2p and Dug3p is restricted to certain yeast/fungi. It thus appears that some fungi have evolved an alternative pathway of degrading GSH, although the reason for the alternative pathway being restricted to these groups is not yet fully understood.