Additional cassettes for epitope and fluorescent fusion proteins in *Candida albicans*

Maryam Gerami-Nejad,1 Keely Dulmage1 and Judith Berman2*

1Department of Genetics, Cell Biology and Development, University of Minnesota, MN, USA
2Department of Microbiology, University of Minnesota, MN, USA

*Correspondence to: Judith Berman, 6-160 Jackson Hall, 321 Church Street SE, Minneapolis, MN 55455, USA. E-mail: jberman@umn.edu

Received: 13 March 2009
Accepted: 26 April 2009

Abstract

Epitope tags that confer specific properties, including affinity for resins or antibodies or detection by fluorescence microscopy, are highly useful for biochemical and cell biological investigations. In *Candida albicans* and several other related yeasts, the CUG codon specifies serine instead of leucine, requiring that molecular tools be customized for use in this important human fungal pathogen. Here we report the construction of a set of plasmids containing 13-Myc, 3HA, GST, V5 or His9 epitope cassettes that facilitate PCR-mediated construction of epitope-tagged proteins. Common primer sets amplify the different tags with two different selectable markers. In addition, we report construction of a codon-optimized *Discosoma* red fluorescent protein (DsRFP) gene. Like mCherryRFP, this DsRFP signal is detectable in transformants at the colony level and is useful in double-labelling experiments with green fluorescent protein (GFP). Finally, we describe a construct that directs PCR-mediated two-step insertion of GFP internal to a coding sequence, which facilitates tagging of secreted proteins, including GPI-anchor cell wall proteins that require endogenous N- and C-termini for function. These reagents expand the repertoire of molecular tools available for working with *C. albicans* and other members of the CUG clade of pathogenic yeasts. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: *Candida albicans*; fluorescent protein; epitope; tagging

Introduction

*Candida albicans* is a major cause of serious mucosal and systemic fungal infections, especially in immunocompromised patients (Pfaller and Diekema, 2007). The use of PCR-mediated transformation has enabled rapid genetic modification of this obligate diploid yeast (Wilson *et al*., 1999). Epitope tags are important tools for performing biochemical analyses, such as affinity purification and quantification of protein levels. Especially useful are a series of cassettes that employ a defined set of common primers, so that different strains carrying epitopes inserted at the same gene can be constructed with a single pair of oligonucleotides (Bahler *et al*., 1998; Krawchuk and Wahls, 1999; Lavoie *et al*., 2008; Longtine *et al*., 1998b; Sung *et al*., 2008). The use of cassettes for tagging proteins with green, cyan and yellow fluorescent proteins (GFP, CFP and YFP) at C-termini has also facilitated the analysis of protein localization and expression within living *C. albicans* cells (Gerami-Nejad *et al*., 2001; Schaub *et al*., 2006). Tagging at the N-terminus has usually involved replacement of the native promoter with a constitutive or regulatable promoter cassette (Gerami-Nejad *et al*., 2004). Double FP labelling experiments are generally performed using YFP together with CFP. However, the development of *Discosoma* red fluorescent protein (Bevis and Glick, 2002) has facilitated the performance of double labelling experiments using GFP and RFP.
In \textit{C. albicans}, the alternative translation of the CUG codon for serine instead of leucine obviates the use of most heterologous vectors with epitope tags. A single \textit{C. albicans} tagging vector that expresses a histidine 6-mer (His6) C-terminal to the protein of interest has been available for a few years (Schaub et al., 2006); a new set of three epitope tagging vectors (expressing HA, TAP or MYC) was recently described (Lavoie et al., 2008). Here we report a series of plasmids that use a shared set of primers for the amplification of five different epitope tags. In addition, we describe a plasmid that permits insertion of GFP, in frame, within an open reading frame. Together, this set of epitope-tagging plasmids expands the repertoire of biochemical and cell biological experiments that can be performed in \textit{C. albicans}.

### Materials and methods

#### Strains, growth conditions and DNA methods

All strains used in this study (Table 2) were derived from BWP17 (Wilson et al., 1999) and were grown at 30°C in rich (YPAD) medium, synthetic complete medium (SC), or synthetic complete medium lacking specific nutrients (SM; Sherman, 1991). \textit{Escherichia coli} strain XLI-blue (Stratagene, La Jolla, CA, USA), growth conditions, DNA manipulations and primer design and synthesis were essentially as described previously (Gerami-Nejad et al., 2004). Transformants were selected on SM medium lacking uridine or histidine and were screened by PCR using oligonucleotide primers listed in Table S1.

#### Plasmid constructions

To generate pMG1874 (HA–URA3) the BamH1/HindIII fragment of pGFP–URA3 (Gerami-Nejad et al., 2001) containing URA3 was ligated to the BglII–HindIII 3HA fragment of pFA6a–3HA–KanMX6 (Longtine et al., 1998a).

### Table 1. PCR primers used to amplify transformation cassettes

<table>
<thead>
<tr>
<th>Primer</th>
<th>Epitope</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>HA, MYC, GST</td>
<td>5′-(gene-specific sequence) GGTGGTGGTGCGGATCCCAGGGTTAATTAA-3′</td>
</tr>
<tr>
<td>F2</td>
<td>V5</td>
<td>5′-(gene-specific sequence) GGTGGTGGTGACCTATCCCTAACCTCCTC-3′</td>
</tr>
<tr>
<td>F3</td>
<td>His9</td>
<td>5′-(gene-specific sequence) GGTGGTGGTGGCAGCACTGAGTATGNTTAT-3′</td>
</tr>
<tr>
<td>F4</td>
<td>RFP</td>
<td>5′-(gene-specific sequence) GGTGGTGGTGGTGGACAGATGTAAGTAATG-3′</td>
</tr>
<tr>
<td>F5</td>
<td>GFP</td>
<td>5′-(gene-specific sequence) TCTAGTACGAATAGTATCCATTG-3′</td>
</tr>
<tr>
<td>F6</td>
<td>M Cherry</td>
<td>5′-(gene-specific sequence) GGTGGTGGTGGTGGAGATGTAAGTAATG-3′</td>
</tr>
<tr>
<td>R1</td>
<td>URA3</td>
<td>5′-(gene-specific sequence) TCTAGTACGAATAGTATCCATTG-3′</td>
</tr>
<tr>
<td>R2</td>
<td>His1</td>
<td>5′-(gene-specific sequence) GAGTGGTGGTGGTGGACAGATGTAAGTAATG-3′</td>
</tr>
<tr>
<td>R3</td>
<td>GFP</td>
<td>5′-(gene-specific sequence) TCTAGTACGAATAGTATCCATTG-3′</td>
</tr>
</tbody>
</table>

### Table 2. Strains used in this study

<table>
<thead>
<tr>
<th>Strain ID</th>
<th>Relevant genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>YJB9949 (BWP17)</td>
<td>URA3Δ::IMM434/URA3Δ::IMM434 His1::HisG His1::HisG Arg4::HisG; Arg4::HisG</td>
<td>Wilson, 1999</td>
</tr>
<tr>
<td>YJB6751</td>
<td>BWP17 CDC3/CDC3::HA-URA3</td>
<td>This study</td>
</tr>
<tr>
<td>YJB6928</td>
<td>BWP17 CDC3/CDC3::MYC-URA3</td>
<td>This study</td>
</tr>
<tr>
<td>YJB8223</td>
<td>BWP17 CDC3/CDC3::GFP</td>
<td>This study</td>
</tr>
<tr>
<td>YJB8250</td>
<td>BWP17 HWP1::GFP</td>
<td>This study</td>
</tr>
<tr>
<td>YJB8600</td>
<td>BWP17 CLB2/CLB2::URA3, CLN1/CLN1::MYC, His1</td>
<td>This study</td>
</tr>
<tr>
<td>YJB8622</td>
<td>BWP17 CLN1/CDC3:V5-URA3</td>
<td>This study</td>
</tr>
<tr>
<td>YJB9139</td>
<td>BWP17 MLC1/MLC1::YFP-URA3, CDC3/CDC3::HA-HIS1</td>
<td>This study</td>
</tr>
<tr>
<td>YJB9452</td>
<td>BWP17 ENO1/ENO1::RFP-URA</td>
<td>This study</td>
</tr>
<tr>
<td>YJB11257</td>
<td>BWP17 ENO1::mCherry-URA3</td>
<td>This study</td>
</tr>
</tbody>
</table>
pMG1921 (HA–HIS1), the HIS1 marker was isolated as a PvuI fragment from pGFP–HIS1 (Gerami-Nejad et al., 2001) and ligated to PvuI-digested pMG1874, resulting in HIS1 replacing the URA3 gene. Analogous approaches were used to insert the 13-MYC as a BglII and HindIII fragment from pFA6a–13myc–KanMX6 (Longtine et al., 1998a) into the BamHI–HindIII fragment of pGFP–URA3 to yield pMG1905 (13-MYC–URA3). To construct pMG2093 (13MYC–HIS1), the HIS1 marker was isolated as a PvuI fragment from pGFP–HIS1 and was ligated into PvuI-digested pMG1905.

To construct pMG2083 (GST–URA3) the BglII–HindIII GST fragment from pFA6a–GST–TRP was ligated to BamHI–HindIII-digested pGFP–URA3. To construct pMG2090 (V5–URA3), the V5 epitope was amplified using forward primer 1490 (designed to include a HindIII site; Table S2, italicized), reverse primer 1491 (designed to include a BamH1 site; Table 1) and ptracer–EF/V5–HIS (Invitrogen, Carlsbad, CA, USA). The resulting PCR product was digested with forward primer 1490 (including a HindIII site; Table S2, italicized), reverse primer 1491 (designed to include a BamH1 site; Table 1) and ptracer–EF/V5–HIS (Invitrogen, Carlsbad, CA, USA). The resulting PCR product was digested with BamH1–HindIII and ligated into BamH1- and HindIII-digested pGFP–URA3. pMG2051 (HIS9–URA3) was constructed by annealing primer 1286, which includes nine GTGs and a BglII site (Table S2), digesting the product with BglII–HindIII and then ligating it to BamH1- and HindIII-digested pGFP–URA3.

A synthetic CaRFP gene was produced at Celtek Genes (Nashville, TN, USA) based on the coding sequence of monomeric DsRed (Bevis and Glick, 2002), which was codon-optimized for C. albicans by changing all CUG codons to TTG, which is the most frequently used Candida codon (designed to add a C-terminal epitope tag to TTG). The complete sequence is provided at http://cbs.umn.edu/labs/berman/epitope.htm. The 675 nt codon-optimized DsRed sequence was inserted into the pUC57 that had been digested with BamH1 and PstI site, to generate pMG2162. Plasmid pRFP–URA (pMG2169) was constructed by digesting pMG2162 with BamH1 and SalI, and ligating the RFP coding sequence adjacent to the URA3 fragment in pGFP–URA3 that had been digested with BamH1 and XhoI.

The GFP–URA3–GFP (GUG) cassette in pMG2082 was constructed by amplifying 456 bp of the C-terminal portion of GFP coding sequence using primers 1402 and 1403 (Table S2) and pGFP–URA3 as template. The PCR product was digested with XbaI and cloned into the XbaI site in pYPB1–ADHp1 (Bertram et al., 1996) to yield plasmid pMG2079, which contains URA3 fused in its 3′ UTR to a truncated GFP lacking the N-terminal 87 codons. To construct GFP–URA3–GFP, the N-terminal 459 bp of GFP coding sequence was amplified from pGFP–URA3 using primers 1400 and 1401 (Table S2). The PCR product was digested with BglII and XhoI and cloned into pMG2079 digested with XhoI and BglII.

To generate pMG2254 (mCherryRFP), plasmid mRFP (yEmRFP; Keppler-Ross et al., 2008), kindly provided by Dr Neta Dean, was amplified using primers 3345 and 3346 (Table S2). The PCR product was cut with both BamH1 and EcoRV and subsequently ligated to BamH1 and EcoRV-digested pGFP–URA3.

Requests for plasmids should be sent to jberman@umn.edu. Electronic sequence files of the tagging cassettes are available at http://www.cbs.umn.edu/labs/berman/fluorescent.shtml under the ‘epitope tags’ link.

Transformation, identification of integration events and protein detection methods

PCR was performed using one of the plasmids described above as template, and sequences homologous to the appropriate target-gene integration site sequence. Primer pairs were designed as indicated in Figure 1 and Table 1. To tag CLB2 and CLN1 with the V5 epitope, CLB2 or CLN1-specific homologous sequence was added to the universal primer sequences described in Table 1. PCR reactions were performed as described previously (Gerami-Nejad et al., 2001). The products from 10 such PCR reactions were pooled, precipitated with ethanol, resuspended in 30 µl water and used to transform C. albicans strain BWP17 (Wilson et al., 1999). Transformants were selected by plating on the appropriate selective media. To identify transformants in which the cassette had correctly integrated into the target gene sequence, genomic DNA was prepared and used as the template in PCR reactions, using one primer that annealed within the transformation module and a second primer that annealed to the target gene locus outside the altered regions. CDC3 was tagged at the C-terminus with...
**HA, MYC, GST** or **HIS9, CDC3**-specific homologous sequence (70 nt) was added to the universal primer sequences described in Table 1.

Immunoblot detection of fusion proteins was performed essentially as described previously (Gerami-Nejad et al., 2004), except that starter cultures were grown for ~12 h in SDC medium, diluted to OD_{600} = 0.8 and grown until cultures reached OD_{600} = 1.0. The cultures were collected by centrifugation and stored at −80°C until analysis. Cells were thawed and resuspended in 50 µl Thorner buffer and heated at 100°C for 3 min, and then vortexed with an equal volume of glass beads for 4 min at 4°C. Prior to analysis on 10% SDS–polyacrylamide gels, samples were denatured at 100°C for 4 min, vortexed for 2 min and lysis was confirmed by microscopy.

Proteins were transferred to a PVDF membrane, blocked for 30 min with 2% skimmed milk in TBST (20 mM Tris, pH 7.6, 137 mM NaCl, 0.1% Tween-20). Membranes were then incubated for 1 h with HRP-conjugated mouse anti-HA antibody (12CAS, 1:1000 in 0.2% skimmed milk in TBST), HRP-conjugated rabbit anti-GST Ab (Z-5: sc-459, 1:1000 in 0.2% skimmed milk in TBST), mouse Anti-c-Myc Ab (1:5000 in 0.2% skimmed milk in TBST) or mouse anti-V5 Ab (1:500 in 0.2% skimmed milk in TBST). Blots probed with non-HRP-conjugated antibodies were then washed in TBST and incubated for 1 h with horseradish peroxidase-conjugated goat anti-Mouse IgG (Santa Cruz Biotechnology, Santa Cruz, CA, USA), diluted 1:10 000 in 0.2% skimmed milk in TBST. The GST blot was washed in water to remove background (Wu et al., 2002), all other Westerns were washed in 0.67 × TBST, and proteins were then detected using Supersignal Femto Chemiluminescent reagent (Pierce, Rockford, IL, USA).

The DsRFP and mCherry variants (Shaner et al., 2008), were inserted at the **ENO1** locus using primers (Table S3) that included 70 bp **ENO1**-homologous sequence plus the universal primer sequences described in Table 1.

Detection of DsRFP and of mCherryRFP fluorescence was performed on 2 day-old colonies using a Nikon SMZ1500 microscope equipped with a Chroma 59-022 filter. Six fluorescent and four non-fluorescent colonies were analysed by PCR using primers 716 and 1363 (Table S3).

### Results and discussion

To produce cassettes for epitope tagging the C-terminus of **C. albicans** genes, we modified GFP-tagging vectors pGFP–URA3 and pGFP–
Cassettes for tagging *C. albicans*

**Figure 2.** Western blots and antibody detection of epitope-tagged strains. Lanes 1, 4, 6 and 9 correspond to the parent strain (BWP17). Lanes 2 and 3, HA-tagged strains YJB6751 (CDC3–HA*URA3*) and YJB9139 (CDC3–HA*HIS1*); lane 5, 6928 (CDC3–MYC); lanes 7 and 8, V5-tagged strains YJB8600 (Clb2–V5) and YJB8622 (Cln1–V5); lane 10, GST-tagged strain YJB8223 (CDC3–GST). Expected sizes of fusion protein products are as follows: 55 kDa, CDC3–HA; 69 kDa, CDC3–MYC; 81 kDa, CLB2–V5; 60 kDa, CLN1–V5; 75 kDa, CDC3–GST

HIS1 by replacing the *GFP* gene with sequences encoding different epitope tags. The plasmids retained the transcription termination sequence from *ADH1* and the relevant selectable marker (either *URA3* or *HIS1*). In general, to add a C-terminal epitope tag to a specific *C. albicans* gene, the cassette of interest was amplified using gene-specific forward and reverse primer pairs that included ~70 bp of homology with the gene of interest, followed by the appropriate forward (F1–F6) and reverse (R1–R3) primer sequences (Figure 1, Table 1) with homology to the tagging cassette. A useful feature of the cassettes is that the HA, MYC and GST epitope tags are amplified using the same forward primer (F1) and the reverse primer corresponding to the selectable marker used (R1 or R2). Similarly, the V5, His9 and DsRFP epitopes are all amplified using the same reverse primer (R1).

To test the usefulness of the cassettes, we integrated the HA, MYC, GST and His9 epitopes at the C-terminus of *CDC3* and the V5 epitope at the C-terminus of *CLB2* and *CLN1* in laboratory strain BWP17. *CDC3* transformants were analysed by PCR for the presence of fragments, using primer 335 from within the cassette together with *CDC3*-specific primer 630 (Table S1). PCR products of the expected size were sequenced to confirm that the insertion was in-frame. For each epitope tagging transformation, PCR products of the expected sizes (Table S4) were detected in at least one and as many as three of the eight screened transformants (12.5–38% bona fide transformants).

To investigate whether the epitope tags were expressed at detectable levels, we separated proteins from *C. albicans* crude lysates by electrophoresis, transferred them to a PVDF membrane and probed them with commercially available antibodies that recognize the individual epitopes (see Materials and methods). Strains expressing fusion proteins all produced a unique prominent band (Figure 2). The HA, V5 and GST fusion proteins ran at the expected electrophoretic mobility, while the Myc fusion protein ran larger than the predicted protein fusion size. This is consistent with the mobility of other Myc-tagged proteins (data not shown). Anti-His6 antibody was less specific (data not shown). Because each epitope is detected with a different antibody, it is difficult to perform a direct comparison of the signals from each epitope.

A recent report of a codon-optimized mCherryRFP found that the fluorescent protein was visually detectable as a purple pigment in colonies of *S. cerevisiae*, but not in *C. albicans* (Keppler-Ross et al., 2008). We compared the appearance of codon-optimized DsRFP to that of mCherryRFP in *C. albicans* colonies. While colony colour was not evident by visual inspection, the ENO1–DsRFP fusion protein was readily detectable in 2 day-old colonies viewed with a Nikon SMZ1500 microscope equipped with a Chroma 59022 (Texas red) filter (Figure 3A). Similar levels of fluorescence were detectable with codon-optimized mCherryRFP after a single day. The brighter appearance of mCherryRFP relative to DsRFP is consistent with observations in other organisms (Shaner et al., 2005). However, the lower excitation and emission wavelengths of DsRFP relative to mCherryRFP are more compatible with some flow cytometry filter sets. Importantly, transformant colonies can
Figure 3. (A) Fluorescence is readily detected at the colony level after 2 days of growth. (B) Immunofluorescence micrographs of ENO1–RFP and ENO1–mCherry expressed from its native promoter. Construction of strains expressing ENO1–RFP is described in Materials and methods. Analysis of fluorescence intensities showed $5 \times$ stronger fluorescence of the strain tagged with ENO1–mCherry than with ENO1–RFP.

be directly screened for both DsRFP and mCherryRFP expression using fluorescence microscopy (Figure 3B). PCR analysis of six colonies that exhibited DsRFP fluorescence revealed that all of them contained the expected junction fragments spanning ENO1 and the DsRFP sequence. As a control, four colonies that did not exhibit DsRFP fluorescence in colonies from the same transformation experiment did not contain junction fragments diagnostic of bona fide insertion of the DsRFP into ENO1. Thus, detection of DsRFP expression at the colony level facilitates the rapid and accurate identification of potential transformants, with low levels of false-positive and false-negative results.

Finally, we generated vector pGUG, in which amplification of a GFP–URA3–GFP cassette yields proteins that are tagged in-frame, within the ORF, via a two-step selection scheme (Figure 4A). A similar approach has been used to epitope tag S. cerevisiae proteins (Schneider et al., 1995). First, a cassette carrying incomplete copies of GFP flanking the counterselectable URA3 marker is inserted at the locus to be tagged using homology to the insertion site (Table S3) and primers F5 and R3 (Table 1). Insertion of the cassette is facilitated by selection for URA3. Subsequently, growth on 5-FOA selects for cells in which recombination between the GFP flanking DNA results in deletion of the interrupting URA3 sequence, while it restores a functional GFP gene that is in-frame with the N-terminal and C-terminal coding sequences of the genes being tagged (Figure 4A).

HWP1 is a cell wall protein in which the 20 N-terminal amino acids are necessary for entry into the secretory pathway (Mao et al., 2003) and the C-terminus is modified by a GPI-anchor that directs attachment of proteins to the outer surface of the plasma membrane and/or to components of the cell wall (Mao et al., 2008). Expression of C-terminally tagged HWP1–GFP results in cytoplasmic localization of the GFP because the protein cannot be appropriately localized (Mao et al., 2003). Guided by the work of Mao et al. (2003, 2008), we inserted pGUG between Gln54 and Thr588, such that the remaining N- and C-terminal signals should be sufficient to localize GFP to the plasma membrane and/or cell wall. As expected, replacement of the middle 534 codons of HWP1 with pGUG, followed by selection for loss of the URA3 sequence on 5-FOA, resulted in localization of Hwp1–GFPp, like native Hwp1p, to the cell periphery (Figure 4B). Thus, the pGUG construct described here greatly facilitates the tagging
Cassettes for tagging *C. albicans*

**Figure 4.** (A) Growth of HWP1–GF–URA3–FP colonies on FOA results in a restoration of the intact GFP tag. (B) Expression of HWP1 internally tagged with GFP results in localization of GFP1p, like native HWP1p, to the cell periphery of proteins that require both N- and C-terminal sequences for proper localization and obviates the multiple steps required to clone GFP within the ORF of interest prior to transformation (Mao *et al*., 2003). Furthermore, unlike cloning approaches used previously, use of pGUG enabled construction of a fusion protein expressed from its native promoter.

In summary, epitope tags and fluorescent protein fusions have become indispensable components of the molecular toolbox for studying the biochemistry and cell biology of all model organisms. Because of its unusual codon usage, most generic epitope tags will not function in *C. albicans*, an important human pathogen. The new cassettes constructed here extend the number of epitope tags available for use in biochemical studies of *C. albicans*. In addition, we produced a synthetic DsRFP cassette useful for fluorescent tagging and studying proteins alone or in double-labelling experiments (in combination with green or yellow fluorescent proteins). Furthermore, like codon-optimized mCherryRFP, codon-optimized DsRFP can be exploited to visually detect transformants 2 days after plating, using a simple fluorescence microscopy screen.

**Acknowledgements**

We thank Mark McClellan and Aaron Christensen-Quick for web site creation and technical assistance. We thank Neta Dean for providing the mCherry GFP plasmid. This work was supported by funding from NIH, Grant Nos R01 AI0624273 and R01 AI075096, to J.B.

**Supporting information**

Supporting information may be found in the online version of this article.

**References**


