The unique C-terminal region of Hap3 is required for methanol-regulated gene expression in the methylotrophic yeast *Candida boidinii*

Saori Oda,[1,†] Hiroya Yurimoto,[1] Nobuhisa Nitta,[1] and Yasuyoshi Sakai[1,2,3,*]

[1]*Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Kitashirakawa-Oiwake, Sakyo-ku, Kyoto 606-8502, Japan*

[2]*Research Unit for Physiological Chemistry, the Center for the Promotion of Interdisciplinary Education and Research, Kyoto University, Japan*

Running title: CbHap3p in methanol-regulated gene expression

*Correspondence: Yasuyoshi Sakai, Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Kitashirakawa-Oiwake, Sakyo-ku, Kyoto 606-8502, Japan. Tel.: +81 75 753 6385; fax: +81 75 753 6454; e-mail: ysakai@kais.kyoto-u.ac.jp

†Present address: Department of Chemistry and Biotechnology, Graduate School of Engineering, Tottori University, 4-101 Koyamacho-Minami, Tottori 680-8552, Japan.

Abstract

The Hap complex of the methylotrophic yeast *Candida boidinii* was found to be required for methanol-regulated gene expression. In this study, we performed functional characterization of CbHap3p, one of the Hap complex components in *C. boidinii*. Sequence alignment of Hap3 proteins revealed the presence of a unique extended C-terminal region, which is not present in Hap3p from *Saccharomyces cerevisiae* (ScHap3p), but is found in Hap3ps of methylotrophic yeasts. Deletion of the C-terminal region of CbHap3p (Δ256-292 or Δ107-237) diminished activation of methanol-regulated genes and abolished the ability to grow on methanol, but did not affect nuclear localization or DNA-binding ability. On the other hand, deletion of the N-terminal region of CbHap3p (Δ1-20) led to not only a growth defect on methanol and a decreased level of methanol-regulated gene expression, but also impaired nuclear localization and binding to methanol-regulated gene promoters. We also revealed that CbHap3p could complement the growth defect of the *Schap3Δ* strain on glycerol, although ScHap3p could not complement the growth defect of a *Cbhap3Δ* strain on methanol. We conclude that the unique C-terminal region of CbHap3p contributes to maximum activation of methanol-regulated genes, while the N-terminal region is required for nuclear localization and binding to DNA.

Introduction

Methylotrophic yeasts, such as *Hansenula polymorpha*, *Pichia pastoris*, and *Candida boidinii*, are unique yeasts that can utilize methanol as a sole carbon and energy source. The promoters of genes encoding methanol-metabolizing enzymes, including alcohol oxidase (AOD), dihydroxyacetone synthase (DAS), glutathione-dependent formaldehyde dehydrogenase (FLD), and formate dehydrogenase (FDH), are highly induced by methanol, and tightly regulated by the presence of alternative carbon sources. Therefore, these promoters have been used for industrial protein production with methylotrophic yeasts as hosts (Gellissen, 2000; Daly & Hearn, 2005; Yurimoto *et al.*, 2011; Vogl & Glieder, 2012). Methanol-regulated gene expression is presumed to be conducted by three distinct pathways. Methanol-regulated genes are completely repressed in the presence of glucose, which requires CbMig1p (glucose repression). Exhaustion of glucose releases glucose-repression, resulting in activation of methanol-regulated genes by CbTrm2p, which does not require methanol for gene activation (derepression). In addition, the presence of methanol induces maximum activation of methanol-regulated genes via CbTrm1p (methanol induction; methanol-specific induction) (Hartner & Glieder, 2006; Sasano *et al.*, 2008; Yurimoto, 2009; Sasano *et al.*, 2010; Zhai *et al.*, 2012).

In a previous study, we identified a multimeric transcription factor, the CbHap complex, which is involved in methanol-regulated gene expression, specifically methanol induction (Oda *et al.*, 2015). The Hap complex is highly conserved among all eukaryotes, from yeasts to humans (Ramil *et al.*, 2000; McNabb & Pinto, 2005; Sybirna *et al.*, 2005; Singh *et al.*, 2011; Ridenour & Bluhm, 2014). It consists of a stable heterotrimer (Hap2p/3p/5p), which binds to a CCAAT consensus sequence and the activator protein Hap4p. In *Saccharomyces cerevisiae*, while ScHap2p, ScHap3p and ScHap5p are constitutively expressed, ScHap4p is regulated in a carbon source-dependent manner at the transcriptional level; it is repressed in the presence

of glucose, and induced by exhaustion of glucose or in the presence of non-fermentable carbon sources such as ethanol or glycerol. The ScHap2p/3p/4p/5p complex activates genes involved in respiratory metabolism and mitochondria biogenesis, and is indispensable for respiratory growth on non-fermentable carbon sources. Although CbHap2p/3p/5p proteins were found to be necessary for maximum activation of methanol-regulated genes and growth on methanol in *C. boidinii,* the CbHap complex was not necessary for growth on non-fermentable carbon sources or for derepression (Oda *et al.*, 2015).

In this study, we further characterized CbHap3p, which was found to contain a unique C-terminal region specific to methylotrophic yeasts, but not present in *S. cerevisiae*. We revealed that the N-terminal and C-terminal regions of CbHap3p have distinct roles during methanol induction in *C. boidinii.*

Materials and methods

Strains, media, and cultivation conditions

The haploid strain *C. boidinii* S2 was used as the wild-type strain (Tani *et al.*, 1985). *C. boidinii* strain TK62 (*ura3*) was used as a host for transformation (Sakai *et al.*, 1991). *S. cerevisiae* strain BY4741 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) was used as a host for transformation (Brachmann *et al.*, 1998).

C. boidinii strains were grown on either YPD medium (2% glucose, 2% Bacto peptone, 1% Bacto yeast extract) or YNB medium (0.17% yeast nitrogen base without amino acids and ammonium sulfate, 0.5% ammonium sulfate). One of the following was used as the carbon source in YNB medium: 2% (w/v) glucose (YND) and 0.7% (v/v) methanol (YNM). *S. cerevisiae* strains were grown on either YPD medium or YPGly medium (2% glycerol, 2% Bacto peptone, 1% Bacto yeast extract). The initial pH of the medium was adjusted to 6.0. All yeasts were cultivated aerobically at 28°C.

Construction of strains expressing domain-deleted protein of CbHap3p-YFP proteins

C. boidinii strains expressing domain-deleted protein of CbHap3p-YFP proteins were constructed as follows. PCR was performed with the primer pairs listed in Table S1 using pCbHAP3-YFP as a template. Amplified fragments were self-ligated to produce pCbHAP3-YFPΔ256-292, pCbHAP3-YFPΔ107-224, pCbHAP3-YFPΔ107-237, pCbHAP3-YFPΔ107-241, pCbHAP3-YFPΔ107-256, pCbHAP3-YFPΔ1-15, pCbHAP3-YFPΔ1-20, pCbHAP3-YFPΔ1-25, pCbHAP3-YFPΔ101-224, pCbHAP3-YFPΔ98-224, pCbHAP3-YFPΔ86-224, respectively. The constructed plasmids were linearized with EcoT22I and used to transform *C. boidinii* strain *Cbhap3Δura3*. The plasmids were integrated into the *ura3* locus in the genome of *C. boidinii*.

106

107 **Fluorescence microscopy and nuclear staining**

108 *CbHAP3-YFP*/*Cbhap3Δ* cells grown to mid-exponential phase in YND, YNE, YNG,
109 YNM or YNO medium were harvested, washed once and fixed in 1 ml of 70% ethanol for 30
110 min at room temperature. Fixed cells were then washed twice, resuspended in 150 µl of
111 sterilized water, and stained with 150 µl of 0.125 µg/ml DAPI
112 (4',6'-diamidino-2-phenylindole) solution. After 10 min of incubation, fluorescence was
113 observed using a fluorescence microscope (Olympus IX81, Tokyo, Japan).

114

115 **Western blot analysis**

116 Yeast cells grown in 5 ml of YNM medium to an OD_{610} of 1.0 were collected and
117 resuspended in 1 ml of lysis buffer (1 % NaOH, 1 % mercaptoethanol), and kept on ice for 10
118 min. Then, 120 µl of 10 % trichloroacetic acid was added and samples were kept on ice for 10
119 min. Samples were centrifuged at 14,000 x g for 10 min at 4°C. After washing twice with cold
120 acetone, pellets were dissolved in distilled water.

121 Samples were separated by 12 % sodium dodecyl-sulfate polyacrylamide gel
122 electrophoresis and blotted onto a nitrocellulose membrane. Detection was performed using
123 anti-AOD or anti-DAS polyclonal antibody and horseradish peroxidase-linked anti-rabbit
124 antibody.

125

126 **Analysis of interaction between CbHap3p and CbHap5p**

127 The interaction of CbHap3p and CbHap5p was investigated as described previously (Oda
128 *et al.*, 2015). *C. boidinii* strains expressing internal amino acids-deleted CbHap3p-HA were
129 constructed as follows. PCR was performed with primers CbHAP3-100E-up and
130 CbHAP3-225E-down, CbHAP3-97K-up and CbHAP3-225E-down, CbHAP3-85E-up and
131 CbHAP3-225E-down using pP_{ACT1}-CbHAP3-HA (Oda *et al.*, 2015) as a template, yielding

pCbHAP3-HAΔ101-224, pCbHAP3-HAΔ98-224, pCbHAP3-HAΔ86-224, respectively. The resulting plasmids were linearized with EcoT22I and used to transform the *Cbhap3Δura3* strain. A *C. boidinii* strain expressing CbHap5p-His was constructed as follows. First, the coding region of the *CbHAP5* gene was amplified by PCR with primers NotI-CbHAP5-fw and NotI-CbHAP5-rv, using genomic DNA as a template. The 7.4-kb NotI fragment of pNOTeI (Sakai *et al.*, 1996) and the 0.9-kb NotI fragment of the coding region of *CbHAP5* were then ligated to yield pNOT-CbHAP5. Using the resulting plasmid as a template, PCR was performed with primers CbHAP5-C-His-up and pNOTeI-His-down. The amplified fragment was then self-ligated to yield pNOT-CbHAP5-His. The plasmid was linearized with EcoT22I and used to transform strain TK62. The resulting strain was named the *CbHAP5-His*/TK62 strain. *CbHAP3-HAΔ101-224*/*Cbhap3Δ*, *CbHAP3-HAΔ98-224*/*Cbhap3Δ*, *CbHAP3-HAΔ86-224*/*Cbhap3Δ* and *CbHAP5-His*/TK62 cells grown in 100 ml YNM medium to an OD_{610} of 1.0 were harvested and used for immunoprecipitation.

Chromatin immunoprecipitation (ChIP) assay

The chromatin immunoprecipitation (ChIP) assay was done as follows. *CbHAP3-YFP*/*Cbhap3Δ* and *ScHAP3-GFP*/*Cbhap3Δ* cells grown to mid-exponential phase in YNM medium were cross-linked by using 1% formaldehyde for 10 min. Immunoprecipitation was performed by using an anti-GFP antibody at a dilution of 1:400 with MAGnifyTM Chromatin Immunoprecipitation System (Invitrogen, Carlsbad, CA).

Quantitative RT-PCR (qRT-PCR)

Yeast cells were pre-cultured in YND medium for 10 hours, and washed twice with distilled water and transferred to YNM at an OD_{610} value of 1.0. After cultivated for 4 or 8 hours, cells were harvested by centrifugation at 3,000 rpm for 5 min at 4°C, and treated with Yeast Processing Reagent (Takara Bio, Otsu, Japan). Total RNAs were extracted from cells

158 using RNeasy Mini Kit (Qiagen, Hilden, Germany). In addition, to eliminate genomic DNA
159 contaminating total RNA, total RNA was treated with DNase I (RNase-Free DNase Set,
160 Qiagen). Reverse transcription was performed with Random primer (Promega, Madison, WI)
161 and ReverTra Ace (Toyobo, Osaka, Japan). For reverse transcription, 1.0 μg of total RNA was
162 used.
163 qRT-PCR was performed with a Light Cycler Instrument (Roche Diagnostics, Lavel,
164 Canada). The PCR reaction was performed with SYBR Premix Ex Taq (Takara Bio.) and the
165 primers for *ACT1*, *DAS1* and *CbHAP3* listed in Table S1. The program was as follows: 10 sec
166 at 95°C, 40 cycles of 5 sec at 95°C of 20°C/sec, 20 sec at 60°Cof 20°C/sec. Amplicon
167 specificity was verified by melting curve analyses conducted at 65 to 95°C (0 sec at 95°C of
168 20°C/sec, 15 sec at 65°C of 20°C/sec, 0 sec at 95°C of 0.1°C/sec). The number of copies of
169 each sample was determined with the Light Cycler software.
170
171 **Construction of yeast strains expressing heterologous *HAP3* genes**
172 Oligonucleotide primers are listed in Table S1. The *Schap3Δ* strain was generated by
173 homologous recombination by replacing the coding region of *ScHAP3* with the KanMX6
174 cassette (Wach, 1996) following amplification by PCR with the primers ScHAP3del-Fw and
175 ScHAP3del-Rv.
176 The *S. cerevisiae* strain expressing CbHap3p (P_{ScHAP3}-*CbHAP3*/*Schap3Δ*) was constructed
177 as follows. First, the *ScHAP3* promoter region and the coding region of *CbHAP3* were
178 amplified by PCR with the primers pRS-ScHAP3pro-Fw and ScHAP3pro-CbHAP3-Rv, using
179 *S. cerevisiae* genomic DNA as a template, and ScHAP3pro-CbHAP3-Fw and
180 pRS-CbHAP3-Rv, using *C. boidinii* genomic DNA as a template, respectively. Then, using
181 the 0.3-kb fragment of the *ScHAP3* promoter and the 0.9-kb fragment carrying full length
182 *CbHAP3* as a template, the PCR was performed with primers pRS-ScHAP3pro-Fw and
183 pRS-CbHAP3-Rv. Finally, the PCR was performed with primers pRS-up and pRS-down,

using pRS316 as a template. The 4.9-kb fragment of pRS316 and the 1.2-kb fragment of P_{ScHAP3}-*CbHAP3* for P_{ScHAP3}-*CbHAP3*/*Schap3Δ* were used to transform strain *Schap3Δ* to uracil prototrophy using the lithium acetate method (Ito *et al.*, 1983).

C. boidinii strains expressing ScHap3p (*ScHAP3*/*Cbhap3Δ*), the ScHap3p-GFP fusion protein (*ScHAP3-GFP*/*Cbhap3Δ*), or the chimeric protein comprised of the full length of ScHap3p and 106 - 292 amino acids of CbHap3p (*Sc-CbHAP3*/*Cbhap3Δ*) were constructed as follows. For the *ScHAP3*/*Cbhap3Δ* strain, the coding region of *ScHAP3* was amplified by PCR with the primers SalI-ScHAP3-Fw and PstI-ScHAP3-Rv using *S. cerevisiae* genomic DNA as a template. For the *ScHAP3-GFP*/*Cbhap3Δ* strain, the coding region eliminating the stop codon of *ScHAP3* and the coding region of *GFP* were amplified by PCR with the primers SalI-ScHAP3-Fw and PstI-ScHAP3-endcodon-Rv, using *S. cerevisiae* genomic DNA as a template, or the primers PstI-GFP-Fw and PstI-GFP-Rv, using pGFP-PTS1 as a template, respectively. For *Sc-CbHAP3*/*Cbhap3Δ*, the coding region of *ScHAP3* and 316 - 876 bp of *CbHAP3* were amplified by PCR with the primers SalI-ScHAP3-Fw and ScHAP3-CbHAP3-Rv, using *S. cerevisiae* genomic DNA as a template, or the primers ScHAP3-CbHAP3-Fw and PstI-CbHAP3-Rv, using *C. boidinii* genomic DNA as a template, respectively. Then, using these two fragments as a template, PCR was performed with the primers SalI-ScHAP3-Fw and PstI-CbHAP3-Rv. Each SalI-PstI fragment (the 0.4-kb fragment of *ScHAP3*, the 0.4-kb fragment of *ScHAP3* excluding the stop codon, and the 0.9-kb fragment of *Sc-CbHAP3*) and the 7.4-kb fragment of pGFP-PTS1 were ligated to yield pP_{ACT1}-ScHAP3, pP_{ACT1}-ScHAP3-end, or pP_{ACT1}-Sc-CbHAP3, respectively. Then, the 7.8-kb PstI fragment of pP_{ACT1}-ScHAP3-end and the 1.7-kb PstI fragment of the coding region of *GFP* were ligated to yield pP_{ACT1}-ScHAP3-GFP. pP_{ACT1}-ScHAP3, pP_{ACT1}-ScHAP3-GFP and pP_{ACT1}-ScCbHAP3 were linearized with EcoT22I and used to transform strain *Cbhap3Δura3* (Oda *et al.*, 2015). The resulting strains were named *ScHAP3*/*Cbhap3Δ*, *ScHAP3-GFP*/*Cbhap3Δ* and *Sc-CbHAP3*/*Cbhap3Δ*, respectively.

Results

CbHap3p has a unique C-terminal region

Hap3p has a core region that is highly conserved in eukaryotes, including yeasts, fungi, plants, and animals. The core region contains the histone fold motif of histone H2B (Baxevanis *et al.*, 1995). In *S. cerevisiae*, the core region was reported to be responsible for formation of the Hap complex and binding to DNA (McNabb *et al.*, 1997).

Sequence alignment of Hap3p from *C. boidinii* and *S. cerevisiae* revealed that, in addition to the conserved N-terminal region, CbHap3p had an extended structure of ca. 190 amino acids long at the C-terminus (Fig. 1). We also found that, in addition to CbHap3p, Hap3ps in *P. pastoris* and *H. polymorpha* had similar extended structures at their C-termini. In particular, a sequence of approximately 40 amino acids at their C-termini (amino acids 256-292 of CbHap3p) showed high similarity among methylotrophic yeasts (Fig. S1). Based on this information, we speculated that the unique C-terminal region of Hap3p in methylotrophic yeasts has a specific function in methanol-regulated gene expression.

The unique C-terminal region of CbHap3p is essential for specific induction by methanol

In a previous study, we showed that strain *Cbhap3*Δ was impaired in methanol induction during methanol-regulated gene expression (Oda *et al.*, 2015). In order to elucidate the functional role of the C-terminal region of CbHap3p in methanol induction, we first identified the essential regions necessary for gene activation, by expressing CbHap3p-YFP mutant proteins in *Cbhap3*Δ cells (Fig. 2). Wild-type CbHap3p-YFP (CbHap3pFL) complemented the growth defect on methanol, and showed induced production of DAS (Fig. 3). Deletion of amino acids 256 to 292 (CbHap3p Δ256-292) abolished the ability to grow on methanol (Fig. 3a), and the *DAS1* transcript (Table S2) and protein (Fig. 3d) levels were reduced.

We deleted the C-terminal region of CbHap3p-YFP from the opposite side, yielding CbHap3p (Δ107-224), CbHap3p (Δ107-237), CbHap3p (Δ107-241), and CbHap3p (Δ107-256), respectively (Fig. 2). CbHap3p (Δ107-237) and CbHap3 (Δ107-241) showed severe growth defects on methanol (Fig. 3a) and decreased amounts of DAS (Fig. 3d). CbHap3p (Δ107-224) caused retarded growth on methanol, but the amount of DAS protein in this strain was comparable to that in the wild-type CbHap3pFL (Fig. 3a and d). As a result, we concluded that the C-terminal 225-292 amino acids are necessary for methanol induction.

CbHap3p needs to be localized to the nucleus, and bind specifically to methanol-regulated promoters for gene activation to occur. We observed localization of CbHap3p-YFP proteins in methanol-induced conditions, and performed ChIP assays with *Cbhap3*Δ cells producing each CbHap3p-YFP-mutant. Fig. 2 summarizes the results of complementation experiments examining growth on methanol, and localization and binding activity of each mutant CbHap3-YFP protein to the *DAS1* promoter. Like CbHap3pFL, CbHap3p (Δ256-292) localized to the nucleus (Fig. 4a) and bound to P_{DAS1} (Fig. 4b). Similarly, both nuclear localization and DNA binding were normal for all of the above tested CbHap3p proteins (Fig. 4). These results indicate that the C-terminal region of CbHap3p (residues 256-292) is not required for nuclear localization and DNA binding but is required for activation of methanol-regulated gene expression.

The N-terminal region of CbHap3p is necessary for binding to DNA

Next we performed deletions of the N-terminal region of CbHap3p in order to determine whether the conserved N-terminal region of CbHap3p is responsible for nuclear localization and binding to the *DAS1* promoter. Similar to CbHap3pFL, the CbHap3p (Δ1-15), deletion mutant missing the N-terminal 15 amino acids, showed nuclear localization and DNA-binding activity (Fig. 4). In contrast, deletion of the N-terminal 20 or 25 amino acids (CbHap3p (Δ1-20) and CbHap3p (Δ1-25), respectively), caused a severe growth defect on methanol (Fig.

3b), and DAS protein was not produced under methanol-induced conditions (Fig. 3d and Table S2). We confirmed that expression levels of CbHap3p (Δ1-20) and CbHap3p (Δ1-25) were comparable to that of CbHap3p (Δ1-15) (Fig. S2). Both CbHap3p (Δ1-20) and CbHap3p (Δ1-25) were diffused in the cytosol and did not bind to P_{DAS1} (Fig. 4). These results indicate that the N-terminal region (from 16 amino acids including the putative DNA binding motif) is essential for nuclear localization and DNA binding. Our previous study suggested that nuclear localization of CbHap3p depended on CbHap5p (Oda *et al.*, 2015); therefore, the N-terminal region of CbHap3p might also be involved in interacting with CbHap5p.

Further internal deletions in CbHap3p-YFP proteins, CbHap3p (Δ101-224), CbHap3p (Δ98-224), and CbHap3p (Δ86-224), were analyzed in *Cbhap3*Δ cells. The function of CbHap3p (Δ101-224) was similar to that of the CbHap3p Δ107-224 mutant (Fig. 3a, c, d and Fig. 4). However, the deletion of residues 98 to 224 (CbHap3p (Δ98-224)) and 86 to 224 (CbHap3p (Δ86-224)) caused growth impairment on methanol (Fig. 3c) and a low level of DAS protein (Fig. 3d). Interestingly, the mutant proteins CbHap3p (Δ98-224) and CbHap3p (Δ86-224) could bind weakly to DNA, but were mostly diffused in the cytosol. Therefore, the region from 86 to 100 may be partially involved in nuclear localization and DNA binding. Since the DNA binding motif (amino acids 41 to 62) is distant from this deleted region (Romier *et al.*, 2003), the region from 86 to 100 may be involved in complex formation with CbHap2p and CbHap5p. To confirm this hypothesis, we performed co-immunoprecipitation analysis to see interaction between CbHap3p and CbHap5. As a result, the presence of His-tagged CbHap5p was detected only from the sample containing HA-tagged CbHap3p (Δ101-224) (Fig. 4c), indicating that CbHap3p (Δ101-224) interacts with CbHap5p, but CbHap3p (Δ98-224) and CbHap3p (Δ86-224) does not.

Taken together, the deletion analyses indicate that the N-terminal region from amino acids 16 to 100 and the C-terminal region from amino acids 225 to 292 are essential for methanol induction.

Functional complementation of Hap3 proteins between *C. boidinii* and *S. cerevisiae*

In order to determine whether CbHap3p functions in *S. cerevisiae*, we constructed a *S. cerevisiae* strain expressing CbHap3p in the *Schap3Δ* background and tested growth on glycerol medium. As shown in Fig. 5a, the *Schap3Δ* strain harboring empty vector pRS316 did not grow on glycerol, but expression of ScHap3p or CbHap3p in the *Schap3Δ* strain restored the ability to grow. These results suggest that CbHap3p has conserved roles and could function as a transcription factor in *S. cerevisiae*, complementing the respiratory growth defect of the *Schap3Δ* strain. Next, we investigated whether ScHap3p can restore the growth defect of the *Cbhap3Δ* strain on methanol by transforming the *Cbhap3Δ* strain with the ScHap3p-GFP expression plasmid. The *ScHAP3-GFP*/*Cbhap3Δ* strain was unable to grow on methanol (Fig. 5b), although complementation of *Cbhap3Δ* with *CbHAP3-YFP* restored its ability to grow. The *ScHAP3*/*Cbhap3Δ* strain also could not grow on methanol (Fig. S3). These results indicate that ScHap3p could not restore the growth defect of the *Cbhap3Δ* strain.

We performed ChIP assays with *ScHAP3-GFP*/*Cbhap3Δ* cells that had been induced by methanol (Fig. 5c). As a result, all tested promoter regions of methanol-inducible genes could be amplified from the template DNA, whereas P_{ACT1} was not amplified. These results indicate that ScHap3p could bind to methanol-inducible promoters in *C. boidinii*, but was unable to function as a transcription factor to restore the growth defect of the *Cbhap3Δ* strain on methanol.

These results support the model that the N-terminal region of CbHap3p is involved in DNA binding, and the C-terminal extended region plays a unique role in methanol induction. To confirm this hypothesis, we constructed a chimeric Hap3 protein that consists of the full length ScHap3p and the region of CbHap3p from amino acids 106 to 292. Production of the chimeric protein *Sc-CbHAP3* in the *Cbhap3Δ*mutant partially restored the ability to grow on

methanol (Fig. 5d), indicating that the C-terminal region of CbHap3p functions in methanol induction.

In conclusion, our results revealed that the unique C-terminal region of CbHap3p is required for activation of methanol-regulated genes but not for nuclear localization and DNA binding, while the N-terminal region is responsible for nuclear localization and binding of CbHap3p to methanol-regulated promoters (Fig. 6).

Discussion

The Hap complex is highly conserved among all eukaryotes and is known to activate genes involved in gluconeogenesis, respiration, and mitochondria biogenesis, and contributes to glucose repression/derepression (Buschlen *et al.*, 2003; McNabb & Pinto, 2005). In a previous study, we demonstrated that the *C. boidinii* Hap complex is involved in methanol-regulated gene expression via methanol induction (Oda *et al.*, 2015), revealing a unique role of the Hap complex in the methylotrophic yeast. It has been of great interest to understand how the Hap complex is able to execute such a specialized function in methanol-regulated gene expression in methylotrophic yeasts.

In this study, the unique C-terminal extended region of CbHap3p, which is not present in ScHap3p, was found to play a critical role in methanol induction. Interestingly, this C-terminal extended region is also present in other methylotrophic yeast strains, *P. pastoris* and *H. polymorpha* (Fig. S1). However, BLAST searches did not find sequences homologous to this C-terminal region in any eukaryotes other than methylotrophic yeasts. In particular, the 37-amino acids sequence identified within the C-terminal region (amino acids 256 to 292) was critical for methanol induction, and was highly conserved among Hap3 proteins from methylotrophic yeasts, suggesting the functional importance of this region. We speculated that the C-terminal region of CbHap3p is responsible for methanol induction after binding to methanol-regulated promoters. This notion was supported by the demonstration that deletion of the C-terminal region abolished induction of methanol-regulated genes, but did not affect nuclear localization and binding to P_{DAS1} (Fig. 3d, 4 and Table S2). Therefore, the C-terminal region is speculated to be involved in recruiting other transcription factors that activate methanol-regulated promoters (Fig. 6).

In contrast to the role of the C-terminal region of CbHap3p, the N-terminal region, which is widely conserved in Hap3 proteins, was found to have conserved functions in nuclear

347 localization and binding to DNA. The identified N-terminal region of CbHap3p that is
348 necessary for growth on methanol corresponds to the ScHap3p region required for growth on
349 lactate (Xing *et al.*, 1993). Moreover, the region of human NF-YB (corresponding to Hap3p)
350 that is necessary for complex formation with NF-YC (corresponding to Hap5p) is also
351 comparable with the identified N-terminal region of CbHap3p (Romier *et al.*, 2003).
352 CbHap3p was shown to interact with CbHap5p, and localize to nucleus. The core regions of
353 CbHap5p are also highly conserved among various eukaryotes (Oda *et al.*, 2015). Therefore,
354 it is strongly suggested that the CbHap3p N-terminal region also participates in complex
355 formation with CbHap2p and CbHap5p (Fig. 6). Although we showed that the CbHap3p
356 N-terminal region functions in *S. cerevisiae*, the growth defect of *Schap3Δ* on glycerol was
357 not recovered by expressing the first 121 amino acids of CbHap3p (data not shown). One
358 possible reason is that the difference of the theoretical pIs of Hap3 proteins (ScHap3p is 4.78
359 while CbHap3p 1-121 is 8.8) altered the specificity of DNA binding.
360 In our previous study, we showed that the CbHap complex localized to the nucleus
361 regardless of the carbon source (Oda *et al.*, 2015), and we confirmed that the transcript level
362 of *CbHAP3* was not increased by methanol (Table S3). Therefore, in addition to the Hap
363 complex, induction of methanol-regulated genes seems to require other some activation
364 factors. In *S. cerevisiae* and other yeasts, Hap4p interacts with the Hap2p/3p/5p heterotrimer
365 (Forsburg & Guarente, 1989; Bourgarel *et al.*, 1999; Sybirna *et al.*, 2005; Sybirna *et al.*, 2010).
366 Indeed, a domain required for recruiting Hap4p to the Hap2p/3p/5p complex (Hap4p
367 recruiting domain) has been identified in Hap5p (McNabb *et al.*, 1997), and this domain is
368 also conserved in CbHap5p. However, this domain is not always required for the function of
369 the Hap complex (Tanoue *et al.*, 2006), indicating that other activators/repressors may interact
370 with other regions of Hap2p/Hap3p/Hap5p. On the other hand, Hap4p recruiting domain is
371 absent in CBF-A/CBF-B/CBF-C, which corresponds to Hap3p/Hap2p/Hap5p in rat. However,
372 the CBF-A/CBF-B/CBF-C complex can activate transcription of target genes without any

other activators like Hap4p (McNabb *et al.*, 1997). So far we could not find a gene highly homologous to *ScHAP4* in the *C. boidinii* draft genome sequence (Oda *et al.*, 2015), but the hypothetical CbHap4p might interact with a putative Hap4p recruiting domain in CbHap5p. Indeed, a constructed strain expressing CbHap5p deleted for the putative Hap4p-recruiting domain still grew normally on methanol (data not shown). This observation suggests that, in the case of *C. boidinii*, some methanol specific transcriptional activators interact with other regions of the Hap complex, including the C-terminal region of CbHap3p.

Hap3p has been reported to be functionally interchangeable between yeast and human cells (Chodosh *et al.*, 1988), indicating that Hap3p function has been highly conserved through evolution. Originally, Hap3ps in methylotrophic yeasts was speculated to have some function as an activator for derepression during growth on non-fermentable carbon sources. However, during evolution, these Hap3 proteins may have acquired the specific function of activating methanol induction with their C-terminal region, and lost the derepression function in methylotrophic yeasts. The identified features of CbHap3p suggest a mechanism for methanol induction by the C-terminal region. To our knowledge, this is the first report identifying the function of the C-terminal region in Hap3p and indicates the presence of conserved machinery for methanol-regulated gene expression mediated by Hap3p in methylotrophic yeasts. This knowledge should contribute to the elucidation of a detailed molecular mechanism of methanol-regulated gene expression.

Acknowledgments

This research was supported in part by Grant-in-Aid for Scientific Research (B) (25281063 and to H.Y.) from the Japan Society for the Promotion of Science and by the Asahi Glass Foundation. It was also supported in part by Advanced Low Carbon Technology Research and Development Program (ALCA) and CREST from Japan Science and Technology Agency.

399 **References**

400

401 **Baxevanis, A. D., Arents, G., Moudrianakis, E. N. & Landsman, D. (1995)**. A variety of
402 DNA-binding and multimeric proteins contain the histone fold motif. *Nucleic Acids Res* **23**,
403 2685-2691.

404 **Bourgarel, D., Nguyen, C. C. & Bolotin-Fukuhara, M. (1999)**. HAP4, the
405 glucose-repressed regulated subunit of the HAP transcriptional complex involved in the
406 fermentation-respiration shift, has a functional homologue in the respiratory yeast
407 *Kluyveromyces lactis*. *Mol Microbiol* **31**, 1205-1215.

408 **Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P. & Boeke, J. D.**
409 **(1998)**. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set
410 of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**,
411 115-132.

412 **Buschlen, S., Amillet, J. M., Guiard, B., Fournier, A., Marcireau, C. & Bolotin-Fukuhara,**
413 **M. (2003)**. The *S. cerevisiae* HAP complex, a key regulator of mitochondrial function,
414 coordinates nuclear and mitochondrial gene expression. *Comp Funct Genomics* **4**, 37-46.

415 **Chodosh, L. A., Olesen, J., Hahn, S., Baldwin, A. S., Guarente, L. & Sharp, P. A. (1988)**.
416 A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally
417 interchangeable. *Cell* **53**, 25-35.

418 **Daly, R. & Hearn, M. T. (2005)**. Expression of heterologous proteins in *Pichia pastoris*: a
419 useful experimental tool in protein engineering and production. *J Mol Recognit* **18**, 119-138.

420 **Forsburg, S. L. & Guarente, L. (1989)**. Identification and characterization of HAP4: a third
421 component of the CCAAT-bound HAP2/HAP3 heteromer. *Genes Dev* **3**, 1166-1178.

422 **Gellissen, G. (2000)**. Heterologous protein production in methylotrophic yeasts. *Appl*
423 *Microbiol Biotechnol* **54**, 741-750.

424 **Hartner, F. S. & Glieder, A. (2006)**. Regulation of methanol utilisation pathway genes in

425 yeasts. *Microb Cell Fact* **5**, 39.

426 **Ito, H., Fukuda, Y., Murata, K. & Kimura, A. (1983)**. Transformation of intact yeast cells
427 treated with alkali cations. *J Bacteriol* **153**, 163-168.

428 **McNabb, D. S. & Pinto, I. (2005)**. Assembly of the Hap2p/Hap3p/Hap4p/Hap5p-DNA
429 Complex in *Saccharomyces cerevisiae*. *Eukaryotic Cell* **4**, 1829-1839.

430 **McNabb, D. S., Tseng, K. A. & Guarente, L. (1997)**. The *Saccharomyces cerevisiae* Hap5p
431 homolog from fission yeast reveals two conserved domains that are essential for assembly of
432 heterotetrameric CCAAT-binding factor. *Mol Cell Biol* **17**, 7008-7018.

433 **Oda, S., Yurimoto, H., Nitta, N., Sasano, Y. & Sakai, Y. (2015)**. Molecular Characterization
434 of Hap Complex Components Responsible for Methanol-Inducible Gene Expression in the
435 Methylotrophic Yeast *Candida boidinii*. *Eukaryot Cell* **14**, 278-285.

436 **Ramil, E., Agrimonti, C., Shechter, E., Gervais, M. & Guiard, B. (2000)**. Regulation of the
437 *CYB2* gene expression: transcriptional co-ordination by the Hap1p, Hap2/3/4/5p and Adr1p
438 transcription factors. *Mol Microbiol* **37**, 1116-1132.

439 **Ridenour, J. B. & Bluhm, B. H. (2014)**. The HAP complex in *Fusarium verticillioides* is a
440 key regulator of growth, morphogenesis, secondary metabolism, and pathogenesis. *Fungal*
441 *Genet Biol* **69**, 52-64.

442 **Romier, C., Cocchiarella, F., Mantovani, R. & Moras, D. (2003)**. The NF-YB/NF-YC
443 structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y.
444 *J Biol Chem* **278**, 1336-1345.

445 **Sakai, Y., Kazarimoto, T. & Tani, Y. (1991)**. Transformation system for an asporogenous
446 methylotrophic yeast, *Candida boidinii*: cloning of the orotidine-5'-phosphate decarboxylase
447 gene (*URA3*), isolation of uracil auxotrophic mutants, and use of the mutants for integrative
448 transformation. *J Bacteriol* **173**, 7458-7463.

449 **Sasano, Y., Yurimoto, H., Kuriyama, M. & Sakai, Y. (2010)**. Trm2p-dependent
450 derepression is essential for methanol-specific gene activation in the methylotrophic yeast

Candida boidinii. *FEMS Yeast Res* **10**, 535-544.

Sasano, Y., Yurimoto, H., Yanaka, M. & Sakai, Y. (2008). Trm1p, a Zn(II)2Cys6-type transcription factor, is a master regulator of methanol-specific gene activation in the methylotrophic yeast *Candida boidinii*. *Eukaryot Cell* **7**, 527-536.

Singh, R. P., Prasad, H. K., Sinha, I., Agarwal, N. & Natarajan, K. (2011). Cap2-HAP complex is a critical transcriptional regulator that has dual but contrasting roles in regulation of iron homeostasis in *Candida albicans*. *J Biol Chem* **286**, 25154-25170.

Sybirna, K., Guiard, B., Li, Y. F., Bao, W. G., Bolotin-Fukuhara, M. & Delahodde, A. (2005). A new *Hansenula polymorpha HAP4* homologue which contains only the N-terminal conserved domain of the protein is fully functional in *Saccharomyces cerevisiae*. *Curr Genet* **47**, 172-181.

Sybirna, K., Petryk, N., Zhou, Y. F., Sibirny, A. & Bolotin-Fukuhara, M. (2010). A novel *Hansenula polymorpha* transcriptional factor *HpHAP4-B*, able to functionally replace the *S. cerevisiae HAP4* gene, contains an additional bZip motif. *Yeast* **27**, 941-954.

Tani, Y., Sakai, Y. & Yamada, H. (1985). Isolation and Characterization of a Mutant of a Methanol Yeast, *Candida boidinii* S2, with Higher Formaldehyde Productivity. *Agric Biol Chem* **49**, 2699-2706.

Tanoue, S., Kamei, K., Goda, H., Tanaka, A., Kobayashi, T., Tsukagoshi, N. & Kato, M. (2006). The region in a subunit of the Aspergillus CCAAT-binding protein similar to the HAP4p-recruiting domain of *Saccharomyces cerevisiae* Hap5p is not essential for transcriptional enhancement. *Biosci Biotechnol Biochem* **70**, 782-787.

Vogl, T. & Glieder, A. (2012). Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N Biotechnol* **30**, 385-404.

Wach, A. (1996). PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. *Yeast* **12**, 259-265.

Xing, Y., Fikes, J. D. & Guarente, L. (1993). Mutations in yeast HAP2/HAP3 define a

hybrid CCAAT box binding domain. *EMBO J* **12**, 4647-4655.

Yurimoto, H. (2009). Molecular basis of methanol-inducible gene expression and its application in the methylotrophic yeast *Candida boidinii*. *Biosci Biotechnol Biochem* **73**, 793-800.

Yurimoto, H., Oku, M. & Sakai, Y. (2011). Yeast methylotrophy: metabolism, gene regulation and peroxisome homeostasis. *Int J Microbiol* **2011**, 101298.

Zhai, Z., Yurimoto, H. & Sakai, Y. (2012). Molecular characterization of *Candida boidinii MIG1* and its role in the regulation of methanol-inducible gene expression. *Yeast* **29**, 293-301.

Figure legends

Fig. 1. (a) Schematic model of CbHap3p and ScHap3p. The core region showing high similarity (gray) and the DNA-binding motif (black) are represented. (b) Alignment of amino acid sequences of CbHap3p and ScHap3p.

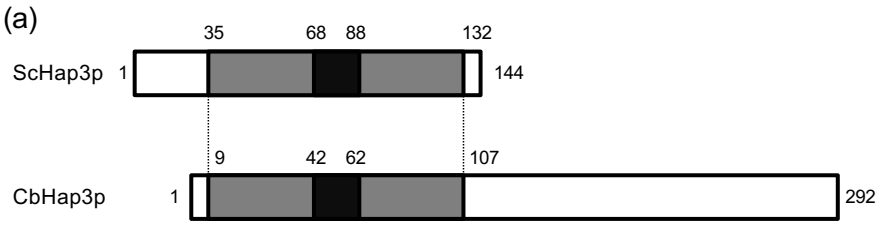
Fig. 2. Deletion analysis of CbHap3p. The 292-amino acids CbHap3p is represented schematically with the N-terminal region including the DNA-binding motif (gray) and the conserved C-terminal domain in methylotrophic yeasts (hatched). Growth on methanol, intracellular localization and DNA-binding activity of CbHap3p are shown for each CbHap3p-YFP mutant protein or expressing strain. Growth on methanol: +++, same growth as CbHap3pFL; ++, partially impaired growth; +, weak growth; -, no growth; Localization: N, nucleus; C, cytosol; DNA-binding: ++, binding similar to the wild-type CbHap3pFL; +, weak binding; -, no binding

Fig. 3. (a-c) Growth of *C. boidinii* strains expressing CbHap3p-YFP variants in YNM medium. (d) Western blot analysis. Cells were incubated in YNM medium for 8 h. Western blot analysis was performed with anti-DAS antibody. Lane 1, *Cbhap3Δ*; 2, Δ86-224; 3, Δ98-224; 4, Δ101-224; 5, Δ1-15; 6, Δ1-20; 7, Δ1-25; 8, Δ107-224; 9, Δ107-237; 10, Δ107-241; 11, Δ107-256; 12, Δ256-292; 13, CbHap3pFL.

Fig. 4. (a) Localization of CbHap3p-YFP. (b) ChIP assay. YFP-tagged CbHap3p variants were immunoprecipitated with (+) or without (-) anti-GFP antibody. IP, Immunoprecipitation; WCE, Whole cell extract. (c) Interaction between internal amino acids-deleted CbHap3p and CbHap5p. Cells expressing internal amino acids-deleted CbHap3p-HA (+) or native CbHap3p (-) and CbHap5p-His (+) were incubated in YNM medium for 8 h. Lysates were

513 immunoprecipitated with anti-HA-tagged MAb magnetic beads. Western blot was performed
514 with anti-His antibody.
515
516 **Fig. 5.** The unique role of CbHap3p in growth on methanol resides in its C-terminal region.
517 (a) The *S. cerevisiae Schap3Δ* strains expressing ScHap3p or CbHap3p were spotted on YPG
518 agar plates, which were incubated for 3 d at 28°C. The *Schap3Δ* strain transformed with the
519 empty pRS316 vector was the negative control. (b) Growth of *Cbhap3Δ* strains expressing
520 CbHap3p-YFP or ScHap3p-GFP in YNM medium. Symbols: closed circles, wild type; open
521 circles, *CbHAP3-YFP*/*Cbhap3Δ*; closed triangles, *SCHAP3-YFP*/*Cbhap3Δ*. (c) ChIP assay
522 was performed with cells grown on methanol. GFP-tagged ScHap3p was immunoprecipitated
523 with (+) or without (-) anti-GFP antibody. IP, Immunoprecipitation; WCE, Whole cell extract.
524 (d) Growth of the *Cbhap3Δ* strain expressing Sc-CbHap3p on methanol. Symbols: closed
525 circles, wild type; open circles, *Sc-CbHAP3* /*Cbhap3Δ*; closed triangles, *Cbhap3Δ*.
526
527 **Fig. 6.** Functional regions of CbHap3p and model for activation of methanol-regulated genes
528 by the Hap complex. The N-terminal region of CbHap3p (amino acids 16-100, gray) is
529 necessary for Hap complex formation (interaction with CbHap2p and CbHap5p) and binding
530 to the promoter. The C-terminal region (amino acids 225-292, hatched) is involved in
531 activation of methanol-regulated genes.

Fig. 1.



(b)

```
ScHap3p       MNTNESEHVSTSPEDTQENGGNASSSGSLQQISTLREQDRWLPINNVARLMKNTLPPSAK 60
CbHap3p       -------------------------MTGNGEFELREQDRWLPIANVARIMKTTLPQTAK 34
                                       : :    ********** ****:**.*** :**
                        DNA binding motif
ScHap3p       VSKDAKECMQECVSELISFVTSEASDRCAADKRKTINGEDILISLHALGFENYAEVLKIY 120
CbHap3p       VSKDAKECMQECVSEFISFITSEASDRCLMEKRKTINGEDILFSMSSLGFENYSEVLKIY  94
              ***************:***:********  :***********:*: :******:******

ScHap3p       LAKYRQQQALKNQLMYEQDDEEVP---------------------------------- 144
CbHap3p       LAKYREQQALRQERGEVRKKTTNRKNSEKLKDNSNDDMELDNGHSLEEGHMVLSENDPSK 154
              *****:****:::    :..

ScHap3p       ------------------------------------------------------------
CbHap3p       DTITGKSKSENNNTDNDTKIKPENEHDEYVQQLYEQDYGGQGDKYSSHNPNYHNNQSVPQ 214


ScHap3p       ------------------------------------------------------------
CbHap3p       LDSLNAPQLVEAAIIKNEAVANDGNDTAAVSVSTADPHTLNLNLNDNVNGVVTETEELAS 274


ScHap3p       ------------------
CbHap3p       LTNGHHGENVLYRYQNGF 292
```

Fig. 2.

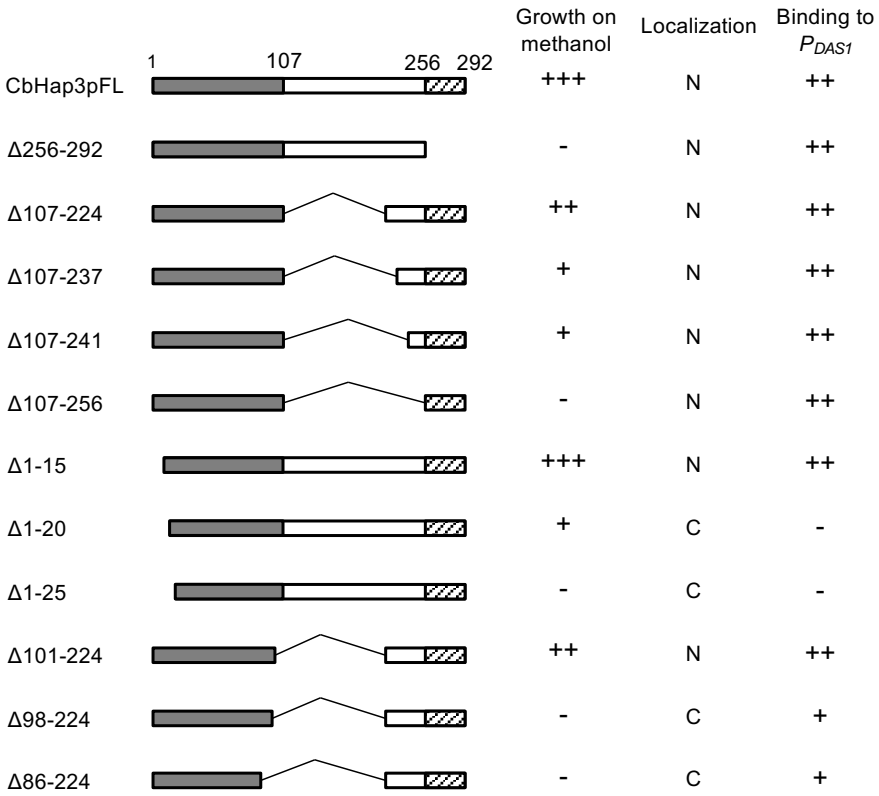
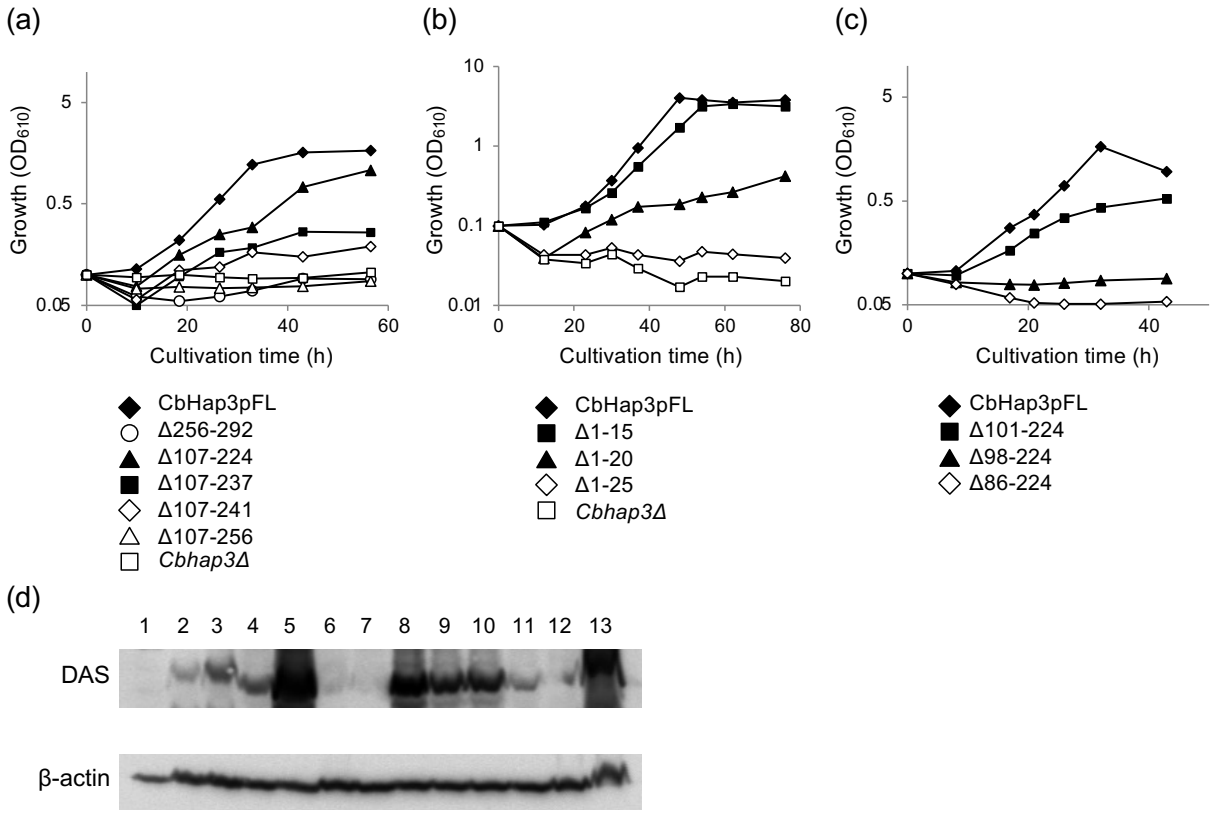

Growth on methanol
Localization
Binding to P_{DAS1}
1
107
256
292
CbHap3pFL
+++
N
++
Δ256-292
-
N
++
Δ107-224
++
N
++
Δ107-237
+
N
++
Δ107-241
+
N
++
Δ107-256
-
N
++
Δ1-15
+++
N
++
Δ1-20
+
C
-
Δ1-25
-
C
-
Δ101-224
++
N
++
Δ98-224
-
C
+
Δ86-224
-
C
+

Fig. 3.


(a)
(b)
(c)
Growth (OD_{610})
Cultivation time (h)
CbHap3pFL
Δ256-292
Δ107-224
Δ107-237
Δ107-241
Δ107-256
Cbhap3Δ
Δ1-15
Δ1-20
Δ1-25
Δ101-224
Δ98-224
Δ86-224

(d)
1 2 3 4 5 6 7 8 9 10 11 12 13
DAS
β-actin

Fig. 4.

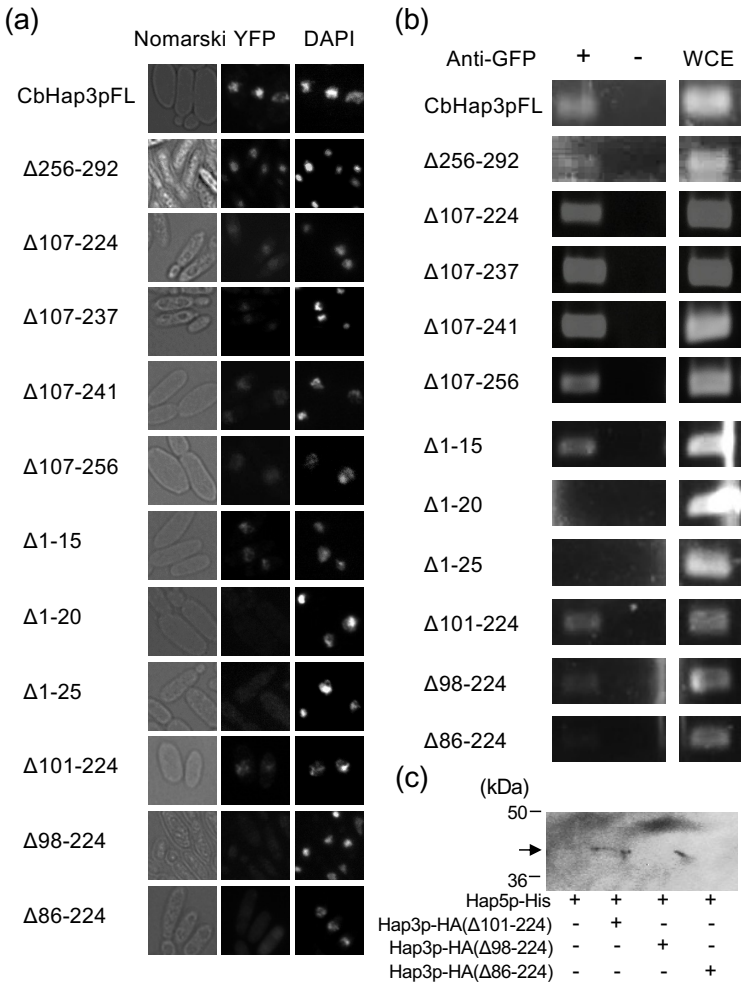

(a)
Nomarski YFP DAPI
CbHap3pFL
Δ256-292
Δ107-224
Δ107-237
Δ107-241
Δ107-256
Δ1-15
Δ1-20
Δ1-25
Δ101-224
Δ98-224
Δ86-224
(b)
Anti-GFP + - WCE
CbHap3pFL
Δ256-292
Δ107-224
Δ107-237
Δ107-241
Δ107-256
Δ1-15
Δ1-20
Δ1-25
Δ101-224
Δ98-224
Δ86-224
(c)
(kDa)
50
36
Hap5p-His + + + +
Hap3p-HA(Δ101-224) - + - -
Hap3p-HA(Δ98-224) - - + -
Hap3p-HA(Δ86-224) - - - +

Fig. 5.

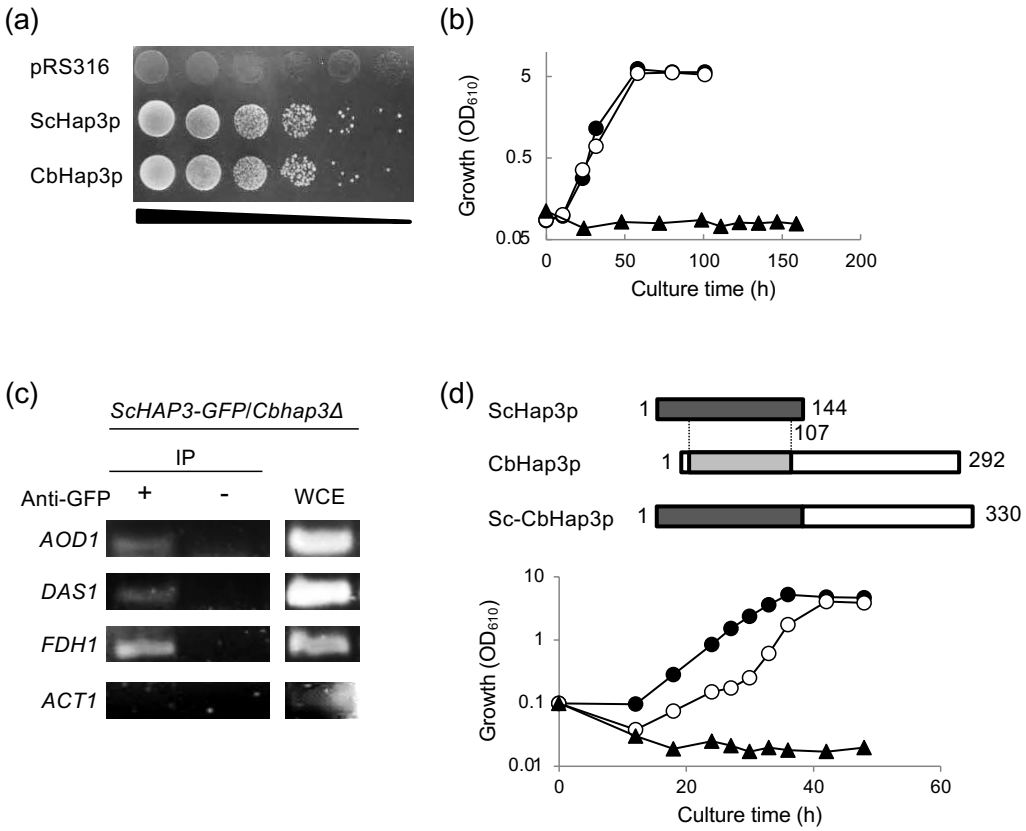

(a)
pRS316
ScHap3p
CbHap3p
(b)
Growth (OD610)
5
0.5
0.05
0
50
100
150
200
Culture time (h)
(c)
ScHAP3-GFP/Cbhap3Δ
IP
Anti-GFP
+
-
WCE
AOD1
DAS1
FDH1
ACT1
(d)
ScHap3p
1
144
107
CbHap3p
1
292
Sc-CbHap3p
1
330
Growth (OD610)
10
1
0.1
0.01
0
20
40
60
Culture time (h)

Fig. 6.

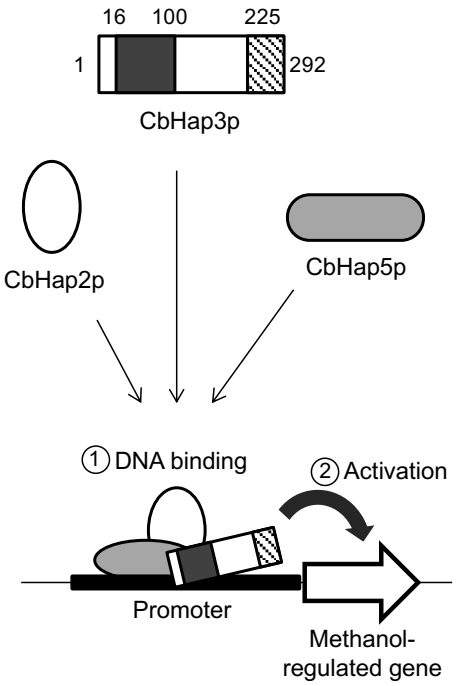

16
100
225
1
292
CbHap3p
CbHap2p
CbHap5p
① DNA binding
② Activation
Promoter
Methanol-
regulated gene

Table S1. Primers used in this study.

Primer	Sequence (5' → 3')
ScHAP3del-Fw	AGCTAGATAGTAACACAAGTGGCACAAACCTCTCGAGAATGTTTAG CTTGCCTCGTCCC
ScHAP3del-Rv	AGCTAGCAACTTTTGCGATCTACCACCTGGTTTTGTCTTCATCGATG AATTCGAGCTCG
pRS-ScHAP3pro-Fw	ACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATCAGAATC AACTTCAAATCACCCTATCTGTG
ScHAP3pro-CbHAP3-Rv	TTCCTGTCATATTCTCGAGAGGTTTGTGCCACTTGTGTTA
ScHAP3pro-CbHAP3-Fw	TCTCGAGAATATGACAGGAAACGGAGAATTTGAATTAAGA
pRS-CbHAP3-Rv	CAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGGTGATATA AACAAATAAATTTGAAATTTGA
pRS-up	ATTCGCCCTATAGTGAGTCGTATTACAATT
pRS-down	CAGCTTTTGTTCCCTTTAGTGAGGGTTAAT
SalI-ScHAP3-Fw	ACGCGTCGACATGAATACCAACGAGTCCGA
PstI-ScHAP3-Rv	AACTGCAGTCAAGGCACCTCTTCGTCGT
PstI-ScHAP3-endcodon-Rv	AACTGCAGAGGCACCTCTTCGTCGTCCTGC
PstI-GFP-Fw	AACTGCAGATGGGTAAAGGAGAAGAACTTT
PstI-GFP-Rv	AACTGCAGTCTGAGTCCGGACTTGTATAGT
ScHAP3-CbHAP3-Rv	CACCTCTTTCAGGCACCTCTTCGTCGTCCTGCTCATACAT
ScHAP3-CbHAP3-Fw	AGAGGTGCCTGAAAGAGGTGAAGTTAGAAAAAAGACAACA
PstI-CbHAP3-Rv	AACTGCAGTTAAAAACCGTTTTGATATCTG
CbHAP3-256L-up	ATTTAAAGTATGAGGATCGGCTGTACTTACCGATACAGC
YFP-startcodon-Fw	GTTTCTAAAGGTGAAGAATTATTCACTGGTGTTGTTCCAATTTTAGT
CbHAP3-106Q-up	TTGTCTAAGTGCTTGTTGTTCTCTATATTT
CbHAP3-225E-down	GAAGCAGCAATCATTAAAAATGAAGCAGTT
CbHAP3-238G-down	GGTAACGATACAGCTGCTGTATCGGTAAGT
CbHAP3-242A-down	GCTGCTGTATCGGTAAGTACAGCCGATCCT
CbHAP3-257N-down	AATTTAAACGATAACGTTAACGGTGTAGTT
CbHAP3pro-ATG-up	CATTGTGATATTAGTTGTATATGTAAGTATGTGTTTAA
CbHAP3-16L-down	TTACCTATTGCCAATGTGGCAAGAATTATG
CbHAP3-21V-down	GTGGCAAGAATTATGAAAACAACTTTACCA
CbHAP3-26K-down	AAAACAACTTTACCACAAACTGCAAAAGTATCA
CbHAP3-100E-up	TTCTCTATATTTAGCTAAATAAATTTTCAA
CbHAP3-97K-up	TTTAGCTAAATAAATTTTCAAAACTTCAGA
CbHAP3-85E-up	TTCAAAACCTAAAGAAGACATTGAAAATAA
NotI-CbHAP5-fw	GCGGCCGCATGAGTAGTAATTCAAGAGAAGACGAAATGTC
NotI-CbHAP5-rv	GCGGCCGCTGCATATTCACCTTGTTGTTGCTGTTGCTGCT
CbHAP5-C-His-up	ATGATGATGTGCATATTCACCTTGTTGTTGCTGTTGCTG
pNOTeI-His-down	CATCACCACTAAGCGGCCGCTAATTCAACAAGTTGTATCTTTTTTTACT GCTCT
AOD1-ChIP-Fw	CCCAGCTTTTCAATTTAATAAAATAGCC
AOD1-ChIP-Rv	GATAGTAAATATAGTAAAATGTGATATGGG
DAS1-ChIP-Fw	ATATTTGGTGGACCTCTCAGTTGCATT
DAS1-ChIP-Rv	AGTCCACTTGACTGAAGACTACGCTAAT
FDH1-ChIP-Fw	TGTTACAATTGTCACAATTCTTGGATATAC
FDH1-ChIP-Rv	AACATCTGACTAGTATTACCATAAATGTAC
ACT1-ChIP-Fw	CCACTGAGTTCCTTCTTTCTGTTT
ACT1-ChIP-Rv	AGAATCTGGGAGGAAAGAATAGAGA
RT-ACT1-fw	GTTGGTAGACCAAGACATCAAGGTATCATG
RT-ACT1-rv	CTTAATTCGTTGTAGAAAGTGTGATGCCAG
RT-DAS1-fw	TGAGAGCATTCCGTTGTTATGTCTTGGA
RT-DAS1-rv	CAACCATACCCATAGCAGAACCAGGATG
RT-CbHAP3-fw	ATACGAACAAGATTATGGTGGACAAGG
RT-CbHAP3-rv	GCAGCTGTATCGTTACCGTCA

Table S2. Transcript level of *DAS1*.

	Relative transcript level (%)[a]	
	Induction time (h)	
	4	8
CbHap3FL	1.00×10^{2}	1.48×10^{2}
Δ256~292	1.20×10^{-2}	2.00
Δ107~224	2.40	88.9
Δ107~237	1.29×10^{-1}	98.6
Δ107~241	8.33×10^{-2}	36.3
Δ107~256	4.14×10^{-3}	0.75
Δ1~15	1.83×10^{2}	1.42×10^{2}
Δ1~20	46.7	2.80
Δ1~25	1.18	2.29
Δ101~224	36.3	17.7
Δ98~224	0.31	4.67
Δ86~224	6.27×10^{-2}	4.25×10^{-3}

[a] Relative transcript level of *DAS1* was standardized against the level of *ACT1*, and then expressed as the relative value to those in cells of CbHap3FL (4 h).

Table S3. Transcript level of *CbHAP3*.

Carbon source	Relative transcript level (%)[a]
Glucose	100
Methanol	103
Ethanol	95.3
Glycerol	71.2

Wild-type cells were incubated in the medium containing indicated carbon sources for 4 h.
[a] Relative transcript level of *CbHAP3* was standardized against the level of *ACT1*, and then expressed as the relative value to those in cells grown on glucose.

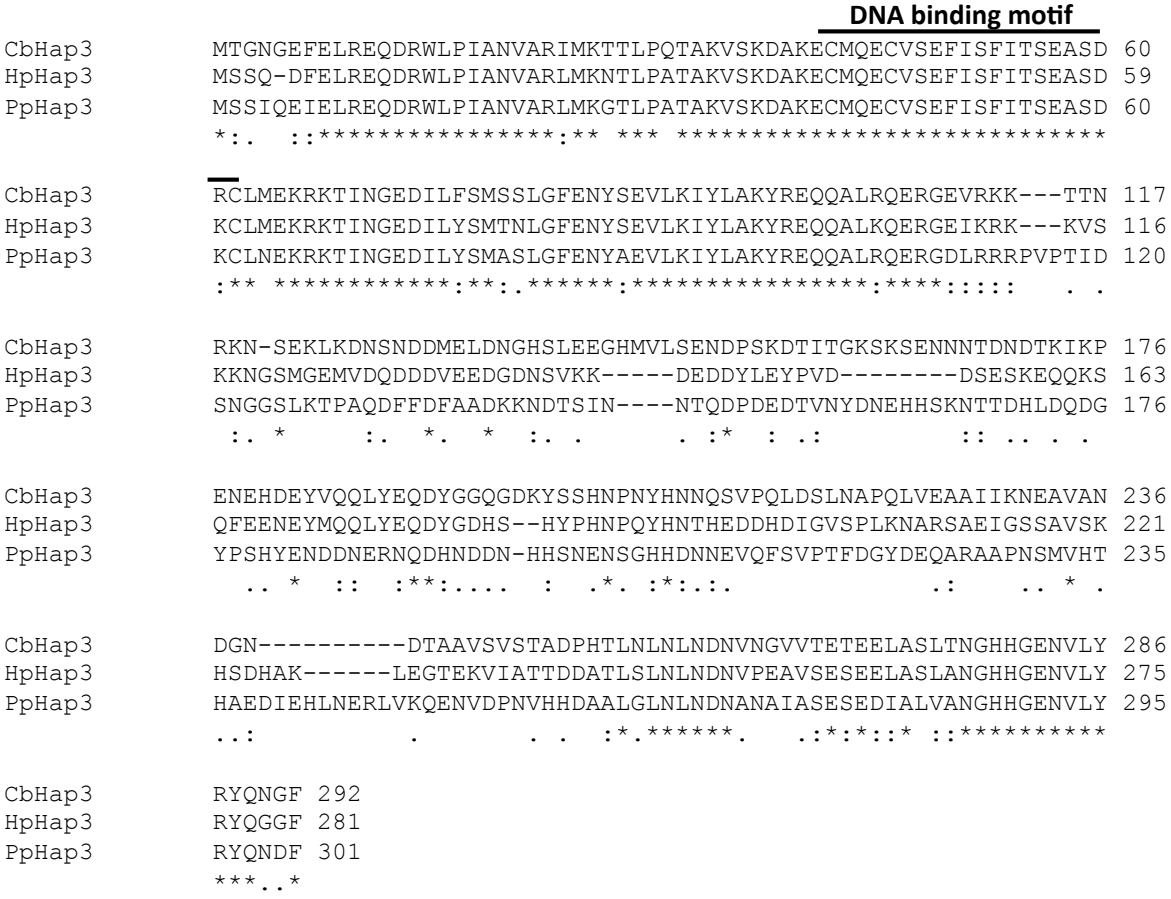


Fig. S1. Alignment of the amino acid sequences of Hap3ps in three methylotrophic yeasts, *C. boidinii* (Cb), *H. polymorpha* (Hp), and *P. pastoris* (Pp).

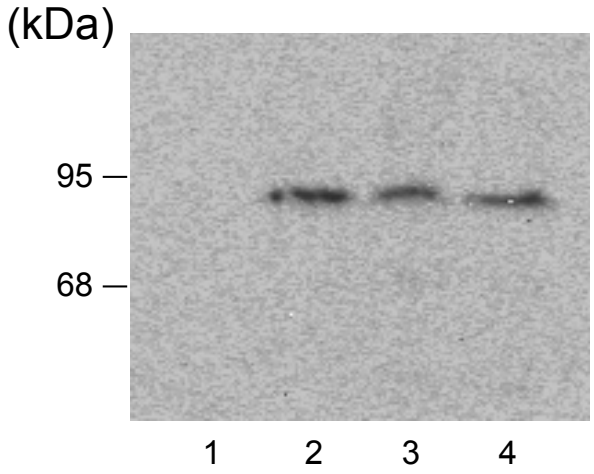


Fig. S2. Western blot analysis of CbHap3p-YFP mutant proteins.
Lane 1, wild type; 2, CbHap3p (Δ1-15)-YFP; 3, CbHap3p (Δ1-20)-YFP;
4, CbHap3p (Δ1-25)-YFP

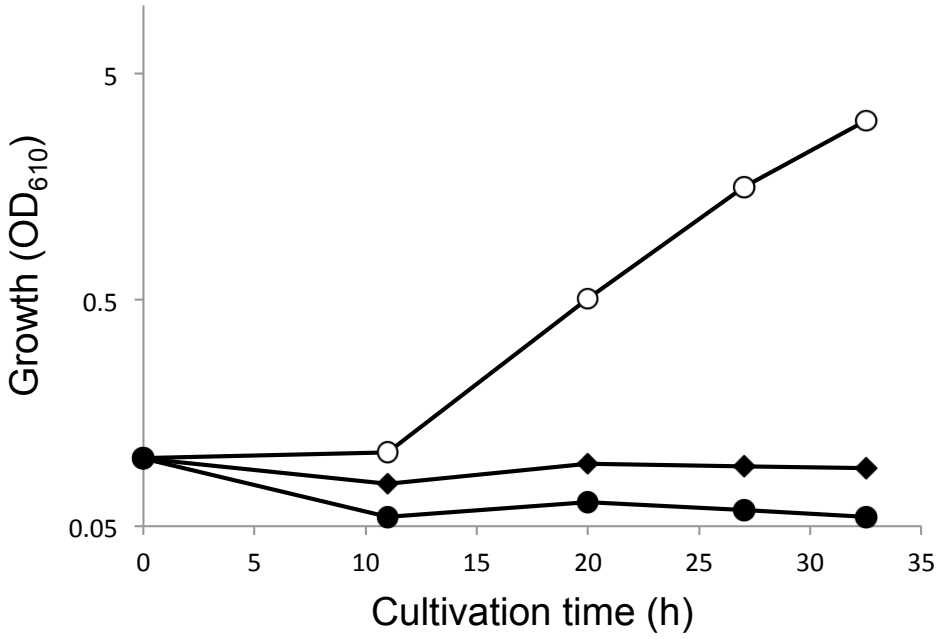


Fig. S3. Growth of *C. boidinii* strain expressing ScHap3p on methanol. Symbols: open circles, wild type; closed diamonds, *ScHAP3*/*Cbhap3Δ*; closed circles, *Cbhap3Δ*.