Nucleic Acids Research, 2000, Vol. 28, No. 7 **1499–1505**

SURVEY AND SUMMARY

Saccharomyces cerevisiae basic helix–loop–helix proteins regulate diverse biological processes

Kelly A. Robinson and John M. Lopes[1,*]

Department of Molecular and Cellular Biochemistry, Loyola University of Chicago, 2160 South First Avenue, Maywood, IL 60153, USA and [1]Department of Biological Sciences, Wayne State University, 5047 Gullen Mall, Detroit, MI 48202, USA

Received December 16, 1999; Revised and Accepted February 14, 2000

ABSTRACT

Basic helix–loop–helix (bHLH) proteins are among the most well studied and functionally important regulatory proteins in all eukaryotes. The HLH domain dictates dimerization to create homo- and heterodimers. Dimerization juxtaposes the basic regions of the two monomers to create a DNA interaction surface that recognizes the consensus sequence called the E-box, 5′-CANNTG-3′. Several bHLH proteins have been identified in the yeast *Saccharomyces cerevisiae* using traditional genetic methodologies. These proteins regulate diverse biological pathways. The completed sequence of the yeast genome, combined with novel methodologies allowing whole-genome expression studies, now offers a unique opportunity to study the function of these bHLH proteins. It is the purpose of this review to summarize the current knowledge of bHLH protein function in yeast.

GENERAL PROPERTIES OF BASIC HELIX–LOOP–HELIX (bHLH) PROTEINS

In mammals, bHLH proteins have critical roles in development, cell growth, differentiation and apoptosis (1). The inaugural members of this family of proteins were c-Myc and MyoD. Both proteins have been studied extensively with respect to functional domains and DNA binding specificity. From these studies it is clear that the role of the HLH domain is to create dimer combinations by interactions between the amphipathic helices (2,3). The consequence of dimerization is the juxtaposition of two regions (extensions of helix 1) that are rich in basic residues (Table 1).

The two basic regions create a DNA-binding interface which interacts with the consensus sequence 5′-CANNTG-3′ (Table 2) (2,3). Since all bHLH proteins bind this core sequence, certain conserved amino acids in the basic region allow for recognition of the core consensus site. Specifically, a highly conserved glutamic acid (E9 in Table 1) contacts the 'CA' nucleotides of the consensus binding site. Other residues in the basic region of each bHLH protein dictate specificity to direct different bHLH proteins to different target sites. This is best evidenced by a mutant form of MyoD that contains a change in the basic region (L14 to R14) (Table 1) which allows it to recognize a c-Myc binding site (5′-CACGTG-3′) instead of a MyoD binding site (5′-CAGCTG-3′) (4). The altered amino acid contacts the fourth nucleotide of the core consensus binding site.

Table 1. Sequence alignment of the basic regions of select bHLH proteins

Protein	1	2	3	4	5	6	7	8	9	10	11	12	13
Pho4p	**K**	**R**	E	S	**H**	K	H	A	**E**	Q	A	**R**	**R**
Ino2p	**R**	**K**	W	K	**H**	V	Q	M	**E**	**K**	I	**R**	**R**
Ino4p	I	**R**	I	N	**H**	V	S	S	**E**	**K**	**K**	**R**	**R**
Cbf1p	**R**	**K**	D	S	**H**	K	E	V	**E**	**R**	**R**	**R**	**R**
Rtg1p	F	**K**	-	-	-	-	-	N	D	**R**	**K**	**R**	**R**
Rtg3p	**K**	**R**	E	F	**H**	N	A	V	**E**	**R**	**R**	**R**	**R**
Sgc1p	Q	**K**	Q	A	**H**	N	K	I	**E**	H	**R**	Y	**R**
c-Myc	**K**	**R**	R	T	**H**	N	V	L	**E**	**R**	Q	**R**	**R**
Max	K	**R**	A	H	**H**	N	A	L	**E**	**R**	**K**	**R**	**R**
MyoD	**R**	**R**	K	A	A	T	M	R	**E**	**R**	**R**	**R**	L
CONSENSUS	**R**	**K**			**H**				**E**	**K**	**K**	**R**	**R**
		K	**R**								**R**	**R**	

BASED ON YEAST PROTEINS

Another important feature of bHLH proteins is their ability to form multiple dimer combinations via the HLH domains. Max can dimerize with itself, c-Myc, Mad and Mxi1 (1). However, there is specificity in terms of the dimers that each bHLH protein can form since c-Myc does not appear to form dimers with MyoD. In the case of Max, the choice for dimerization is dependent on partner availability. The amount of Max is constant and abundant relative to c-Myc or Mad. The c-Myc protein is transiently expressed in growth factor-stimulated

*To whom correspondence should be addressed. Tel: +1 313 993 7816; Fax: +1 313 577 6891; Email: jlopes@sun.science.wayne.edu

cells and the Mad protein is expressed during differentiation. Thus, c-Myc/Max dimers are found in growth-stimulated cells and Mad/Max dimers are observed during differentiation.

The ability of bHLH proteins to form multiple dimer combinations is an efficient mechanism for regulation of gene expression since different dimers are likely to target different sets of genes. Therefore, it will be important to determine how bHLH proteins coordinate gene expression in response to extracellular factors. This requires understanding the regulation of bHLH function and identification of the bHLH target genes. It is also important to identify all of the binding partners for each bHLH protein. While it has been laborious to address these questions in higher eukaryotic systems, yeast now provides a unique opportunity to quickly examine these issues using genome-wide approaches. The answers to some of these questions are already becoming clear from past recent studies of yeast bHLH proteins. Here, we review the current knowledge of yeast bHLH proteins and their target genes as a prelude to the anticipated onslaught of information obtained from genome-wide strategies.

Table 2. Compilation of *S.cerevisiae* bHLH proteins and their targets

bHLH Protein	Target sequence[a]	Target genes	References
Pho4p	CACGTG	*PHO5, PHO8, PHO10, PHO11, PHO81, PHO84*	60-64
Ino2p/Ino4p	CATGTGAAAT and AATGTGAAAT	*INO1, CHO1, CHO2, OPI3, INO2, ADK1, ATP2, CCT1, CKI1, CPT1, CTR1, CYP51, EPT1, ERG1, ERG3, ERG12, FAS1, FAS2, FAS3, FPP1, ITR1, ITR2, MFA1G, OLE1, PHO5, PMA1, PSD1, PYK1, SLC1, SPE2, Ty1-H3, Ty3-1*	17
Ino4p/ ?	CATTTG ?	*CTR1, INO1*	34, 36
Ino4p/MSTF	?	*FAS1, INO2, INO4*	37
Cbf1p	CACGTG	*MET2, MET3, MET8, MET10, MET14, MET16, MET25, SAM2, TRP1, GAL2,* CEN's	36, 44, 46, 48
Rtg1p/Rtg3p	GGTCAC	*CIT2, CIT1, ACOI, IDH1, IDH2*	50, 54
Sgc1p	Unknown	*ENO1, ENO2, TDH, TPI1, PGK1, PGM1, PYK1*	35

[a]The core bHLH sequence is underlined.

THE *PHO4* SYSTEM

The *PHO4* (YFR034C) gene was the first yeast gene shown to encode a bHLH protein (5). Sequence comparison to a large number of mammalian and *Drosophila* bHLH proteins revealed similarity in the bHLH region. The product of the *PHO4* gene binds to a sequence that contains the core bHLH binding site, 5′-CACGTG-3′, which is found in the promoters of several genes involved in phosphate utilization (Table 2). The ability of Pho4p to bind this sequence was initially demonstrated by mobility shift assays and footprinting experiments using the promoter of the *PHO5* gene. This promoter contains two Pho4p binding sites (UAS_p1 and UAS_p2) which flank a Pho2p binding site (6). The binding of Pho4p to these sites is required for the ~1000-fold induction of *PHO5* gene expression in response to phosphate starvation. Binding of Pho4p to UASp1 and to a lesser degree UASp2 is enhanced cooperatively by binding of Pho2p (7). Pho2p binding is believed to disrupt intra-molecular interactions within an internal repression domain of Pho4p, thereby increasing access to the transcriptional activation domain (7,8).

Table 3. Mutations of a conserved Glu (E9) in the basic domains of bHLH proteins

Protein	Mutation	Phenotype	Reference
Pho4p	Glu_Gln	No defect in DNA binding	9
Pho4p	Glu_Asp Glu_Asn Glu_Leu	Defective in DNA binding	9
Cbf1p	Glu_Gly Glu_Val	Methionine auxotroph, defective in binding CDEI element	38
Cpf1p	Glu_Ala	Methionine prototroph, defective in binding CDEI element	37
Sgc1p	Glu_Gln	Suppressor of *gcr1* phenotypes (*i.e.*, no growth on glucose and reduced expression of *ENO1-lacZ* gene)	34

Curiously, the core binding sites for Pho4p and Cbf1p (another bHLH protein) are identical (Table 2). Although both proteins bind the same sequence there is no evidence for overlapping functions. Over-expression of the *PHO4* gene does not complement the methionine auxotrophy of a *cbf1* mutant strain and over-expression of the *CBF1* gene does not regulate *PHO5* gene expression (9). An explanation for this apparent discrepancy lies in the fact that the specificity for binding is dictated by sequences that flank the core bHLH binding site. For example, a T flanking 5′ to the core sequence inhibits the binding of Pho4p but not Cbf1p (9). It is therefore not surprising to find that the genes regulated by Pho4p contain UAS elements that lack the T residue and that Cbf1p target sites typically contain the T residue (9). The inability of Pho4p to recognize the core sequence preceded by the T residue is known to be due to the presence of a specific Glu residue in the basic region (E3 in Table 1) (9). When this Glu is changed to the Cbf1p counterpart (D3 in Table 1), the mutant Pho4p is now able to recognize the Cbf1p optimal binding site (core sequence flanked by the T).

As is the case with c-Myc and MyoD, a highly conserved Glu (E9 in Table 1) in the basic region is important for Pho4p binding to the core sequence. Analysis of the protein crystal structure demonstrated that this conserved Glu binds to the nucleotides ‘CA’ of the consensus binding site (10). Changing this Glu residue to an Asp, Asn or Leu residue completely inhibits binding to the core sequence (Table 3). Interestingly,

changing it to Gln does not affect the ability of Pho4p to bind the core sequence. This is curious given that this Glu residue is conserved in virtually every known bHLH protein (for example, see Table 1).

The Pho4p regulatory cascade is also one of the best understood regulatory systems in yeast. The ability of Pho4p to function as a transcriptional activator depends on its phosphorylation state. When cells are grown in the presence of high concentrations of inorganic phosphate, Pho4p is hyper-phosphorylated (11). Pho4p is phosphorylated by a complex, which is encoded by the *PHO80* and *PHO85* genes. This kinase complex bears remarkable similarity to cyclin/cyclin-dependent kinase (CDK) complexes (11,12). Immunoprecipitation studies show that Pho80p (cyclin) and Pho85p (CDK) interact with Pho4p as a complex (11). The hyper-phosphorylated Pho4p interacts with the export protein Msn5p, which shuttles Pho4p into the cytoplasm (13). Once in the cytoplasm, Pho4p is unable to activate transcription of *PHO5*. As inorganic phosphate becomes limiting, Pho4p becomes de-phosphorylated. The de-phosphorylated Pho4p is transported back into the nucleus by the import protein Pse1p. There are six serine–proline dipeptides within Pho4p designated SP1–6. Mutational analysis reveals that the serine–proline dipeptides SP2 and SP3 are specifically required for export into the cytoplasm while SP4 is required for import into the nucleus (14).

The phosphorylation of Pho4p in the nucleus under inducing conditions is prevented by the product of the *PHO81* gene. The role of this gene was initially suggested by its similarity to a mammalian CDK inhibitor (15). The similarity is restricted to a region that contains ankyrin repeats. Co-immunoprecipitation experiments show that Pho81p interacts with the Pho80p:Pho85p complex and inhibits its activity as a kinase. It is also known that the region containing the ankyrin repeats is sufficient for its inhibitory function (15,16). Consistent with these results is the observation that *PHO81* gene expression is also regulated in response to phosphate. Thus, *PHO81* gene expression is induced when cells are grown in limiting phosphate (inducing conditions) which reduces the phosphorylation of Pho4p (active state). In support of this, experiments where *PHO81* expression is driven by the *GAL1* promoter show that induction of *PHO5* gene expression requires both induction of *PHO81* expression and phosphate starvation (16). Now that the components of this regulatory cascade are in hand, it should be possible to define the signaling mechanism.

THE *INO2/INO4* SYSTEM

The *INO2* (*SCS2*, *DIE1*, YDR123C) and *INO4* (YOL108C) genes are required for derepression of phospholipid biosynthetic gene expression in response to inositol deprivation. Strains containing mutations in either of these two genes are unable to derepress transcription of the *INO1* and *CHO1* (17,18) phospholipid biosynthetic genes when cells are grown in the absence of inositol. The inability to derepress expression of the *INO1* gene results in the characteristic inositol auxotrophy which is the hallmark of *ino2* and *ino4* mutant strains (17,18). The *INO2* and *INO4* genes were cloned by complementation of this inositol auxotrophy (19,20). DNA sequence analyses of the cloned genes revealed a high degree of similarity to the bHLH region of the Myc family of proteins (Table 1) (19,20).

While the products of the *INO2* and *INO4* genes have been shown to form a heterodimer *in vivo* (21) and *in vitro* (22), studies using the yeast two hybrid system suggest that neither protein is capable of homodimerizing (23). A functional analysis of the two proteins reveals that Ino2p contains a transcriptional activation domain (N-terminal) whereas Ino4p does not contain this type of domain (23). Collectively, these observations suggest that Ino4p is required for dimerization with Ino2p, which functions to activate transcription. This organization is reminiscent of the mammalian Myc and Max proteins where Myc has the transcriptional activation function but must dimerize with Max to bind target sequences (1). Another similarity between these two systems is the observation that, like Myc, transcription of the *INO2* gene is regulated and that the amount of *INO2* expression is limiting relative to *INO4* expression (24). It is presumed that the excess of Max over Myc permits it to dimerize with other proteins (Mad and Mxi1) in the absence of Myc (1). If the similarities persist between these two systems, it would suggest that Ino4p also forms multiple dimer combinations with other yeast bHLH proteins.

The mechanism for derepression of the phospholipid biosynthetic genes in response to inositol deprivation has not been established. However, it is clear that regulation of transcription of the *INO2* regulatory gene must play an important role in this response. It has been shown that expression of a *cat* reporter gene driven by the *INO2* promoter is regulated in response to inositol in a pattern that is identical to that of Ino2p target genes such as *INO1* and *CHO1* (24). That is, *INO2-cat* expression is maximal when cells are grown in the absence of inositol and repressed when grown in the presence of inositol. As is the case with the Ino2p target genes, expression of the *INO2-cat* gene is regulated by the *INO2* and *INO4* genes. However, the regulation of *INO2* expression cannot be solely accountable for the response to inositol since expression of the *INO1* and *CHO1* genes is sensitive to inositol even when *INO2* expression is driven by the *GAL1* promoter (inositol-insensitive) (25). This inositol-responsive repression is likely to be mediated by the product of the *OPI1* gene (17,18). The mechanism of action of the *OPI1* gene is still under investigation. However, since *OPI1* also represses transcription of the *INO2* gene (24), derepression of the target genes necessitates prior derepression of *INO2* gene expression. The current model for regulation of phospholipid biosynthetic gene expression has two components. The first component is that derepression of *INO2* gene expression is required for derepression of target gene expression in the absence of inositol. The second component is that the *OPI1* negative regulatory gene represses transcription of the *INO2* gene and the Ino2p target genes in the presence of inositol.

The product of the *SIN3* gene is another factor involved in the regulation of the phospholipid biosynthetic gene expression. This gene functions as a negative regulator of the phospholipid biosynthetic genes (26). Although its mechanism of action is not known, it is clear that Sin3p exerts its effects through the UAS_{INO} element suggesting that Sin3p might interact with the Ino2p/Ino4p bHLH complex (27). It is intriguing that the human Sin3p functions as a repressor by interacting with Mad and Mxi1 when they are complexed with Max (28,29).

It has become increasingly obvious that the function of the *INO2* and *INO4* gene products extends beyond the scope of phospholipid biosynthesis. Several genes, whose expression is regulated by either *INO2*, or in response to inositol, have been

identified (17,18) (Table 2). The majority of these genes function in some aspect of membrane biogenesis, however, a small number of genes (e.g. *ADK1*, *PHO5*, *PYK1*, *PMA1* and *MFA1G*) do not appear to have any role in this process. The pleiotropic phenotype of *ino2* mutant strains also suggests a more general role for Ino2p. These strains have defects in nuclear segregation, bud formation and sporulation, display an aberrant oversized morphology and over-express the *PIS1* gene (30,31).

An examination of the promoters for each of the genes whose expression is inositol-responsive, or *INO2*-dependent, identified a binding site for the Ino2p:Ino4p heterodimer (consensus: 5′-CATGTGAAAT-3′) (Table 2). This element (called UAS_{INO}/ICRE) is necessary and sufficient to bind the Ino2p:Ino4p heterodimer and for inositol-specific regulation (32,33). Predictably, the first six nucleotides of the consensus Ino2p:Ino4p binding site are the core element required for the binding of bHLH proteins. However, an A residue can substitute for the C residue at the first position of the UAS_{INO} sequence (Table 2) (32,33). It is also known that the two nucleotides 5′ to the UAS_{INO} element and the nucleotides at positions 7 and 8 play a role in optimizing its function (32).

The potential for Ino4p to form dimers with multiple proteins (discussed above) raises a question concerning the DNA binding specificity of different Ino4p complexes. One answer to this question may be that different complexes recognize variations of the UAS_{INO} element. Consistent with this hypothesis, an *INO4*-dependent/*INO2*-independent promoter element in the *CTR1/HMN1* promoter has the sequence, 5′-CATTTG-3′ (Table 2) (34). This same sequence has been shown to function as a weak inositol-unresponsive UAS element when fused to a *lacZ* reporter gene (32).

Another question that needs to be addressed is: what are the other binding partners for Ino4p? Mutant alleles of several of the known yeast bHLH genes are not inositol auxotrophs suggesting that these proteins do not play a direct role in phospholipid biosynthesis [B.P.Ashburner and J.M.Lopes, unpublished data; (35)]. However, both the *CTR1/HMN1* and *INO1* genes have been shown to contain an *INO2*-independent, *INO4*-dependent, UAS element (34,36) suggesting that Ino4p must have an alternate partner. Moreover, another study suggests that Ino4p interacts with a myristoylation-sensitive transcription factor (MSTF) to regulate expression of the *INO2*, *INO4* and *FAS1* genes (Table 2) (37). This study showed that expression of the *INO2* and *FAS1* genes is elevated in a *nmt1* mutant strain (temperature-sensitive defect in N-myristoylation) while *INO4* expression is decreased. However, at elevated temperatures expression of the *FAS1* gene becomes *INO2*-independent in a *nmt1* mutant strain (37).

THE *CBF1* SYSTEM

The *CBF1* (*CPF1*, *CEP1*, YJR060W) gene encodes a member of the bHLH family which has been designated as Cbf1p, Cpf1p and Cp1p. These proteins were initially identified by their ability to specifically bind a region present in centromeres called CDEI (5′-RTCACRTG-3′) (38,39). The multiple names for this gene and its protein product are partly due to the fact that two proteins of different sizes (39 and 64 kDa) were found to bind the CDEI element suggesting that they might be different proteins. However, the cloning of the *CBF1* gene confirmed that these proteins were encoded by the same gene and that the observed difference in sizes of the proteins were likely due to proteolytic degradation of the 64 kDa protein to the 39 kDa form (39). The *CBF1* gene encodes a 351 amino acid protein with a predicted molecular weight of 37 kDa (38,39). However, when transcribed and translated *in vitro* this gene yields a product that migrates as a 60 kDa protein on an SDS–PAGE gel. The altered mobility is presumably due to its unusual amino acid composition, which includes regions that are highly negatively and positively charged (39).

The identification of the *CBF1* gene facilitated genetic analyses, which show that Cbf1p is required for chromosomal segregation. Strains bearing *cbf1*Δ alleles display several phenotypes including slow growth, increased chromosomal loss, sensitivity to microtubule-disrupting drugs (e.g. thiabendazole and benomyl) and methionine auxotrophy (36,39). The methionine-dependent growth was unique because these same strains do not have a growth requirement for tryptophane, adenine, histidine, leucine or uracil (36,39). Since Cbf1p was originally identified as a centromere-binding protein, the methionine auxotrophy was surprising and suggested that Cbf1p has a role in transcriptional regulation of the *MET* genes. Thus, Cbf1p is a unique member of the bHLH family because it has two distinct functions, chromosome segregation and transcriptional control.

A clear aspect of Cbf1p function is that the centromere function is mechanistically different from the transcription function. This is supported by the existence of mutants that are defective in either centromere function or are methionine auxotrophs (40,41). Moreover, it has been observed that mutant alleles of the *SPT21*, *SIN3*, *CCR4* and *RPD3* genes suppress the methionine auxotrophy of a *cbf1* mutant strain but *spt21* and *sin3* mutant alleles do not suppress the chromosome loss phenotype (42). Thus, the role of Cbf1p in centromere function and transcription must be mechanistically distinct and may involve different Cbf1p domains.

Several lines of investigation suggest that the bHLH domain and an adjacent C-terminal region encoding a leucine zipper domain are required for Cbf1p function. Mutational studies reveal that the bHLH region and the leucine zipper region are both required for Cbf1p dimerization and function (41,43,44). Another observation that ascribes significance to the bHLH–leucine zipper (zip) region is that the *CBF1* gene from *Kluyveromyces lactis* will complement the methionine auxotrophy of a *Saccharomyces cerevisiae cbf1*Δ allele (45). This is significant because the only appreciable homology between the *CBF1* genes of the two organisms resides in the bHLH–zip region (86% identity). A third line of evidence shows that Cbf1p generated by translation *in vitro* binds DNA as a dimer and that the leucine zipper region is required for dimerization and binding to the CDE1 element (43).

The sequence requirements of the CDEI element have been established by examining the effect of point mutants using a chromosomal loss assay (46). These experiments establish that the optimal sequence requirements for the CDEI element are 5′-CACGTG-3′ (Table 2). As expected this sequence includes the consensus bHLH binding site. The CDEI sequence requirements have not been investigated with respect to either the binding affinity of Cbf1p or the methionine auxotrophy, although several genes have been identified which have CDEI elements in their promoters (47–49) (Table 2). The consensus

derived from aligning the CDEI elements found in promoter sequences is indistinguishable from the CDEI sequence requirements obtained by assaying for centromere function.

The specific function of Cbf1p in expression of the methionine biosynthetic genes remains unresolved. Initial reports, using northern blot and primer extension analyses failed to observe any defects in expression of *MET25*, *GAL2* and *TRP1* genes in cells containing a *cbf1* null mutant allele even though the promoters of all three genes contain CDEI elements (39,40). Other studies show a pronounced effect of a *cbf1* null mutant allele on the derepression rates of *MET3*, *MET10*, *MET14*, *MET16* and *MET25* gene expression (49). These experiments demonstrated that *CBF1* is not absolutely required for *MET3* and *MET25* expression but rather required for complete and efficient derepression. However, the former studies examined *MET25* expression immediately after methionine was removed from the media and then again at a very late stage of growth (40). Thus, they were unable to detect the lag in derepression seen in the latter studies (50) and therefore concluded that Cbf1p had no effect on *MET25* expression. However unlike the *MET25* gene, *CBF1* does seem to be absolutely required for *MET10*, *MET14* and *MET16* expression (50). Consistent with these studies, another report demonstrated that *CBF1* is required for expression of a *lacZ* reporter gene under the control of the *MET16* CDE1 element (48).

Although Cbf1p is a positive regulator of the methionine biosynthetic genes, it does not appear to function as a classic transcriptional activator since a lexA–Cbf1p fusion fails to induce transcription of a reporter gene (50). Instead Cbf1p appears to influence chromatin structure at both the centromeres and the promoters of CDEI-containing genes (e.g. *MET25*, *MET16*, *TRP1*, *GAL2*) (39,47,51). Cbf1p creates a nucleosome-free region surrounding the CDEI elements (47).

The fact that a *cbf1*Δ mutant does not completely eliminate *MET3*, *MET10* and *MET25* gene expression suggests that other factors may be required. This is borne out by the fact that each of the CDEI-containing promoters is dependent on other transcriptional regulatory proteins such as *GCN4* (48) and *MET4* (50) for their expression. In the case of the *MET16* gene, Cbf1p is nearly 6-fold more effective when the CDEI site is located upstream of the Gcn4p binding site (natural location) relative to when it is placed downstream of the Gcn4p binding site (48). Therefore, the function of Cbf1p may be to recruit Gcn4p to the *MET16* promoter.

The idea that Cbf1p recruits other transcriptional activators to the promoters of the *MET* genes is further supported by studies involving the leucine zipper proteins Met4p and Met28p. Mobility shift assays demonstrated that Met4p and Met28p along with Cbf1p bind to the UAS_{MET16} (51). However, neither Met4p nor Met28p can bind the UAS_{MET16} without Cbf1p. The association and dissociation rates of Cbf1p in the presence and absence of Met28p, suggest that Met28p is required for maintaining Cbf1p on the DNA. These experiments suggest that Cbf1p dimers recruit Met28p and the transcriptional activator protein, Met4p to the CDEI element. Once the Met28p:Met4p is bound to DNA, Met4p activates transcription, while Met28p helps maintain the stability of the complex (51,52).

In spite of all the work that has been done, the results of two previous studies need to be addressed. One study using mutations in the basic region shows that methionine-independent growth does not require DNA binding but that centromere function does require DNA binding (Table 3) (40). However, another study shows that several mutations in the bHLH–zip region which abolish DNA binding concomitantly results in methionine auxotrophy (Table 3) (41). It is obvious that more research is needed to understand these results and the complete functional role of Cbf1p.

THE *RTG1/RTG3* SYSTEM

The *RTG1* (YOL067C) and *RTG3* (YBL103C) genes were identified as regulators of *CIT2* gene expression (53,54). Expression of the nuclear-encoded *CIT2* gene is subject to regulation in response to the functional state of the mitochondria, a process that has been termed retrograde regulation (53). In ρ^o petite strains (lacking mitochondrial DNA) *CIT2* expression is induced 6–30-fold relative to isogenic respiratory competent strains (ρ^+). This observation was employed to isolate mutants that fail to induce *CIT2* expression in ρ^o strains. Mutants were isolated by screening for colonies that failed to express a *CIT2–lacZ* reporter gene in a ρ^o mutant background. To clone *RTG1* and *RTG3*, the respective mutant strains were transformed with a YCp50-based yeast genomic library. Transformants were screened for their ability to restore expression of the *CIT2–lacZ* reporter gene. Subsequent sequence analysis revealed a high degree of similarity to several bHLH proteins.

Western blot analysis demonstrated that Rtg1p levels are not increased in a ρ^o strain (55). However, *RTG3* mRNA steady-state levels are derepressed in a ρ^o strain (54). This regulation is reminiscent of *INO2*/*INO4*, suggesting that the constitutive expression of *RTG1* may allow it to form multiple dimer combinations.

The *cis*-acting sequence (UAS_r; R-box) required for the induction of *CIT2* expression is defined by the sequence 5′ -GGTCAC-3′, which does not conform to the classic E-box sequence that binds other bHLH proteins (Table 2). The unusual DNA binding sequence requirements are not entirely surprising given that the basic region for Rtg1p deviates from the consensus basic regions of other yeast bHLH proteins (Table 1). The *CIT2* promoter contains two inverted copies of the R-box separated by 28 bp. While both R-boxes can form an *RTG1*/*RTG3*-dependent complex, the two sites act synergistically *in vivo* (54).

Activation by a Gal4p–Rtg3p fusion protein can occur in the absence of *RTG1* and *RTG2*, suggesting that Rtg3p is responsible for activation (56). However, a Gal4p–Rtg1p fusion protein cannot activate transcription of the *lac*Z reporter gene in the absence of *RTG3*. This evidence suggests that Rtg1p is responsible for recruiting the activation domain of Rtg3p to the UAS_r element, allowing for transcriptional activation of the *CIT2* gene.

Recent experiments have discovered that the tricarboxylic acid cycle genes *CIT1*, *ACO1*, *IDH1* and *IDH2* also require *RTG1* and *RTG3* for full expression in ρ^o strains (Table 2) (57). However unlike *CIT2*, these genes are also expressed in ρ^+ strains and require the *HAP2*, *HAP3* and *HAP4* genes for expression.

THE *SGC1* SYSTEM

The *SGC1* (*TYE7*, YOR344C) gene is required for expression of glycolytic genes in yeast (Table 2) (35). Expression of the glycolytic genes (e.g. *ENO1* and *ENO2*) normally requires the product of the *GCR1* positive regulatory gene, which is also required for growth on media containing glucose as a carbon source. A genetic selection for mutants that simultaneously restored both growth on glucose and expression of an *ENO1–lacZ* reporter gene to a *gcr1* mutant strain yielded two dominant *SGC1* alleles. The dominant mutant and wild-type alleles of the *SGC1* gene were cloned and sequence inspection identified this gene as a member of the bHLH family (35). Northern blot analysis shows that the *SGC1-1* mutant allele, in a *gcr1* background, restores expression of several glycolytic genes although the most pronounced effect is on the *ENO1*, *PGK1* and *PYK1* genes (35). The *SGC1* gene does encode a positive regulator of glycolytic gene expression since an *sgc1* mutant strain is defective in expression of the *ENO1* gene and its gene product (enolase) (35). It is also clear that Sgc1p and Gcr1p function to stimulate expression of the enolase genes (*ENO1* and *ENO2*) through parallel pathways since a *gcr1*, *sgc1* double mutant strain is more defective in enolase gene expression than either of the single mutant strains. What is not known is whether Sgc1p binds a consensus bHLH binding site. The sequence similarity of the basic region to other bHLH proteins suggests that this is likely to be the case (Table 1). Interestingly, one of the dominant mutant alleles contained a Glu→Gln change in the basic region (Table 3). As discussed above, this Glu residue is highly conserved among the yeast bHLH proteins but a change to Gln did not eliminate the DNA binding function of Pho4p (9).

POTENTIAL *S.cerevisiae* bHLH PROTEINS

The completion of the yeast genome sequencing effort has revealed new putative bHLH proteins. Two yeast genes of known function encode potential bHLH proteins. *HMS1* encodes a probable transcription factor with similarities to the Myc family. Over-expression of *HMS1* results in filamentous growth and it is therefore believed to be a regulator of pseudohyphal differentiation (58). *PHD1* encodes a transcription factor involved in regulating filamentous growth. The over-expression of *PHD1* causes invasive filamentous growth in rich media (59). Finally, three ORFs encode potential bHLH proteins: YGR290W, YLR002C and YPL165C. However, it is likely that as our ability to predict protein structure and function improves, the number of bHLH proteins in yeast will increase.

FUTURE DIRECTIONS

As we have seen, bHLH proteins play an integral role in the regulation of a number of important pathways in yeast. However, there are important questions that remain to be answered. First, there is currently little knowledge about the signaling mechanisms that regulate these yeast bHLH systems. There is also currently no published evidence for bHLH proteins forming multiple heterodimer combinations. However, the constitutive expression of both *INO4* and *RTG1* suggests that multiple heterodimer formation is possible. The discovery of multiple bHLH heterodimers will likely elucidate novel regulatory pathways in yeast. Therefore, this molecular and genetically tractable organism will continue to lend great insight into the function and regulation of bHLH proteins.

ACKNOWLEDGEMENTS

The authors thank Kyle Gardenour and Mohan Rao Kaadige for comments on the manuscript. J.M.L. is supported by a grant from the American Cancer Society (RPG-97-002-01).

REFERENCES

1. Amati,B. and Land,H. (1994) *Curr. Opin. Genes Dev.*, **4**, 102–108.
2. Ma,P.C.M., Rould,M.A., Weintraub,H. and Pabo,C.O. (1994) *Cell*, **77**, 451–459.
3. Ferré-D'Amaré,A.R., Prendergast,G.C., Ziff,E.B. and Burley,S.K. (1993) *Nature*, **363**, 38–44.
4. Blackwell,T.K., Huang,J., Ma,A., Kretzner,L., Alt,F.W., Eisenman,R.N. and Weintraub,H. (1993) *Mol. Cell. Biol.*, **13**, 5216–5224.
5. Berben,G., Legrain,M., Gilliquet,V. and Hilger,F. (1990) *Yeast*, **6**, 451–454.
6. Vogel,K., Hörz,W. and Hinnen,A. (1989) *Mol. Cell. Biol.*, **9**, 2050–2057.
7. Barbaric,S., Münsterkötter,M., Goding,C. and Hörz,W. (1998) *Mol. Cell. Biol.*, **18**, 2629–2639.
8. Shao,D., Creasy,C.L. and Bergman,L.W. (1996) *Mol. Gen. Genet.*, **251**, 358–364.
9. Fisher,F. and Goding,C.R. (1992) *EMBO J.*, **11**, 4103–4109.
10. Shimizu,T., Toumoto,A., Ihara,K., Shimizu,M., Kyohoku,Y., Ogawa,N., Oshima,Y. and Haloshima,T. (1997) *EMBO J.*, **16**, 4689–4697.
11. Kaffman,A., Herskowitz,I., Tjian,R. and O'Shea,E.K. (1994) *Science*, **263**, 1153–1156.
12. Hirst,K., Fisher,F., McAndrew,P.C. and Goding,C.R. (1994) *EMBO J.*, **13**, 5410–5420.
13. Kaffman,A., Rank,N.M., O'Neill,E.M., Huang,L.S. and O'Shea,E.K. (1998) *Nature*, **396**, 482–486.
14. Komeili,A. and O'Shea,E.K. (1999) *Science*, **284**, 977–980.
15. Schneider,K.R., Smith,R.L. and O'Shea,E.K. (1994) *Science*, **266**, 122–126.
16. Ogawa,N., Noguchi,K.-I., Sawai,H., Yamashita,Y., Yompakdee,C. and Oshima,Y. (1995) *Mol. Cell. Biol.*, **15**, 997–1004.
17. Greenberg,M.L. and Lopes,J.M. (1996) *Microbiol. Rev.*, **60**, 1–20.
18. Paltauf,F., Kohlwein,S.D. and Henry,S.A. (1992) In Jones,E.W., Pringle,J.R. and Broach,J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces: Gene Expression*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 415–499.
19. Hoshizaki,D.K., Hill,J.E. and Henry,S.A. (1990) *J. Biol. Chem.*, **265**, 4736–4745.
20. Nikoloff,D.M., McGraw,P. and Henry,S.A. (1992) *Nucleic Acids Res.*, **20**, 3253.
21. Nikoloff,D.M. and Henry,S.A. (1994) *J. Biol. Chem.*, **269**, 7402–7411.
22. Ambroziak,J. and Henry,S.A. (1994) *J. Biol. Chem.*, **269**, 15344–15349.
23. Schwank,S., Ebbert,R., Rautenstrauß,K., Schweizer,E. and Schüller,H.-J. (1995) *Nucleic Acids Res.*, **23**, 230–237.
24. Ashburner,B.P. and Lopes,J.M. (1995) *Mol. Cell. Biol.*, **15**, 1709–1715.
25. Ashburner,B.P. and Lopes,J.M. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 9722–9726.
26. Hudak,K.A., Lopes,J.M. and Henry,S.A. (1994) *Genetics*, **136**, 475–483.
27. Slekar,K.H. and Henry,S.A. (1995) *Nucleic Acids Res.*, **23**, 1964–1969.
28. Ayer,D.E., Lawrence,Q.A. and Eisenman,R.N. (1995) *Cell*, **80**, 767–776.
29. Schreiber-Agus,N., Chin,L., Chen,K., Torres,R., Rao,G., Guida,P., Skoultchi,A.I. and DePinko,R.A. (1995) *Cell*, **80**, 777–786.
30. Anderson,M.S. and Lopes,J.M. (1996) *J. Biol. Chem.*, **271**, 26596–26601.
31. Hammond,C.I., Romano,P., Roe,S. and Tontonoz,P. (1993) *Cell. Mol. Biol. Res.*, **39**, 561–577.
32. Bacchawat,N., Ouyang,Q. and Henry,S.A. (1995) *J. Biol. Chem.*, **270**, 25087–25095.
33. Schüller,H.-J., Richter,K., Hoffmann,B., Ebbert,R. and Schweizer,E. (1995) *FEBS Lett.*, **370**, 149–152.
34. Li,Z. and Brendel,M. (1993) *Mol. Gen. Genet.*, **241**, 680–684.
35. Nishi,K., Park,C.S., Pepper,A.E., Eichinger,G., Innis,M.A. and Holland,M.J. (1995) *Mol. Cell. Biol.*, **15**, 2646–2653.
36. Swift,S. and McGraw,P. (1995) *Nucleic Acids Res.*, **26**, 2865–2872.

37. Cok,S.J., Martin,C.G. and Gordon,J.I. (1995) *Nucleic Acids Res.*, **23**, 1426–1433.
38. Cai,M. and Davis,R.W. (1990) *Cell*, **61**, 437–446.
39. Mellor,J., Jiang,W., Funk,M., Rathjen,J., Barnes,C.A., Hinz,T., Hegemann,J.H. and Philippsen,P. (1990) *EMBO J.*, **9**, 4017–4026.
40. Mellor,J., Rathjen,J., Jiang,W. and Dowell,S.J. (1991) *Nucleic Acids Res.*, **19**, 2961–2969.
41. Foreman,P.K. and Davis,R.W. (1993) *Genetics*, **135**, 287–296.
42. McKenzie,E.A., Kent,N.A., Dowell,S.J., Moreno,F., Bird,L.E. and Mellor,J. (1993) *Mol. Gen. Genet.*, **240**, 374–386.
43. Dowell,S.J., Tsang,J.S.H. and Mellor,J. (1992) *Nucleic Acids Res.*, **20**, 4229–4236.
44. Masison,D.C., O'Connell,K.F. and Baker,R.E. (1993) *Nucleic Acids Res.*, **21**, 4133–4141.
45. Mulder,W., Winkler,A.A., Scholten,I.H.J.M., Zonneveld,B.J.M., De Winde,J.H., Steensma,H.Y. and Grivell,L.A. (1994) *Curr. Genet.*, **26**, 198–207.
46. Niedenthal,R., Stoll,R. and Hegemann,J.H. (1991) *Mol. Cell. Biol.*, **11**, 3545–3553.
47. Kent,N.A., Tsang,J.S.H., Crowther,D.J. and Mellor,J. (1994) *Mol. Cell. Biol.*, **14**, 5229–5241.
48. O'Connell,K.F., Surdin-Kerjan,Y. and Baker,R.E. (1995) *Mol. Cell. Biol.*, **15**, 1879–1888.
49. Kuras,L. and Thomas,D. (1995) *FEBS Lett.*, **367**, 15–18.
50. Thomas,D., Jacquemin,I. and Surdin-Kerjan,Y. (1992) *Mol. Cell. Biol.*, **12**, 1719–1727.
51. Kuras,L., Barbey,R. and Thomas,D. (1997) *EMBO J.*, **16**, 2441–2451.
52. Kuras,L. and Thomas,D. (1995) *Mol. Cell. Biol.*, **15**, 208–216.
53. Liao,X. and Butow,R.A. (1993) *Cell*, **72**, 61–71.
54. Jia,Y., Rothermel,B., Thornton,J. and Butow,R.A. (1997) *Mol. Cell. Biol.*, **17**, 1110–1117.
55. Rothermel,B.A., Shyjan,A.W., Etheredge,J.L. and Butow,R.A. (1995) *J. Biol. Chem.*, **270**, 29476–29482.
56. Rothermel,B.A., Thornton,J.L. and Butow,R.A. (1997) *J. Biol. Chem.*, **272**, 19801–19807.
57. Liu,Z. and Butow,R.A. (1999) *Mol. Cell. Biol.*, **19**, 6720–6728.
58. Lorenz,M.C. and Heitman,J. (1998) *Genetics*, **150**, 1443–1457.
59. Gimeno,C.J. and Fink,G.R. (1994) *Mol. Cell. Biol.*, **14**, 2100–2112.
60. Bajwa,W., Meyhack,B., Rudolph,H., Schweingruber,A.-M. and Hinnen,A. (1984) *Nucleic Acids Res.*, **12**, 7721–7739.
61. Bun-ya,M., Nishimura,M., Harashima,S. and Oshima,Y. (1997) *Mol. Cell. Biol.*, **11**, 3229–3238.
62. Plankert,U., Purwin,C. and Holzer,H. (1991) *Eur. J. Biochem.*, **196**, 191–197.
63. Venter,U. and Hörz,W. (1989) *Nucleic Acids Res.*, **17**, 1353–1369.
64. Yoshida,K., Ogawa,N. and Oshima,Y. (1989) *Mol. Gen. Genet.*, **217**, 40–46.