UvA-DARE (Digital Academic Repository)

Sequence of the HAP3 transcription factor of *Kluyveromyces lactis* predicts the presence of a novel 4-cysteine zinc-finger motif

Mulder, W.; Scholten, I.H.J.M.; de Boer, R.; Grivell, L.A.

DOI

[10.1007/BF00279755](https://doi.org/10.1007/BF00279755)

Publication date

1994

Published in

MGG. Molecular & General Genetics

[Link to publication](#)

Citation for published version (APA):

Mulder, W., Scholten, I. H. J. M., de Boer, R., & Grivell, L. A. (1994). Sequence of the HAP3 transcription factor of *Kluyveromyces lactis* predicts the presence of a novel 4-cysteine zinc-finger motif. *MGG. Molecular & General Genetics*, 245, 96-106.
<https://doi.org/10.1007/BF00279755>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

ORIGINAL PAPER

Wietse Mulder · Inge H. J. M. Scholten
René W. de Boer · Leslie A. Grivell

Sequence of the HAP3 transcription factor of *Kluyveromyces lactis* predicts the presence of a novel 4-cysteine zinc-finger motif

Received: 27 December 1993 / Accepted: 23 March 1994

Abstract The *Kluyveromyces lactis* homologue of the *Saccharomyces cerevisiae* HAP3 gene was isolated by functional complementation of the respiratory-deficient phenotype of the *S. cerevisiae* *hap3::HIS4* strain SHY40. The *KlHAP3* gene encodes a protein of 205 amino acids, of which the central B-domain of 90 residues is highly homologous to HAP3 counterparts of *S. cerevisiae* and higher eukaryotes. The protein contains a novel 4-cysteine zinc-finger motif and we propose by analogy that all other homologous HAP3 proteins contain the same motif, with the position containing the third cysteine being occupied by a serine residue. In contrast to the situation in *S. cerevisiae*, disruption of the *KlHAP3* gene in *K. lactis* does not result in a respiratory-deficient phenotype and the growth of the null strain is indistinguishable from wild type. There is also no effect on the expression of the carbon source-regulated *KlCYC1* gene, suggesting either a different role for the HAP2/3/4 complex, or the existence of a different mechanism of carbon source regulation. Sequence verification of the *S. cerevisiae* HAP3 locus reveals that, just as in *K. lactis*, a long open reading frame (ORF) is present upstream of the HAP3 gene. These highly homologous ORFs are predicted to have at least eight membrane-spanning fragments, but do not show significant homology to any known sequence present in databases. The *ScORFX* gene is transcribed in the opposite direction to *ScHAP3*, but, in contrast to an earlier report by Hahn et al. (1988), the transcripts of the two genes do not overlap. The model proposed by these authors, in which the *ScHAP3* gene is regulated by an anti-sense non-coding mRNA, is therefore not correct.

Key words HAP3 · *Saccharomyces cerevisiae* · *Kluyveromyces lactis* · Zinc finger · Carbon source regulation

Introduction

Mitochondrial biogenesis in the yeast *Saccharomyces cerevisiae* is regulated primarily at the level of transcription (Grivell 1989). Since respiratory function is not required under fermentative conditions, many genes coding for mitochondrial proteins are subject to carbon catabolite repression. Thus, genes coding for components of the respiratory chain, like *CYC1* (Guarente et al. 1984) and *QCR8* (DeWinde and Grivell 1992) are repressed when the cells are grown on glucose and transcription is induced roughly ten fold when the cells are grown on a non-fermentable carbon source. The main transcriptional regulator of this induction is thought to be the heterotrimeric transcription-activation complex HAP2/3/4. The genes for the proteins of this complex were isolated via a selection procedure that exploited the ability of the subunits of the complex to bind to UAS_{UP1} of the *CYC1* gene and to activate transcription (Pinkham and Guarente 1985; Hahn et al. 1988; Forsburg and Guarente 1989). Based on in vitro mobility-shift experiments, Olesen et al. (1987) concluded that HAP2 and HAP3 bind to the DNA in an interdependent manner. HAP4 protein does not bind DNA directly, but augments the binding of HAP2 and HAP3 (Forsburg and Guarente 1989). Mutations in any of the three HAP genes lead to a non-functional complex, with a consequently diminished induction of genes coding for mitochondrial proteins. Correspondingly, mutant cells are unable to grow on non-fermentable carbon sources. In *S. cerevisiae* the HAP2 and HAP3 genes are expressed constitutively at a low level, whereas the expression of the HAP4 gene is glucose repressible. This suggests that HAP4 is a key regulator of the HAP2/3/4 complex in its response to carbon

Communicated by C. Hollenberg

W. Mulder · I. H. J. M. Scholten · R. W. de Boer
L. A. Grivell (✉)
Section for Molecular Biology, Institute for Molecular Cell Biology,
Biocentrum Amsterdam, Kruislaan 318, 1098 SM Amsterdam,
The Netherlands

source (Forsburg and Guarente 1989). Genes corresponding to *HAP2* and *HAP3* have been isolated from a number of organisms (Chodosh et al. 1988a; Hooft van Huijsduijnen et al. 1990; Maity et al. 1990; Vuorio et al. 1990; Olesen et al. 1991; Li et al. 1992b). Chodosh et al. (1988b) have shown that human CP1 and the *S. cerevisiae* HAP2/3/4 complex bind to the same sequence, CCAAT, and make similar DNA contacts and, in addition, that the HAP2 and HAP3 homologous subunits are functionally interchangeable in vitro. The CP1 complex consists of more than two subunits, but none of the additional subunits appear to correspond to the *S. cerevisiae* HAP4 factor.

For HAP2, sequence comparison and deletion analysis showed that the 65-amino acid core domain is necessary and sufficient for complex formation and DNA binding (Olesen and Guarente 1990; Olesen et al. 1991). For HAP3, the central B-domain of 90 amino acids is thought to be required both for complex formation and DNA binding (Hooft van Huijsduijnen et al. 1990; Maity et al. 1990; Vuorio et al. 1990; Becker et al. 1991). So far, inspection of the sequences of either B-domain has failed to reveal known DNA-binding motifs (Struhl 1989). However, based on results obtained using the chelator orthophenanthroline, Hooft van Huijsduijnen et al. (1987) suggested that the mammalian NF-Y protein complex is a metallo-protein and that most probably Zn^{2+} is involved. Recently the same authors proposed a half-finger model in which the cysteine residues at position 68 and 72 of HAP3, together with the histidine residues at position 196 and 200 of HAP2 (numbering according to *S. cerevisiae* HAP genes), form a motif reminiscent of the 2*cys-2*his zinc-finger motif (Hooft van Huijsduijnen et al. 1990). We report here the isolation of the *HAP3* gene of the yeast *Kluyveromyces lactis* and show that the protein resembles other HAP3 factors, but is exceptional in having a novel 4-cysteine zinc-finger motif. Unlike the corresponding *S. cerevisiae* disruptant, the *KIHAP3* null mutant is not respiratory deficient, suggesting either that the HAP2/3/4 complex in this yeast may

fulfil a different function or that escape from glucose repression is mediated by different or additional factors.

Materials and methods

Strains, media and genetic methods

Yeast strains used are listed in Table 1. Both *S. cerevisiae* and *K. lactis* cells were grown in YPD (1% yeast extract, 2% Bacto-peptone and 2% glucose), YPEG (1% yeast extract, 2% Bacto-peptone, 2% w/v ethanol and 2% glycerol) or WO (0.67% Yeast Nitrogen Base without amino acids and 2% glucose) supplemented with the appropriate amino acids and nucleotides. For plates 2% Bacto-agar was added. Standard methods were used for genetic analysis (Campbell and Duffus 1988), yeast transformation (Dohmen et al. 1991) and *Escherichia coli* transformation (Sambrook et al. 1989).

Construction of a *K. lactis* genomic DNA library

Genomic DNA was isolated from *K. lactis* CBS2359 and partially digested with *Sau3A*. Fragments were separated in a 10%–40% sucrose gradient. Fragments ranging from 4 to 10 kb were ligated into the *Bam*HI site of *YepLac195* (Gietz and Sugino 1988). After transformation into *E. coli* about 25 000 transformants were obtained. More than 90% of the transformants contained the *YepLac195* vector with an insert (white colonies). Transformants were pooled and plasmid DNA was isolated.

Isolation and sequence determination of the *K. lactis* HAP3 gene

In five separate transformations, 5 µg of the *K. lactis* genomic library was transformed into SHY40. Transformants were selected for uracil prototrophy at 30°. After 3 days, approximately 10 000 transformants were replica-plated onto YPEG. Twenty-four transformants complemented the respiration deficiency of the *hap3*-null strain. Vector DNA was isolated (Soni and Murray 1992) and, after amplification in *E. coli*, retransformed into SHY40 to confirm complementation. Restriction maps were constructed of the complementing plasmids and the *KIHAP3* gene was located by Southern blotting, using a 1030 bp fragment of *ScHAP3* (*Xho*I-*Bam*HI) as probe. The complementing clones could be divided into two classes containing overlapping DNA inserts. Suitable fragments were subcloned in

Table 1 Yeast strains

Strain	Genotype	Source
<i>Saccharomyces cerevisiae</i>		
DL1	a, <i>his3-11,15,leu2-3,112, ura3-251,328, 372</i>	VanLoon et al. (1984)
BWG1-7a	a, <i>leu2-3, 2-112, his4-519, ade1-100, ura3-52</i>	Guarente and Mason (1983)
JO1-1A	As BWG1-7a but <i>hap2</i>	Olesen and Guarente (1990)
SHY40	As BWG1-7a but <i>hap3::HIS4</i>	Hahn et al. (1988)
SLF401	As BWG1-7a but <i>hap4::LEU2</i>	Forsburg and Guarente (1989)
<i>Kluyveromyces lactis</i>		
CBS2359	a	CBS*
MW179-1D	α, <i>lac4-8, uraA, leu2, metA1, trp1, adeW, adeR</i>	M. Wesolowski-Louvel
MW203-5C	a, <i>leu2, trp1, his2-2</i>	M. Wesolowski-Louvel
WMD001	MW179-1D × MW203-5C; <i>trp1/trp1, leu2/leu2</i>	This study
WMD7300-D1	As WMD001 with <i>HAP3/hap3::ScLEU2</i>	This study
WMH7302-D1	<i>uraA, leu2, trp1, metA1, adeR, his2-2, hap3::ScLEU2</i>	This study

* Central Bureau voor Schimmekulturen (Delft)

Table 2 Primers

Gene	Primer	Sequence	Position relative to ATG	Reference
Bluescript KS	BKS	CGAGGTCGACGGTATCG		Stratagene
Bluescript SK	BSSK	TCTAGAACTAGTGGATC		Stratagene
<i>ScHAP3</i>	PH154	GCGATAAAAAATGGGATG	– 1571/ – 1555	Hahn et al. (1988)
<i>ScHAP3</i>	PH155	GCGTAAATGCCTTGTTTC	– 1274/ – 1258	Hahn et al. (1988)
<i>ScHAP3</i>	PH156	GCCAATATGGCGATGCC	– 977/ – 961	Hahn et al. (1988)
<i>ScHAP3</i>	PH157	CAGTGGAACCCGATCC	– 678/ – 662	Hahn et al. (1988)
<i>ScHAP3</i>	PH158	GTAGCTCCTGTTGTGGAC	– 377/ – 360	Hahn et al. (1988)
<i>KlHAP3</i> (LSP)	PH91	CTTGGTTAGGTCAATCAAGAC	– 1469/1449	This study
<i>KlHAP3</i> (GSP)	EP1	CCGTAAATAGTCTTACGCTT	248/229	This study
<i>ScLEU2</i> (MSP)	PH89	GACAAGGAGGAGGGCACCAC	1333/1352	Oliver et al. (1992)

vector pBluescript II⁺ (Stratagene), either KS or SK, and the DNA sequence of both strands was determined according to Sanger et al. (1977). For sequence verification of the *S. cerevisiae HAP3* locus the vector pSH95 (Hahn et al. 1988) was used together with primers PH154–158 (Table 2).

KlHAP3 gene disruption and confirmation by Polymerase chain reaction (PCR)

For disruption of the chromosomal copy of the *KlHAP3* gene, a 2200 bp fragment containing the *S. cerevisiae LEU2* gene was blunted and cloned in the *EcoRI* site of the *KlHAP3* gene (Fig. 1). A 2900 bp *NdeI-NsiI* fragment was isolated and transformed into the diploid *K. lactis* strain WMD001. Transformants were selected for leucine prototrophy and screened for the presence of the correct gene disruption by the PCR technique recently described by Sathe et al. (1991), using primers specific for the *HAP3* locus (LSP), the *HAP3* gene (GSP) and the *ScLEU2* marker (MSP) (Table 2). Cells were either directly added to the reaction mixture or chromosomal DNA was isolated (Soni and Murray 1992) prior to performing the PCR reaction. Only one transformant appeared to have a correctly disrupted *KlHAP3* gene. This diploid, WMD7300-D1, was subjected to random sporulation and gene disruption was confirmed by Southern blotting of chromosomal DNA isolated from several spores. Phenotypes of the spores were determined by using media containing different carbon sources. The presence of selective markers was determined using drop-out media (Campbell and Duffus 1988).

Plasmids

A 2400 bp *DraI-BamHI* fragment of the *ScHAP3* gene and a 2650 bp *HincII-SmaI* fragment of the *KlHAP3* gene were blunted and separately cloned in the unique blunted *Asp718* site of vector pRS306/K2-*ScCEN16* resulting in plasmids pRS/*ScHAP3* and pRS/*KlHAP3* respectively. pRS306/K2-*ScCEN16* is a low-copy-number *S. cerevisiae-K. lactis* shuttle vector (Heus et al. 1993).

RNA analysis

Total RNA isolation and Northern blot analysis were performed as described (DeWinde and Grivell 1992). Poly(A)⁺ mRNA was isolated as described by the manufacturer, using a Poly (U) Sepharose 4B column (Pharmacia). For hybridisation, a 500 bp *EcoRI-HindIII* fragment of plasmid pART6 containing the *K. lactis CYC1* gene

(Freire-Picos et al. 1993) and a 900 bp *EcoRI-HindIII* fragment of plasmid pKL7 containing the *K. lactis* actin gene (Deshler et al. 1989) were used. To detect the different transcripts of the *ScHAP3* gene loci the following fragments were used: *DraI-BglIII*, 1036 bp (*ScORFX* 3 kb); *XhoI-SacI*, 232 bp (*ScHAP3*), both isolated from plasmid pSH95 (Hahn et al. 1988). Quantitation of Fig. 5 was performed with a Datacopy GSplus laserscanner (Xerox) using the Scan Analysis densitometer application (Biosoft) with matched background subtraction.

Accession numbers

The nucleotide sequences of the *KlHAP3* locus and the corrected *ScHAP3* locus will appear in the EMBL, Genbank and DDBJ Nucleotide Sequence Databases under the accession numbers L25779 and M20318, respectively.

Results

Isolation of the *K. lactis HAP3* gene

Owing to disruption of the *hap3* gene the *S. cerevisiae* SHY40 strain is unable to respire and is therefore unable to grow on non-fermentable carbon sources. The *K. lactis HAP3* gene was isolated by functional complementation of this phenotype (detailed in Materials and methods). Plasmids were isolated and, after subcloning, the DNA sequence of the insert was determined. The restriction map of the *K. lactis HAP3* locus and the open reading frames (ORF) found are depicted in Fig. 1. The DNA sequence, together with the translated ORFs, is shown in Fig. 2. Pulsed field gel electrophoresis shows that the *KlHAP3* locus is located on chromosome V (M. Wesółowski-Louvel, personal communication).

To determine whether the homologous *ScHAP3* and *KlHAP3* genes affect the growth rate of *S. cerevisiae* differently, both were cloned in the *S. cerevisiae/K. lactis* shuttle vector pRS306/K2-*ScCEN16* and transformed into the *S. cerevisiae hap3* null strain (SHY40). No growth rate difference was observed on glucose medium, but on ethanol/glycerol medium the *KlHAP3* transformants grew significantly more slowly than the

Fig. 1 **A** Restriction map of the *KIHAP3* gene locus on the chromosome. The open reading frames (ORFs) corresponding to the *KIHAP3* and *KIORFX* proteins are indicated by *stippled* and *white bars*, respectively. The C-terminal end of ORFX has not been determined. Two overlapping clones, #3.1 and #3.2, were isolated from the *K. luyveromyces lactis* genomic library and relevant restriction sites are indicated. Sites depicted in *italics* are derived from vector YepLac195. For disruption of the *KIHAP3* gene, a 2200 bp fragment containing the *ScLEU2* gene was cloned into the *EcoRI* site. After transformation of the *NdeI-NsiI* fragment into *K. lactis* diploid WMD001, colonies were selected for leucine prototrophy. **B** For the isolation of transformants having a correctly disrupted *KIHAP3* gene, DNA was isolated and subjected to polymerase chain reaction (PCR) analysis. The locus-specific primer (LSP) was used together with either the marker-specific primer (MSP) or gene-specific primer (GSP). **C** After completion of the PCR reaction, one-tenth of the

reaction volume was subjected to electrophoresis on an agarose gel. Lane 1, λ DNA/*HindIII*; lanes 2, vector pFLU1/*HincII*; lanes 1 and 6, spore derived from diploid A; lanes 2 and 7, spore derived from diploid B; lanes 3 and 8, diploid A; lanes 4 and 9, diploid B; lanes 5 and 10, wild-type CBS2359. Lanes 1–4, LSP and MSP; lanes 5–9, LSP and GSP. Note that for both diploids A and B in lanes 9 and 10, respectively, the wild-type and the disrupted *KIHAP3* locus give rise to different fragment lengths. **D** Diploid WMD7300-D1 was sporulated. DNA was isolated from spores WMH7301/7302/7302 and used for Southern blotting. For hybridisation probe I, an *EcoRI* fragment derived from #3.1 was used. In the *StuI* digestion, the 3.5 kb fragment containing part of the *KIHAP3* gene and the *ScLEU2* gene in the 7301/7302/7303 lanes is less clear owing to the small size of the region capable of hybridisation. The double band of 1.3 kb in the wild-type lane is thus not visible. (*B* *BamHI*; *E* *EcoRI*; *H* *HindIII*; *K* *KpnI*; *P* *PstI*; *S* *Sall*; *Sm* *SmaI*; *St* *StuI*; *X* *XbaI*)

ScHAP3 transformants. These results suggest that the hybrid HAP2/3/4 complex activates mitochondrial biogenesis less efficiently than the complex containing the endogenous *ScHAP3* protein. As yet, however, alternative explanations, including suboptimal synthesis, or increased turnover of the *KIHAP3* protein in *S. cerevisiae* cannot be completely excluded.

Moreover, *KIHAP3* is unable to replace either *ScHAP2* or *ScHAP4* protein in *S. cerevisiae*, as was determined by transforming the original plasmid #3.1 into both a *Schap2* null (JO1-1A) and a *Schap4* null (SLF401) strain. After selection for uracil prototrophy, neither of the transformants was able to grow on a non-fermentable carbon source.

1	GATCACCGAGCAACACTTTGGAAAAACGCTCGAAGATACCATTCAGGAAAGGAATG	59	1140	GCACCTAGCAAGTGTTCACCAAGAAATTAACCCCTTTTCAACGTTTGTGCCAACATTA	1199
<	I V P A V S Q F F A E F I G N L S L F P			V L L T N L L F N V G K S L T Q G L M L	
60	GAATAGAAAGCAGTAAGTTCTTATTTGTTTCAAAAATGAGGTACTTGACCACCTTAGATC	119	1200	GGAATAAGACTCCATTGGCGAATTTGTTCATATGCTCATCTTGACATCTTCTTACCA	1259
I	Y F C Y T R I T E L F S T S S W K S G			F L V G N A F K N M H E N K V N M	
120	CAATGACAAATTTCAATAAGTATGACGAATGATTGGTCCGAATATTATAATGAAGAGTG	179	1260	AATGATATGAAGTAACCTTGCACCCCTTCTTAAAGAACAAATGCTGAGCACTAGCACCTACA	1319
I	V F K L L Y S S N I P G F I I I F L S				
180	ATAAGTACAGATAAACTTGGTTAGGTCAATCAAGACCTTCATAGATAACCATAAGTTCT	239	1320	GCAGTATAACCAATCCTATCGATGAGCACTGGCCGTTTATGGATCTTATTAGTGCTCT	1379
L	Y L Y F K T L D I L V K M S L W L N K				
240	TCTTGACGATAGCCTTAACAAGGCGCTAAAAACAAGCGTAATGCCTCCTCAATCGGAG	299	1380	TCTTTGCCAATGTATACATTGAAAAATCTCGATCTTTGACGTTGAGCAATTCAAAAATCA	1439
K	C C L R L L A A L F L R L A E E I P A				
300	CAACAATAGTCTTGTATCAAGAGCGTAATTTGAAAGAAGAGATATATACCTTGCT	359	1440	TCAACAAACCTTAAATACCCATTCATGACTATTAACTACATAATAAAATACTAAGTGT	1499
F	L L R T I L S G Y N S L L S Y I G Q E				
360	CTTCCACCGTGCACAAAGATTTATGATTAGCTTATCACCTTCGTGTAACAAATGCTTGA	419	1500	TTAAGGTCCTTCACTGATTGGTCTGTATGAAGTCAGAGCTAGCGACACCTACTAAAT	1559
E	V T C L S N I I L K D G E T L L H K F				
420	AGCATAGCTGGAATACACTTTTTTGAAGTCTGAACGGTATCATTTTGGAATAGTATG	479	1560	ATGATACAGAGACAGAAGCTGAGAAGCAAGAGAGGAGCAATAACTATCTTAAACGAATA	1619
C	L Q F Y V K K F H Q V T D N Q F Y Y S			M D T E T E A E K Q E R H N N Y L N E L	
480	ACTCTTGCTTGTTCAGGCTGATTTTGTCACTTGTAGCTATAACGCTCTCCAGTAG	539	1620	GCGGAGCAAGATCGATGGCTTCGGATTAACAACGTCGCAAGACTAATGAAAAACAGTTG	1679
E	Q R T E P R I K T L K Y S Y R E G T T			A E Q D R W L P I N N V A R L M K N T L	
540	TATAGACCTCTGAGTAATATTTAGGAGTAACATGCAAGTAAAGTCTCGCTGAACGA	599	1680	CCAGCGACCAACAGGTATCTAAGATGCGAAGGAATGATGCAAGAGTGCCTAAGCGAA	1739
Y	V E Y Y N Y C A L L T M A H V L			P A T T K V S K D A K E C M Q E C V S E	
600	GCTTACCAACTGCAAAATGCAAAAATAGCGATCCCTTCTTGTGTAATCATCATGCAACT	659	1740	TTTCATATCGTTTGTCACTAGCGAGCGCTGTATAGATGTACCACTGGGAAGCGTAAGACT	1799
K	G V A F A F I A I G E Q K Y D D H L E			F I S F V T S E A C D R C T S G K R K T	
660	CTAGCCCACTCCGTTCCACTTTCTCAAACAGACAATATAAAGTTTCGCAACAC	719	1800	ATTAACGGTGAAGATATTCTTCTATCGCTACATGCATTGGGATTCGAGAACTATGCCGAA	1859
L	G V G N V M K E F W V V I I F N A V C			I N G E D I L L S L H A L G F E N Y A E	
720	ATGCAAGGTGACACCTAACTCTCATATTTGGATCTGATAAGTGAATGAGTAAATATT	779	1860	GTGCTCAAGATTTACCTAGCAAAATACAGACAACACAGCGCAATAAAGAAATCAGATGAT	1919
A	F T V G L S E Y K S R I F H N L L Y Q			V L K I Y L A K Y R Q Q A I K N Q M M	
780	GATGTACCAAGTAGAAGCGTTCCTGACCAACTCAATCAATATACTAAGCCATATAAGGA	839	1920	TACCTAAGGAGGATGTGAAGGTTTCATACAGAGAAGAATCAAGGCGAGAAGACGCA	1979
H	V L Y F P E S V L E I L I S L W I L F			Y P K E D V E G S Y T E E E S R A E D A	
840	ATATGGAGGCTTTGAAGTATGGAAGTCAATGAAGTACGAATTTAGATTAGATATTGCC	899	1980	GATCAGCTGACACAGGAAGGCAGCAACCTCTTCATCTGTGCTTCTTCAGCTTCCAGC	2039
I	S A K F Y P L D I F Y S N L N S Y Q W			D H V T Q E R Q Q P L H P V P S S A S T	
900	ATGAAATTAGGATAATAGACAATGGTAACCAATCAACAACGGTATATAACCAAAATTA	959	2040	ACGACAAGCAGAGGATTCGGCATCAGAACATGTACAGTCTGCTTCAAAAAGCACTCAT	2099
S	I L I I S L P L G I V F P I Y G F N I			T T S T E V S A S E H V Q S S K S T H	
960	TCAGGGATTGCAAGTTCTCGTATCTCCATCTCTACTTTTTCCAATTTGCCGCTAT	1019	2100	AATCCAGCACAGCTCAACCACGGATTGTCTGCAACACCCACTGGGATCGCGCCATCT	2159
L	S Q L T R T D E G D E V K E L K G S D			N P A Q A Q P R I V V Q Q P L G S R H S	
1020	CTGTGTAGATTTGATCCGTAAGGTGAGATCCGAATCGATTCTCTACTGAAGAATAGTA	1079	2160	ACTGCGGATCACTTGTAACTTACCTCTTCTCACTATAACAAACGCATCTATATATAT	2219
T	T S K I R L T S I R I S E R S F F L V			T A D H L *	
1080	CTGTAGAAAGTAAAACTTAGGAAAGAGGTAATACCAAAAATTCGGGGGGATAAATAAC	1139	2220	ATATATATATATAGATATGTA	2279
T	S L L F E L F S T I G F I R P S L Y R			TGGATGTACTTAAATATTAAACGACCAACGGCTCAATAC	
			2280	ATACCAAGGAGGATCTTAACGTACAATTATGCAATTTGGCAAGCAAGTGGTACGATA	2339
			2340	GGTCCGCTTTCAGACATTATATTTCCGTATTAGTCTAAAATGTCATTTTCCGGTGATA	2399
			2400	TTTGCGCCACCGGTATTTATAATATTGTAAGATAAATACCAATTTATACAATAAGGAA	2459
			2460	CACCTTGGATTCCGGTAACATTTTAAGAGGAAAAAGGCCT	2499

Evolutionary comparisons

The *K. lactis* HAP3 gene codes for a protein of 205 amino acids with a calculated molecular weight of 23 kDa and an estimated pI of 5.3. The sequence was compared with sequences present in the EMBL and Swissprot databases using the (T)FASTA algorithms. High similarity scores were found with *S. cerevisiae* HAP3 (Hahn et al. 1988), mammalian NF-YB/CP1-A (Hooft van Huijsduijnen et al. 1990; Maity et al. 1990; Vuorio et al. 1990; Becker et al. 1991) and the recently isolated NF-YB equivalents of man, chicken, toad, lamprey and maize (Li et al. 1992b). In addition, similarity was found with a DNA region downstream of the *amdR* gene of *Aspergillus nidulans* (Andrianopoulos and Hynes 1990). Sequence alignment of the homologous HAP3 proteins is shown in Fig. 3.

As has been proposed by Li et al. (1992b), the HAP3 protein can roughly be divided into three domains:

(1) The small N-terminal A-domain is only weakly conserved. Although its sequence and domain length are conserved in higher eukaryotes it differs significantly in both yeasts and maize. In *K. lactis* the A-

Fig. 2 DNA sequence and translation of the *KIORFX* and *KIHAP3* genes. The *KIORFX* gene starts at position 1251 and is oriented oppositely to the *KIHAP3* gene, which is located from position 1560 to 2179. The stop codon for *KIORFX* is outside the sequenced region. Downstream of the *KIHAP3* gene, perfect matches to the recently determined consensus poly (A) site signal in *Saccharomyces cerevisiae* (Irniger et al. 1992) are underlined

domain is only 20 amino acids long. The only observable conservation is the preference for acidic residues, which gives this region a rather hydrophilic character.

(2) The central B-domain is highly conserved both in length (90 amino acids) and sequence. The *K. lactis* and *S. cerevisiae* sequences are 89% identical in this region, while the *K. lactis* and human comparison reveals 66% identity. These values rise to 98% and 88% respectively if conservative substitutions are included. The C-terminal end of the B-domain coincides with the conserved position of an intron in the higher eukaryotes (Li et al. 1992a). In the *A. nidulans* HAP3 counterpart, this position is marked by an abrupt decrease in homology and the occurrence of a consensus donor splice site-GTAAGA-(Kulmburg et al. 1991). We therefore suggest that *A. nidulans* resembles the higher

Domain A

Human	MDGDSSTTDASQLGISADYIGGSHYVIQPHD-DTEDSMN--DHEDTNGSKESF	50
Rat	MDGDSSTTDASQLGISADYIGGSHYVIQPHD-DTEDSMN--DHEDTNGSKESF	50
Mouse	MDGDSSTTDASQLGISADYIGGSHYVIQPHD-DTEDSMN--DHEDTNGSKESF	50
Chicken	MDGDSSTTDASQLGISADYIGGSHYVIQPHD-DTEDSMN--DHEDTNGSKESF	50
Xenopus	*-----*	
Lamprey	MDADGSTTDASQLGITGDIYIGGGHYVLCSSDGDAGSIASGDHRESGGSKIRY	53
Maize	MAEAPASPGGGGSGHESGSPRGGGGGGSV	29
<i>A.nidulans</i>	*-----*	
<i>S.cerevisiae</i>	MNTNESEHVSTSPEDTQENGNGNASSSGSLQQISTL	35
<i>K.lactis</i>	MDTETEAEKQERHNNYLNEL	20

Domain B

Human	REQDIYLPINVARIMKNAIPQTGKIAKDAKECVQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	140
Rat	REQDIYLPINVARIMKNAIPQTGKIAKDAKECVQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	140
Mouse	REQDIYLPINVARIMKNAIPQTGKIAKDAKECVQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	140
Chicken	REQDIYLPINVARIMKNAIPQTGKIAKDAKECVQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	140
Xenopus	----- VQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	
Lamprey	REQDIYLPINVARIMKTIPSSGKIAKDAKECVQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	143
Maize	REQDIYLPINVARIMKTIPSSGKIAKDAKECVQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	119
<i>A.nidulans</i>	----- SEASEKCCQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	
<i>S.cerevisiae</i>	REQDRWLPINVARLMKNTLPSAKVSKDAKECMQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	125
<i>K.lactis</i>	REQDRWLPINVARLMKNTLPSAKVSKDAKECMQECVSEFISFITSEASERCHQEKRRKTINGEDILFAMSTLGFD SYVEPLKLYLQKFR	110
	C---C---F---C---C---KR	
	zinc finger	

Domain C

Human	EAMKGEKGIGGAV-TATDGLSEELTEEAFTNQLPAGLITADGQQQNMVYTTTSYQQISGVQIQFS	205
Rat	EAMKGEKGIGGAV-SATDGLSEELTEEAFTNQLPAGLITADGQQQNMVYTTTSYQQISGVQIQFS	205
Mouse	EAMKGEKGIGGAV-SATDGLSEELTEEAFTNQLPAGLITADGQQQNMVYTTTSYQQISGVQIQFS	205
Chicken	EAMKGEKGIGG-----*	
Xenopus	EAMKGEKGIGGTV-TTGDGLGEDLAEEFTSQIPAGLITADGQQQNMVYTTTSYQQISGVQIQFS	
Lamprey	ESMKGEKGINATVVTITDGLSEELTEESFGHATSIITADGQQQNMVYTTTSYQQISGVQIQFS	209
Maize	E-MEGISKLTAKSSDGSIKKDALGHVGASSAAEGMGQQGAYNQGMGYMQPQYHNGDISNV	179
<i>A.nidulans</i>	-----*	
<i>S.cerevisiae</i>	QQQALKNLMEQDHEEVP	144
<i>K.lactis</i>	QQQALKNMYPKEEVEGSGYTEESRAEDADHVTQERQQPLHPVPSSASTTTSTEVSASEHVQSSSKSTHNPAQAQPRIVVQQPLGSRHSTADHI	205

eukaryotes in having a split NF-YB/HAP3 gene and that the position of the intron is also conserved in this organism.

The B-domain is thought to be involved in subunit interaction and DNA binding (Hooft van Huijsduijnen et al. 1990; Maity et al. 1990; Vuorio et al. 1990; Becker et al. 1991). For *S. cerevisiae*, evidence has been presented that supports this hypothesis (Olesen and Gaurent 1990).

Except for *K. lactis* HAP3, all the homologous HAP3 proteins share the protein sequence motif CVSExISF (L, I, V, M, T)SEAS (D, E) (K, R) C, which is present in the B-domain. The *K. lactis* HAP3 protein differs from the other sequences in having the last conserved serine residue replaced by a cysteine residue. As described below we propose that this sequence is part of a novel four-cysteine zinc-finger motif.

(3) The carboxy-terminal C-domain of HAP3 differs significantly between lower and higher eukaryotes. *K. lactis* has a C-domain of 95 amino acids, which is long compared with the corresponding region of *S.*

Fig. 3 Amino acid alignment of all known HAP3 protein sequences. The proteins are divided into three domains according to Li et al. (1992b). Homologous substitutions are *stippled* and non-identical residues are *boxed*. Unsequenced regions of chicken, *Xenopus* and *Aspergillus nidulans* are depicted as *broken lines*. Asterisks indicate that the start or stop codon has not been determined. Arrows indicate the positions of introns in the higher eukaryotic versions. Residues that are part of the putative novel zinc-finger motif are shown as in *reverse*. This novel zinc-finger motif is depicted below domain B

cerevisiae HAP3, which is only 18 amino acids in length. The C-domains of all homologous HAP3 factors are rich in glutamine residues. This feature is present in activation domains of many transcription factors and it is tempting to speculate that HAP3 is directly involved in *trans*-activation. Positions 124–149 in the C-domain of KIHAP3 contain predominantly acidic amino acids and are predicted to form an α -helix. This feature is also characteristic of many transcription

factor activation domains. Finally, positions 156–179 in the C-domain of KIHAP3 are serine/threonine-rich with 10 serines and 5 threonines present in 24 residues. The HAP2 proteins of both *S. cerevisiae* and *Schizosaccharomyces pombe* also contain serine/threonine-rich regions, although unlike the situation in KIHAP3 these are present at the N-terminus of the protein.

K. lactis HAP3 contains a novel four-cysteine zinc-finger motif

Careful examination of the *K. lactis* HAP3 protein sequence shows that the only four cysteine residues present are arranged in a configuration that is reminiscent of the 2*cys-2*his zinc-finger motif in other transcription factors (indicated in Fig. 3). Other characteristics of this motif are also present: a phenylalanine residue at position 4 in the loop and three basic residues C-terminal of the last cysteine. We therefore propose that the HAP3 protein of *K. lactis* contains a zinc-finger motif. This motif is unusual in being closely related to the 2*cys-2*his zinc-finger motif while possessing four cysteine residues. Since HAP3 factors from yeast to man display functional conservation, we propose that all other HAP3 proteins contain a similar, novel zinc-finger motif. This motif has not been recognised so far because these proteins contain a serine residue at the position occupied by the third cysteine residue. In addition, the maize HAP3 protein contains a threonine residue at the normal first cysteine position. This serine and probably also the threonine need not hamper function. As in other zinc-binding proteins, such as phosphomonoesterase (Coleman 1992), the oxygen atom of these residues should be capable of assisting in the positioning of the Zn^{2+} atom.

Disruption of the *K. lactis* HAP3 gene does not affect growth on a non-fermentable carbon source

Although the HAP2 and HAP3 proteins are highly conserved between various organisms, it is only in the yeast *S. cerevisiae* that the trimeric HAP2/3/4 complex seems to be involved in mitochondrial biogenesis and, in particular, in the carbon source response. Disruption of any of the three HAP genes in *S. cerevisiae* results in inability to respire and hence to grow on non-fermentable carbon sources. To see whether this holds for the yeast *K. lactis*, the KIHAP3 gene was disrupted as detailed in Materials and methods. One-step gene disruption in *S. cerevisiae* is usually highly efficient, yielding large numbers of transformants containing the desired disruption. This is not the case in *K. lactis*, in which the disruption fragment seems to integrate more or less randomly in the genome.

To screen the transformants rapidly a recently developed PCR technique was used (Sathe et al. 1991).

This method is schematically depicted in Fig. 1 together with the result of the KIHAP3 gene disruption. Disruption was confirmed for spore WMH7302-D1 by Southern blotting (Fig. 1). We were surprised to find that all leucine prototrophic spores grow on medium containing ethanol/glycerol as sole carbon source without showing significant phenotypic growth differences relative to spores containing an intact KIHAP3 gene. In contrast to *S. cerevisiae*, the KIHAP3 gene in *K. lactis* is therefore not essential for growth on non-fermentable carbon sources.

To check whether gene disruption affects the mRNA level of genes whose counterparts in *S. cerevisiae* show carbon source-dependent expression, we carried out Northern analysis of the KICYC1 gene, encoding iso-1-cytochrome *c* (Freire-Picos et al. 1993). In *S. cerevisiae*, the ScCYC1 gene is poorly expressed when the cells are grown in the presence of glucose and is induced more than ten fold when the cells are grown solely on a non-fermentable carbon source (Zitomer and Lowry 1992). This carbon source-dependent expression is mainly regulated by the HAP2/3/4 complex. As in the promoter of the ScCYC1 gene, the promoter of the KICYC1 gene contains a consensus recognition site for the HAP2/3/4 complex and as Fig. 4 shows, the expression of the KICYC1 gene is also carbon source regulated in *K. lactis*. The expression of the KICYC1 gene in ethanol/glycerol grown cells is roughly seven fold higher than in glucose-grown cells. However, Fig. 4 shows that KIHAP3 gene disruption affects neither the carbon source-dependent expression of the KICYC1 gene, nor the level of mRNA produced.

Fig. 4 Strain WMH7302-D2 (*Kihap3::ScLEU2*) was transformed with either pRS/ScHAP3 or pRS/KIHAP3. Transformants and untransformed WMH7302-D2 were grown in rich medium containing the carbon source indicated. RNA was isolated and 10 µg was used on a native agarose gel for Northern blot analysis

Gene order at the *HAP3* locus is conserved

In *K. lactis* the *ORFX* gene, coding for an unknown protein, is located upstream of the *KIHAP3* gene. Alignment of the DNA sequence of the *HAP3* loci of *S. cerevisiae* and *K. lactis* reveals that the DNA sequence of this *KIORFX* is highly conserved in *S. cerevisiae* (Fig. 5), suggesting that the region upstream of the *ScHAP3* gene in *S. cerevisiae* could code for a protein homologous to *KIORFX*. This contrasts with the observation of Hahn et al. (1988) that a 3 kb mRNA derived from this region does not contain any sizeable ORFs. To clarify these conflicting observations we verified the DNA sequence of this region, using plasmid pSH95, and found that the published sequence contains multiple errors. The corrected *ScHAP3* locus reveals the presence of the homologous *ScORFX*. The *ScORFX* and *KIORFX* proteins are at least 428 and 417 amino acids long, respectively. Stop codons for both proteins have not yet been found, because the coding sequences extend beyond existing clones. *ScORFX* and *KIORFX* are 62% identical (82% homologous) over the region so far sequenced and do not show significant homology to any sequence present in the database (Fig. 6). *ORFX* is a rather basic protein and contains a high percentage of leucine and isoleucine residues. Predictions according to the Chou and Fasman (1978) and Garnier et al. (1978) algorithms, show that the protein is likely to be membrane associated, with at least eight putative β -sheet membrane spanning-domains separated by turns (Klein et al. 1985).

KIORFX and *KIHAP3* are anti-sense oriented but their transcripts do not overlap

Hahn et al. (1988) have reported that the *ScHAP3* locus encodes two overlapping but anti-sense mRNAs. The

Fig. 5 DNA identity matrix showing that homology between the regions containing the *K. lactis* and *S. cerevisiae* *ORFX* and *HAP3* genes is restricted to the coding sequences. Only the 1–2400 bp region is used. The *ORFX* and *HAP3* genes are indicated by arrows. The line indicates that, with a window of 30, identity is at least 65%

HAP3 mRNA (550 nucleotides) and a transcript of approximately 3000 nucleotides in length. This latter RNA was found to lack any sizeable ORFs. Hahn et al. and other authors have hypothesised that this anti-sense transcript might function to regulate the expression of the *HAP3* gene. Our observation that the anti-sense oriented mRNA does code for protein *ORFX* together with the determination of the transcription start site of this messenger by Hahn et al. (1988) would suggest that this mRNA has an unusual untranslated leader of 785 nucleotides. Since *S. cerevisiae* mRNAs normally have 5'-leaders ranging in length from 50 to 350 nucleotides, we carried out transcript mapping by Northern blotting on both total RNA and poly (A)⁺ RNA. For the hybridisation we used DNA probes containing only the coding sequence of either the *ScORFX* gene or the *ScHAP3* gene. The results depicted in Fig. 7 show first, in contrast to the findings of Hahn et al., that the *ScORFX* transcript is polyadenylated and second, that each gene-specific probe hybridises only with its own transcript. The anti-sense oriented transcripts derived from the *ScORFX* and *ScHAP3* genes do not overlap and this observation refutes the proposed regulation model of Hahn et al. (1988).

Discussion

This report describes the isolation and characterisation of the *K. lactis* gene coding for the *KIHAP3* protein. The *KIHAP3* protein is functional in an *S. cerevisiae* *hap3* null strain in respect of the complementation of the respiratory-deficient phenotype of this mutant. In *S. cerevisiae* the *ScHAP3* protein is thought to be part of the heteromeric trimer *HAP2/3/4* and this result thus suggests that the *KIHAP3* protein is capable of association with both *ScHAP2* and *ScHAP4* and that it can subsequently activate the expression of genes involved in mitochondrial biogenesis in *S. cerevisiae*. The slightly slower growth rate of *Schap3* null transformants containing the *KIHAP3* gene on a low-copy-number vector suggests that the heterologous complex functions less efficiently compared with the homologous *ScHAP3* protein. The long C-terminal extension of the *KIHAP3* protein is perhaps responsible for this effect. However, the possibility of less efficient synthesis or increased turnover cannot be excluded.

The C-domain of the *KIHAP3* protein shares features with activation domains of transcription factors. The *K. lactis* factor may thus combine functions that in *S. cerevisiae* are fulfilled by separate proteins, one possibility being that it overcomes the need for the third activating protein *HAP4*. This would be analogous to the results obtained by Olesen and Guarente (1990), who showed that fusion of the transcription activation domain of *GAL4* to the *ScHAP2* protein rescues

a *Schp4* null mutant. However, the observation that KIHAP3 cannot complement the respiratory-deficient phenotype of either *Schp2* null or *Schp4* null strains shows that the transcription activation capacity of the KIHAP3 protein may also be dependent on other species-specific factors. This dependence may even differ from gene to gene. This idea is supported by the observation that in *S. cerevisiae* ScHAP2 and ScHAP3 regulate expression of the *HAR3* gene, even in the absence of ScHAP4 (B. Daignan Fornier and M. Bolotin-Fukuhara, personal communication).

The alignment of the homologous HAP3 proteins shows that the central B-domain is highly conserved in evolution. This is consistent with the important functions this region of 90 amino acids probably has in subunit association and DNA binding, as has been suggested by several authors (Hooft van Huijsduijnen et al. 1990; Maity et al. 1990; Vuorio et al. 1990; Becker et al. 1991). The exact relationship between these two activities is as yet unclear, since DNA binding, as measured by gel-mobility shift assays (Hahn et al. 1988), requires the simultaneous presence of at least ScHAP2 and ScHAP3 and is clearly limited by the amount of either factor present in wild-type cells. The presence of a novel zinc-finger motif in the KIHAP3 protein leads us to speculate that the KIHAP3 protein and all homologous HAP3 proteins can bind directly to the DNA. Limitations on the sensitivity of the gel-mobility assay and/or a requirement for ScHAP2 in stabilization of a DNA-ScHAP3 complex may explain why such complexes have not been detected so far.

The conservation of gene order of *ORFX* and *HAP3* between *S. cerevisiae* and *K. lactis* is not unprecedented. Other examples are the gene clusters *RP59/S24/L46* (Bergkamp-Steffens et al. 1992), *APA2/QCR7* (Mulder et al. 1994b) and *FPS/QCR8* (Mulder et al. 1994a). This phenomenon does not necessarily

Fig. 6 Amino acid alignment of ORFX of *S. cerevisiae* (Sc) and *K. lactis* (Kl). Identical amino acids are indicated by an asterisk and homologous substitutions by a dot. Broken lines refer to a gap in the alignment and the arrows indicate that the C-termini have not been determined. The predicted membrane-spanning fragments are shown by stippled boxes

reflect a functional relationship between these genes. The genes encode two very different proteins within unlinked pathways, as Hahn et al. (1988) have reported that disruption of the region containing the *ORFX* gene yields cells with no detectable mutant phenotype and no obvious defects in carbon source regulation. Our transcript mapping demonstrates that despite the anti-sense nature of the transcripts of these two genes, they do not overlap. This observation refutes the hypothesis in which the synthesis of the ScHAP3 protein is regulated by an anti-sense transcribed RNA. The mismapping by Hahn et al. (1988) of the transcription start site for the *ORFX* mRNA is probably due to the probe they used in the S1 nuclease assay: most probably this fragment does not contain the *ScORFX* start site at all. Our examination of the promoter of the *ScHAP3* gene shows that it lacks any obvious elements for known regulatory factors. This, when taken together with the observation that the gene is constitutively transcribed, leads to the suggestion that post-transcriptional events regulate the roles played by the HAP3 protein. Although the additional features displayed by KIHAP3 suggest important roles for this protein, disruption of the gene in *K. lactis* does not show a noticeable mutant phenotype. This is in marked contrast to *S. cerevisiae* in which gene disruption leads to respiratory deficiency. The same lack of phenotype is displayed by disruption of the *K. lactis* *HAP2* gene (C. C. Nguyen and M. Bolotin-Fukuhara, personal communication). The reason for this difference in

Fig. 7 Total RNA was isolated from *S. cerevisiae* strain DL1 grown on rich medium containing 2% ethanol and 2% glycerol. Part was used to isolate poly (A)⁺ RNA. For Northern blot analysis 20 µg of total RNA and an amount of the poly (A)⁺ fraction corresponding to 40 µg total RNA was used. To detect the *ORFX* and *HAP3* mRNAs and to show that these transcripts do not overlap probe A, fragment *DraI*-*BglII*, and probe B, fragment *XhoI*-*SacI* were used, respectively. *ScPIM1* encodes a mitochondrial ATP-dependent protease as recently reported by VanDyck et al. (1994)

phenotype of gene disruptants, between *S. cerevisiae* and *K. lactis*, is not yet clear. Possibly *K. lactis* makes use of a different (set of) factor(s) for carbon source regulation. Alternatively, HAP2 and HAP3 fulfil similar roles, but one can take over the function of the other when disrupted. The difference in phenotype may also be related to differences in HAP2/3-dependent expression of nuclear genes required for maintenance and replication of the mitochondrial genome in the two yeasts. For reasons that are as yet unclear, disruption of either *ScHAP2* or *ScHAP3* in *S. cerevisiae* causes a high degree of rho⁰ induction (Mattoon et al. 1990). This phenomenon is not observed in *K. lactis*. We are trying to resolve these differences by construction of

a *Klhap2/Klhap3* double disruptant and study of the effects on both growth and gene expression.

Acknowledgements We would like to thank M. E. Cerdan, L. A. Guarente, S. Hahn, J. J. Heus, J. J. Rossi, and M. Wesólowski-Louvel for providing strains and plasmids. We are grateful to M. Bolotin-Fukuhara for communication of unpublished results and to M. Wesólowski-Louvel for the determination of the chromosomal location of the *KIHAP3* gene. We gratefully acknowledge J. H. de Winde for helpful and stimulating discussions. This work was supported in part by grants from the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for the Advancement of Research (NWO).

References

- Andrianopoulos A, Hynes MJ (1990) Sequence and functional analysis of the positive acting regulatory gene *amdR* from *Aspergillus nidulans*. *Mol Cell Biol* 10:3194–3203
- Becker DM, Fikes JD, Guarente L (1991) A cDNA encoding a human CCAAT-binding protein cloned by functional complementation in yeast. *Proc Natl Acad Sci USA* 88:1968–1972
- Bergkamp-Steffens GK, Hoekstra R, Planta RJ (1992) Structural and putative regulatory sequences of *Kluyveromyces* ribosomal protein genes. *Yeast* 8:903–922
- Campbell I, Duffus JH (eds) (1988) *Yeast – A practical approach*. In: Rickwood D, Hames BD (eds) *The practical approach series*. IRL Press, Oxford
- Chodosh LA, Baldwin AS, Carthew RW, Sharp PA (1988a) Human CCAAT-binding proteins have heterologous subunits. *Cell* 53:11–24
- Chodosh LA, Olesen J, Hahn S, Baldwin AS, Guarente L, Sharp PA (1988b) A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally interchangeable. *Cell* 53:25–35
- Chou PY, Fasman GD (1978) Empirical predictions of protein conformations. *Annu Rev Biochem* 47:251–276
- Coleman JE (1992) Zinc proteins: enzymes, storage proteins, transcription factors and replication proteins. *Annu Rev Biochem* 61:897–946
- Deshler JO, Larson GP, Rossi JJ (1989) *Kluyveromyces lactis* maintains *Saccharomyces cerevisiae* intron-encoded splicing signals. *Mol Cell Biol* 9:2208–2213
- DeWinde JH, Grivell LA (1992) Global regulation of mitochondrial biogenesis in yeast: ABF1 and CPF1 play opposite roles in regulating expression of the *QCR8* gene, encoding subunit VIII of the mitochondrial ubiquinol cytochrome *c* oxidoreductase. *Mol Cell Biol* 12:2872–2883
- Dohmen RJ, Strasser AWM, Höner CB, Hollenberg CP (1991) An efficient transformation procedure enabling long-term storage of competent cells of various yeast genera. *Yeast* 7:691–692
- Forsburg SL, Guarente L (1989) Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. *Genes Dev* 3:1166–1178
- Freire-Picos MA, Rodriguez-Torres AM, Ramil E, Cerdan ME, Breunig KD, Hollenberg CP, Zitomer RS (1993) Sequence of a cytochrome *c* gene from *Kluyveromyces lactis* and its upstream region. *Yeast* 9:201–204
- Garnier J, Osguthorpe DJ, Robson B (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* 120:97–120
- Gietz RD, Sugino A (1988) New-yeast *Escherichia coli* shuttle vectors constructed with in vitro mutagenized genes lacking six-base pair restriction sites. *Gene* 74:527–534
- Grivell LA (1989) Nucleo-mitochondrial interactions in yeast mitochondrial biogenesis. *Eur J Biochem* 182:477–493

- Guarente L, Mason T (1983) Heme regulates transcription of the *CYC1* gene of *Saccharomyces cerevisiae* via an upstream activation site. *Cell* 32:1279–1286
- Guarente L, Lalonde B, Gifford P, Alani E (1984) Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. *Cell* 36:503–511
- Hahn S, Pinkham J, Wei R, Miller R, Guarente L (1988) The HAP3 regulatory locus of *Saccharomyces cerevisiae* encodes divergent overlapping transcripts. *Mol Cell Biol* 8:655–663
- Heus JJ, Bloom KS, Zonneveld BJM, Steensma HY, VandenBerg JA (1993) Chromatin structure of *Kluyveromyces lactis* centromeres in *K. lactis* and *Saccharomyces cerevisiae*. *Chromosoma* 102:660–667
- Hooft van Huijsduinen RAM, Bollekens J, Dorn A, Benoist C, Mathis D (1987) Properties of a CCAAT box-binding protein. *Nucleic Acids Res* 15:7265–7282
- Hooft van Huijsduijnen R, Li XY, Black D, Matthes H, Benoist C, Mathis D (1990) Co-evolution from yeast to mouse: cDNA cloning of the two NF-Y (CP-1/CBF) subunits. *EMBO J* 9:3319–3127
- Irniger S, Sanfacon H, Egli CM, Braus GH (1992) Different sequence elements are required for function of the Cauliflower Mosaic Virus polyadenylation site in *Saccharomyces cerevisiae* compared with in plants. *Mol Cell Biol* 12:2322–2330
- Klein P, Kanehishisa M, DeLisi C (1985) The detection and classification of membrane spanning proteins. *Biochim Biophys Acta* 815:468–476
- Kulmburg P, Prange T, Mathieu M, Sequeval D, Scazzocchio C, Felenbok B (1991) Correct intron splicing generates a new type of a putative zinc-binding domain in a transcriptional activator of *Aspergillus nidulans*. *FEBS Lett* 280:11–16
- Li XY, Hooft van Huijsduijnen RH, Mantovani R, Benoist C, Mathis D (1992a) Intron-exon organization of the NF-Y genes - tissue-specific splicing modifies an activation domain. *J Biol Chem* 267:8984–8990
- Li XY, Mantovani R, Hooft van Huijsduijnen R, Andre I, Benoist C, Mathis D (1992b) Evolutionary variation of the CCAAT-binding transcription factor NF-Y. *Nucleic Acids Res* 20:1087–1091
- Maity SN, Vuorio T, Decrombrugghe B (1990) The B-subunit of a rat heteromeric CCAAT-binding transcription factor shows a striking sequence identity with the yeast HAP2 transcription factor. *Proc Natl Acad Sci USA* 87:5378–5382
- Mattoon JR, Caravajal E, Guthrie D (1990) Effects of HAP mutations on heme and cytochrome formation in yeast. *Curr Genet* 17:179–183
- Mulder W, Scholten IHJM, Nagelkerken B, Grivell LA (1994a) Sequence of *KIQC8*. EMBL/Genbank Accession Number X76026
- Mulder W, Scholten IHJM, VanRoon H, Grivell LA (1994b) Sequence of *KIQC7*. EMBL/Genbank Accession Number X76027
- Olesen JT, Guarente L (1990) The HAP2 subunit of yeast CCAAT transcriptional activator contains adjacent domains for subunit association and DNA recognition: model for the HAP2/3/4 complex. *Genes Dev* 4:1714–1729
- Olesen J, Hahn S, Guarente L (1987) Yeast HAP2 and HAP3 activators both bind to the *CYC1* upstream activation site, UAS2, in an interdependent manner. *Cell* 51:953–961
- Olesen JT, Fikes JD, Guarente L (1991) The *Schizosaccharomyces pombe* homolog of *Saccharomyces cerevisiae* HAP2 reveals selective and stringent conservation of the small essential core protein domain. *Mol Cell Biol* 11:611–619
- Oliver SG et al. (1992) The complete DNA sequence of yeast chromosome III. *Nature* 357:38–46
- Pinkham JL, Guarente L (1985) Cloning and molecular analysis of the *HAP2* locus: a global regulator of respiratory genes in *S. cerevisiae*. *Mol Cell Biol* 5:3410–3416
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467
- Sathe GM, O'Brien S, McLaughlin MM, Watson F, Livi GP (1991) Use of polymerase chain reaction for rapid detection of gene insertions in whole yeast cells. *Nucleic Acids Res* 19:4775
- Soni R, Murray JAH (1992) A rapid and inexpensive method for isolation of shuttle vector DNA from yeast for the transformation of *E. coli*. *Nucleic Acids Res* 20:5852
- Struhl K (1989) Helix-zinc-helix, zinc-finger and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. *Trends Biochem Sci* 14:137–140
- Van Loon APM, Van Eijk E, Grivell LA (1983) Biosynthesis of the ubiquinol-cytochrome c reductase complex in yeast. Discoordinate synthesis of the 11-kd subunit in response to increased gene copy number. *EMBO J* 2:1765–1770
- VanDyck L, Pearce DA, Sherman F (1994) *PIM1* encodes a mitochondrial ATP-dependent protease that is required for mitochondrial function in the yeast *Saccharomyces cerevisiae*. *J Biol Chem* 269:238–242
- Vuorio T, Maity SN, Decrombrugghe B (1990) Purification and molecular cloning of the A-chain of a rat heteromeric CCAAT-binding protein—sequence identity with the yeast HAP3 transcription factor. *J Biol Chem* 265:22480–22486
- Zitomer RS, Lowry CV (1992) Regulation of gene expression by oxygen in *Saccharomyces cerevisiae*. *Microbiol Rev* 56:1–11