Two Proteins with Different Functions Are Derived from the \textit{KlHEM13} Gene\textsuperscript{7}

Ángel Vizoso Vázquez, Moisés Blanco,† Justyna Zaborowska,‡ Pilar Soengas,§
M. Isabel González-Siso, Manuel Becerra, Esther Rodríguez-Belmonte, and M. Esperanza Cerdañ*

Departmento Biología Celular y Molecular, Universidad de A Coruña, Facultad de Ciencias, Campus de A Zapateira s/n 15071, A Coruña, Spain

Received 5 May 2011/Accepted 30 July 2011

Cell complexity depends on regulatory mechanisms that control gene expression and adapt levels of specific proteins to internal and environmental changes. One gene can produce one protein or several proteins through mechanisms controlling transcription, mRNA processing, translation, and post-translational events. In yeast, changes in gene transcription and the regulation of mRNA translation play critical roles in the response to oxygen availability.

In the yeast \textit{Kluyveromyces lactis}, the \textit{HEM13} gene (\textit{KlHEM13}) encodes coproporphyrinogen oxidase (Cpo), which catalyzes the sixth step of heme biosynthesis. Cpo promotes oxidative decarboxylation of the 2- and 4-propionyl groups of coproporphyrinogen III to form the vinyl groups of protoporphyrinogen IX. Northern blot analyses have revealed that \textit{KlHEM13} is a hypoxic gene and that its transcription is regulated by oxygen and heme (10). In \textit{Saccharomyces cerevisiae}, the gene homologous to \textit{KlHEM13} is essential in the absence of a heme precursor under aerobicosis (normoxia, with approximately 20% O\textsubscript{2}) and hypoxic conditions (less than 0.05% O\textsubscript{2}). \textit{KlHEM13} complements a \textit{Δhem13} mutation in \textit{S. cerevisiae} and restores its capacity to grow in the absence of hemin (protoporphyrin IX). Northern blot analysis suggests that \textit{KlHEM13} is a hypoxic gene and that its transcription is regulated by oxygen and heme (10). In \textit{Saccharomyces cerevisiae}, the gene homologous to \textit{KlHEM13} is essential in the absence of a heme precursor under aerobicosis (normoxia, with approximately 20% O\textsubscript{2}) and hypoxic conditions (less than 0.05% O\textsubscript{2}). \textit{KlHEM13} complements a \textit{Δhem13} mutation in \textit{S. cerevisiae} and restores its capacity to grow in the absence of hemin (protoporphyrin IX) (2). Sequence alignment of the protein derived from the \textit{KlHEM13} open reading frame (ORF) with the \textit{S. cerevisiae} homolog demonstrates the presence of 52 additional amino acids at the N terminus of \textit{Kicpo}. This extension includes the amino acids between the first and second in-frame ATG codons in the \textit{KlHEM13} ORF. Therefore, computer-aided translation from the second in-frame ATG (the second translation initiation codon [TIC]) produces a \textit{Saccharomyces}-like Cpo, while translation from the first in-frame ATG (TIC-1) gives rise to an N-terminally extended Cpo (Fig. 1A).

This study demonstrates that in \textit{K. lactis}, \textit{KlHEM13}, a single-copy-number gene, produces two proteins that differ in the N-terminal region (l-Kicpo and s-Kicpo). Diverse evidence supports a mechanism in which the use of alternative transcription start sites (TSS) allows two different polypeptides to be synthesized. Complementation with s-Kicpo, but not with l-Kicpo, restores the growth of \textit{K. lactis} \textit{Δhem13} mutants in media without hemin supplementation. Differential use of TSS pools during hypoxia and other functional data indicate that l-Kicpo plays a role in the transcriptional regulation of \textit{KlHEM13} expression. A hypothetical model of the evolutionary origins and coexistence of these two proteins in \textit{K. lactis} is discussed.

**MATERIALS AND METHODS**

**Yeast strains and growth conditions.** The \textit{K. lactis} strains used in this study were NRRL-Y1140 (CBS2359), MW190-9B (MATa lac4-8 ura3-51 Rag1), and CBS4930 (CBS2359 \textit{Δ(ku80)}). The \textit{S. cerevisiae} strain BY4741 (MATa his3Delta1 leu2Delta2120 met15Delta0 ura3Delta0), used for efficient recombination, was obtained from EUROSCARF (http://web.uni-frankfurt.de/fb15/mikro/euroscarf/yeast.html). Yeast cells were grown at 30°C in YPD (2% glucose, 2% Bacto peptone, 1% yeast extract) or synthetic complete medium (CM) (29). Hypoxic conditions were generated in anaerobic jars with the AnaeroGen gas pack system (Oxoid Ltd.). For hypoxic growth, media were supplemented with Tween and ergosterol as described previously (2). When required, media were supplemented with hemin at a final concentration of 50 μg/ml.

**Nomenclature used for gene regions, primers, and plasmids.** In the numbering of sequences, the A in the first ATG of the \textit{KlHEM13} ORF was considered position +1; negative positions were counted 5’-wards and positive positions 3’-wards. Using this nomenclature for the \textit{KlHEM13} ORF, TIC-1 was at position +1, TIC-2 at position +157, and the TGA stop codon at position +1126. In the primer sequences, capital letters represent the genomic sequences and lowercase letters represent additional sequences added for design convenience. Restriction sites are shown in italics. The letters “w” and “c” at the end of primer names indicate the forward and reverse strands, respectively.

\textsuperscript{7}Published ahead of print on 5 August 2011.

\textsuperscript{*}Corresponding author. Mailing address: Dpto. de Biología Celular y Molecular, F. Ciencias, Universidad de A Coruña, Campus de A Zapateira s/n, A Coruña 15071, Spain. Phone: 34 9 81 167000. Fax: 34 9 81 167065. E-mail: esper.cerdan@udc.es.

†Present address: INIBIC, Complejo Hospitalario Universitario, As Xubias, A Coruña, Spain.

‡Present address: Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, United Kingdom.

§Present address: Misión Biológica de Galicia, CSIC, Apartado 28, 36080 Pontevedra, Spain.
Construction of the K. lactis Δhem13 strains. The KIHEM13 gene was amplified from MW190-9B genomic DNA using PCR with primers ECV185Mw (ggggacagtATTTTGGCATGGAAAGGAAGG, at position +1025) and ECV185Mc (ggggacagtACTAGTGGAGAGGCTGGT, at position +1751) and was cloned into the KpnI site of the YIplac204 vector (9). This construct was used as a template for inverse PCR carried out with the divergent primers ECV250Mc (ggggacagtCTAGCTAGTTTCCACCG, at position +1253) and ECV251Mw (ggggacagtGGGTCGACAATACTTTT, at position +1046) to obtain a linear cassette. In order to achieve the constructions of null phenotypes were accomplished using Cpo enzymatic assays and phenotypic analyses of the growth defect in CM in the absence of hemin. (D) Western blot of proteins expressed in MW190-9B from the pSki1-KIHEM13-3HA construct C-terminally tagged with HA and hybridized with anti-HA (left) or Ab-l-ΔCpo, which is directed against the peptide underlined at the N terminus in panel A (right). Hybridization with ab21054, an HRP-conjugated anti-histone H3 antibody from Abcam (H3), was used as a loading control. l-ΔCpo, protein translated from the first TIC; s-ΔCpo, protein translated from the second TIC. The strains were grown under aerobic (O2) or hypoxia (H). The molecular sizes of the proteins used as molecular markers (PageRuler prestained protein ladder; Fermentas) are given to the left of each panel.

**FIG. 1.** One gene and two proteins. (A) Alignment of the N-terminal regions of the ORFs translated from the KIHEM13 (Kl) and ScHEM13 (Sc) genes. The first and second in-frame Mets from the K. lactis protein are boldface and underlined. (B) Verification of the null strains. (Top) Schemes showing the primers used. (Bottom) PCR amplifications from MW190-9B (lanes 1, 3, 5, and 6) or MW190-9B-Δhem13 (lanes 2, 4, 7, and 8) DNA with primers ECV261w and ECV262c (lanes 1 and 2), ECV95w and ECV96c (lanes 3 and 4), ECV349Mw and K3 (lanes 5 and 7), or K2 and ECV185Mc (lanes 6 and 8). Lane M, 1-kbp GeneRuler (Fermentas). (C) Growth of wild-type and Δhem13 mutants were used as templates for PCR with primers ECV250Mc (ggggacagtGTGGAGAAGGTGCATGGG, at position +1025) and ECV185Mc (ggggacagtACGACGAAGTTTGTTATTATTTTTT, at position +1046) were designed flanking KIHEM13, but external to the regions of homology used for the recombination event. Amplification with each pair of external/internal primers produced PCR products of the expected sizes. The expected size of the fragment obtained with ECV261w and ECV262c (described above) and the absence of amplification of an internal fragment of the ORF were tested with primers ECV95w (GGGAGTCTCTGTACCGCAA, at position +687) and ECV96c (GAGACGGGATCCTGGAGT, at position +1008). Further verifications of null phenotypes were accomplished using Cpo enzymatic assays and phenotypic analyses of the growth defect in CM in the absence of hemin.

Cell fractionation. K. lactis cells (Δhem13 and wild type) were grown in heme-supplemented medium as described above. Cells were harvested at an optical density at 600 nm (OD600) of 2, washed (to remove adsorbed heme), and broken with glass beads (diameter, 0.45 mm) in 0.1 M potassium phosphate buffer, pH 7.2. The cell extract was centrifuged for 5 min at 3,000 × g to remove intact cells. The supernatant is referred to as the homogenate (H). An aliquot of the homogenate was centrifuged for 30 min at 45,000 × g to sediment membranous fractions containing nuclei (MF). The supernatant was referred to as the “soluble fraction” (SF). The pellet (the MF) was resuspended in potassium phosphate buffer and was solubilized by sonication.

Cpo activity. Cpo activities were measured as described previously (18) on a Photon Technology International (PTI) spectrophotometer in 1-ml cuvettes at 30°C. Tween 80, EDTA, diithiothreitol (DTT), and freshly prepared coproporphyrinogen solutions were added as described previously (18). The coupling enzyme was highly purified S. cerevisiae protoporphyrinogen oxidase (100 U/assay). The excitation wavelength was 410 nm, and the emission wavelength was 632 nm. One unit of Cpo activity is defined as the amount of enzyme that catalyzes the formation of 1 nmol of protoporphyrinogen per hour at 30°C under standard conditions.
C-terminal tagging of KlCpo. KlHEM13 was modified using the pET6a-3HA-kanMX6 cassette (20). The 3HA-kanMX6 cassette was amplified using PCR with two primers, ECV436Mw (CAAGTATTGGAAGACCCCTTGTAAGGTGTT gagatccgctgtaata) and ECV437Mw (ATCATAGAGACTCTCAAAATAAA AGGCATgaattgtgtgtatt), which were designed to fuse the KlHEM13 ORF in frame to a DNA sequence encoding three copies of the hemagglutinin (HA) epitope. They had regions of homology to the pET6a-3HA-kanMX6 cassette (lowercase letters in sequences) and to KlHEM13 (capital letters), near a BsmI restriction site at position +1127 that overlapped with the TGA stop codon of KlHEM13. Plasmid pSK1-KlHEM13 (2) was linearized by partial digestion with BsmI and was then transformed into the 3HA-kanMX6 PCR product to cotransform the S. cerevisiae strain BY4741. The transformed cells were selected in CM without uracil (CM-Ura) and CM-Ura plates with 40 μg/ml G418. Growth in CM-Ura selected cells carrying plasmid pSK1-KlHEM13, and growth in media with G418 selected recombinant plasmids carrying the tagged protein. The recombinant plasmid was isolated from the transformed BY4741 cells and was amplified in DH10B bacteria by using LB medium supplemented with kanamycin (20 μg/ml) for selection. After verification by sequencing, plasmid pSK1-KlHEM13-3HA was used to transform the K. lactis strain MW190-9B.

Protein extracts. Protein extracts for Western blotting and chromatin immunoprecipitation (ChIP) assays were obtained in phosphate-buffered saline (PBS) (0.8% NaCl, 0.02% KCl, 0.144% Na2HPO4, 0.024% KH2PO4 [pH 7.4]), and those for the measurement of β-galactosidase activity were obtained in Z-Buffer, as described previously (2). Protein concentrations were determined using the method of Bradford (3).

Western blotting. Western blotting and control of loading were carried out as described previously (1). The HA probe (F-7; sc-7392; Santa Cruz Biotechnology, Inc.), directed against the HA-tagged proteins (l-KlCpo and s-KlCpo) at their C termini, was used according to the manufacturer's instructions. A polyclonal antibody specific to KlCpo, directed to its N terminus (Ab-1126; Sigma-Aldrich Co.), was obtained by chemical synthesis of the TETOPRPHVLTHSK peptide (positions T34 to T47) and rabbit immunization (at the facilities of GenScript USA Inc.). Membranes were incubated with a 1:325 dilution of Ab-1126 in PBS with 0.25% nonfat dry milk and 0.1% Tween 20 at room temperature for 25 min. The secondary antibody, horseshadish peroxidase (HRP)-conjugated goat anti-rabbit IgG (sc-2004; Santa Cruz Biotechnology, Inc.), was diluted 1:200 in PBS with 0.1% Tween 20, and membranes were incubated at room temperature for 25 min. An HRP-conjugated antibody against human histone H3 (ab21054; Abcam) was used as a loading control, since it cross-reacts with the homologous protein from K. lactis. The immune complexes on the membrane were visualized using dianaminobenzidine (Sigma-Aldrich Co.) and H2O2.

Protein identification by mass spectrometry. Samples were analyzed by matrix-assisted laser desorption–tandem time of flight (MALDI-TOF/TOF) mass spectrometry by following procedures described previously (24). Peptides were identified using ProteinPilot software, version 4.0 (AB Sciex Pte. Ltd.). Mass spectrometry by following procedures described previously (24). Peptides were identified using ProteinPilot software, version 4.0 (AB Sciex Pte. Ltd.).

β-Galactosidase activity. Protein extracts for Western blotting and control of loading were carried out as described previously (1). The HA probe (F-7; sc-7392; Santa Cruz Biotechnology, Inc.), directed against the HA-tagged proteins (l-KlCpo and s-KlCpo) at their C termini, was used according to the manufacturer's instructions. A polyclonal antibody specific to KlCpo, directed to its N terminus (Ab-1126; Sigma-Aldrich Co.), was obtained by chemical synthesis of the TETOPRPHVLTHSK peptide (positions T34 to T47) and rabbit immunization (at the facilities of GenScript USA Inc.). Membranes were incubated with a 1:325 dilution of Ab-1126 in PBS with 0.25% nonfat dry milk and 0.1% Tween 20 at room temperature for 25 min. The secondary antibody, horseshadish peroxidase (HRP)-conjugated goat anti-rabbit IgG (sc-2004; Santa Cruz Biotechnology, Inc.), was diluted 1:200 in PBS with 0.1% Tween 20, and membranes were incubated at room temperature for 25 min. An HRP-conjugated antibody against human histone H3 (ab21054; Abcam) was used as a loading control, since it cross-reacts with the homologous protein from K. lactis. The immune complexes on the membrane were visualized using dianaminobenzidine (Sigma-Aldrich Co.) and H2O2.

Protein identification by mass spectrometry. Samples were analyzed by matrix-assisted laser desorption–tandem time of flight (MALDI-TOF/TOF) mass spectrometry by following procedures described previously (24). Peptides were identified using ProteinPilot software, version 4.0 (AB Sciex Pte. Ltd.).

Directed mutagenesis. Plasmid pSK1-KlHEM13-3HA was subjected to directed mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
direct mutagenesis in order to eliminate the codon for the second in-frame Met to Ala, and ECV120Mc (AGGCGAAGAATTTTACTCCTG) were used to
RESULTS

Only one gene encodes Cpo in K. lactis. A null allele of KIHEM13 was constructed in the K. lactis haploid strains CBSA80 and MW190-9B. Correct integration of the disruption cassette was verified and is demonstrated for strain MW190-9B (Fig. 1B). Similar results were obtained with CBSA80 (data not shown). Both ΔKlhem13 null strains were unable to grow in the absence of hemin (Fig. 1C). Cpo enzymatic activity was determined in CBSA80 and its CBSA80-ΔKlhem13 derivative using fluorometric assays (18). Cpo enzymatic activity was absent in the CBSA80-ΔKlhem13 strain. Cpo activity in the strain with wild-type HEM13, CBSA80, was detected only in the soluble fraction (40 ± 4 U/μg protein), not in the membranous fraction. This result is in agreement with the subcellular location of Cpo activity in the soluble cytosolic fraction of K. lactis cells, as reported for S. cerevisiae (17). Analysis of the complete sequence of the K. lactis genome with BLAST tools at Genolevures (http://www.genolevures.org) yielded no other K. lactis ORFs with homology to HEM13 genes. These data confirm that only one gene, KIHEM13 (KLLA0F18546g), encodes a functional Cpo enzyme in K. lactis.

Two KICpo proteins are translated from KIHEM13. In order to ascertain which of the two in-frame TICs present in KIHEM13 was used for translation, a recombinant K. lactis strain expressing KICpo tagged at the C terminus with three copies of HA was constructed. After growth of this strain under aerobic or hypoxic conditions, protein extracts were obtained; 30 μg of total protein was fractionated using electrophoresis in 12% polyacrylamide gel electrophoresis (PAGE) gels, and Western blotting was performed with an HA-specific primary antibody. Two proteins with different Mr were obtained (Fig. 1D). The Mr calculated for the slowly migrating band had the expected size for the KICpo-3HA construct translated from TIC-1 (l-KICpo). The Mr calculated for the rapidly migrating band was coincident with that expected for the protein translated from TIC-2 (s-KICpo). In order to verify the nature of these two proteins, the corresponding bands were fractionated by PAGE, excised from the gel, and analyzed using MALDI-TOF/TOF. The sequences of the internal peptides obtained confirmed that the two proteins were KICpo, but peptides from the N terminus and C terminus were not present in the pools detected, and therefore it was not possible to verify by these methods whether the proteins differed in their N-terminal regions. In order to test whether l-KICpo contained the N-terminal region corresponding to translation from TIC-1 and whether s-KICpo lacked this region, Western blotting was performed using a specific antibody (Ab-l-KICpo) directed against a peptide encoded by the region between the first and second TICs (underlined in Fig. 1A). The results (Fig. 1D) revealed that this antibody interacts with l-KICpo but not with s-KICpo. Considering that (i) both proteins contained the C-terminal tag, (ii) only l-KICpo interacted with Ab-l-KICpo, (iii) experimentally calculated Mr were as expected from TIC-1 and TIC-2 translation, and (iv) no introns could be identified when KIHEM13 was analyzed using the yeast intron program EXPLORA (14), it was concluded that the two KICpo proteins detected were derived from translation initiation from TIC-1 (l-KICpo) and TIC-2 (s-KICpo). The production of different proteins by translation from different TICs has been described previously in eukaryotic systems, including yeasts (21, 22).

KIHEM13 is transcribed from two different tss pools, and selection of a specific tss depends on the availability of oxygen and heme. To better characterize the promoter regions in the 5' upstream region of KIHEM13 and their relative positions to the TICs, the precise positions of tss on the KIHEM13 promoter were determined using primer extension. The data (Fig. 2) confirmed the presence of two tss pools. One pool (tss-1) was located 5' of TIC-1 and included three weak starts at G positions −87, −80, and −69. A second tss pool (tss-2), between TIC-1 and TIC-2, included two strong starts at C positions +46 and +56 (positions −111 and −101 if the relative positions from TIC-2, which would be the only possible starting AUG codon for an mRNA transcribed from this tss-2 pool in this case, are considered). The positions of the determined tss relative to the TICs are represented on the KIHEM13 sequence (Fig. 2B). The tss pools were detected during aerobicosis or hypoxia. However, the tss at position −69 in the tss-1 pool is specific for hypoxia and disappeared when hemin was added (Fig. 2A). The long transcript could be produced to a minor extent from tss-1 under both aerobic and hypoxic conditions. The short transcript would be more abundant than the larger transcript under both conditions, but production of the long transcript would increase during hypoxia owing to heme-regulated use of the tss at the −69 position (Fig. 2C).

In order to confirm the functionality of these tss pools, total RNA was isolated from K. lactis cells grown under aerobicosis or hypoxia, and cDNA was obtained. This cDNA was used as a template in PCR amplifications using two pairs of primers, ECV551Mw/ECV550Mc and ECV548Mw/ECV550Mc. The long transcript, starting 5' of TIC-1, is present in the total RNA isolated from K. lactis cells under aerobic and hypoxic conditions (Fig. 3A). The upstream regions of KIHEM13 from −1025 to TIC-1 (pTIC-1-lacZ) and to TIC-2 (pTIC-2-lacZ) were fused with the reporter gene lacZ, and β-galactosidase activity in protein extracts obtained from transformed K. lactis cells after aerobic and hypoxic growth was measured. The results presented in Fig. 3B indicate that the sequence located upstream of TIC-1 had poor activity in terms of promoting the expression of the reporter gene. When this sequence was extended 3' of TIC-2, its promoter activity increased by 3 orders of magnitude. From these data it was concluded that the importance of the region between TIC-1 and TIC-2 in the control of KIHEM13 transcription, under aerobic and hypoxic conditions, is due to the presence of two major initiation sites in the tss-2 pool. However, minimal transcription is possible through the tss-1 pool in the region upstream of TIC-1. These results are compatible with the use of alternative tss to generate two KIHEM13 transcripts differing at their 5' termini.

s-KICpo but not l-KICpo expression complements the heme deficiency of the ΔKlhem13 mutant strain. The results indicated that the two proteins, l-KICpo and s-KICpo, were obtained from the single gene KIHEM13. Therefore, the question of whether the two proteins were functionally equivalent was investigated. Constructs that expressed only one of the two proteins were obtained from a heterologous LAC4 promoter, avoiding the co-expression caused by the use of the native KIHEM13 promoter. The pCXJ18-l-KICpo (expressing l-KICpo) and pCXJ18-s-
KlCpo (expressing s-KlCpo) constructs and the control pCXJ18-KlHEM13 construct (expressing both l-KlCpo and s-KlCpo) are described in Materials and Methods. These plasmids were used to transform the MW190-9B and MW190-9B-H9004 hem13 strains, and the selected transformed cells were plated on the selective synthetic medium CM-Ura with galactose as a carbon source in the absence or presence of hemin. Complementation of the heme deficiency of the H9004 hem13 mutant strain implies that the protein expressed from the plasmid construct has Cpo activity. This was evident with the pCXJ18-s-KlCpo construct and the control pCXJ18-KlHEM13 construct but not with pCXJ18-l-KlCpo. Therefore, only s-KlCpo complements the mutant phenotype (Fig. 4A). However, despite the lack of complementation, expression of l-KlCpo from the pCXJ18-KlHEM13 and pCXJ18-l-KlCpo constructs in the H9004 hem13 strain was confirmed using Western blotting with the l-KlCpo-specific antibody (Fig. 4B). In conclusion, l-KlCpo is expressed from these constructs but lacks Cpo activity.

Autoregulation of the KlHEM13 promoter. The l-KlCpo protein does not complement the heme deficiency of a ΔHem13 mutant strain (Fig. 4A). In addition, Cpo activity in the K. lactis wild-type strain was present only in the cytosolic soluble fraction (SF). These data excluded the possibility that s-KlCpo could be present in the membranous fraction that contains the nuclei (MF) but did not exclude the possibility that l-KlCpo could be present in this fraction, since it has no catalytic activity. Western blotting was carried out with protein extracts from both fractions, using anti-HA and the specific antibody Ab-l-KlCpo. The data presented in Fig. 5A demonstrate that l-KlCpo is found in both the soluble and membranous fractions, while s-KlCpo is present only in the soluble fraction. As expected, histone H3 was detected in the total extract (H) and the MF but not in the SF.

In order to explore whether l-KlCpo could have a nuclear function and regulate KlHEM13 transcription, the interaction between the protein and the promoter was investigated in vivo using ChIP. The K. lactis strain MW190-9B was transformed with pSK1-KlHEM13-HA; protein extracts were obtained after aerobic or hypoxic growth; and ChIP assays were performed. The results revealed that under hypoxic conditions, an HA-tagged immune-precipitated protein was complexed with DNA in a region that extended from −494 to −200 in the KlHEM13 promoter and with a smaller subregion from −434 to −200 in the same promoter (Fig. 5B, lanes 8). To prove unambiguously that l-KlCpo, and not s-KlCpo, was binding to the KlHEM13 promoter, a mutant KlHEM13 gene without the ATG codon for the second in-frame Met was obtained (pSK1-KlHEM13-3HA-2TIC mutant). The K. lactis strain MW190-9B was transformed with this mutant construct, and ChIP assays were performed after growth during hypoxic conditions. The results obtained confirm that l-KlCpo binds to the KlHEM13 promoter (Fig. 5C).

A functional role for this region of the KlHEM13 promoter, which is coimmunoprecipitated with the protein in ChIP assays, was evident in terms of autorepression during hypoxia. The promoter region of KlHEM13 from −494 to +190 was fused in frame to the lacZ reporter in the pXW1 plasmid to make the construct pTIC-2lacZ. Cells from the K. lactis strain MW190-9B and its ΔHem13 derivative were transformed with this construct, and β-galactosidase activity was

FIG. 2. Determination of transcriptional start sites in the KlHEM13 upstream region. (A) Primer extension analysis. Aerobic or hypoxic conditions are indicated as in Fig. 1. +Hem, medium supplemented with hemin, -, negative control. (B) Sequence of the upstream region of KlHEM13 and the ORF (partial). The positions of primers ECV300Mc and ECV405Mc are underlined. The ATGs of the first and second TICs and the experimentally determined tss are underlined. (C) Scheme showing the relative positions and use of tss pools and TICs.
determined in protein extracts obtained from cells grown under hypoxic conditions. Comparative analysis of reporter gene activity demonstrated that the expression directed by the \( \text{KlHEM13} \) promoter was lower in strain MW190-9B than in its \( \text{Klhem13} \) derivative (Fig. 5D). These data are compatible with a regulatory feedback control on \( \text{KlHEM13} \) expression by l-\( \text{KlCpo} \).

**DISCUSSION**

Sequence alignment of the protein derived from \( \text{KlHEM13} \) with \( \text{S. cerevisiae} \) Cpo shows the presence of 52 additional amino acids at the N terminus of \( \text{KlCpo} \). This extension includes the amino acids between two in-frame TICs in the \( \text{KlHEM13} \) ORF. Other proteins from ascomycetes with homology to \( \text{KlCpo} \) accessed through Ge `nolevures (http://www.genolevures.org/) do not have similar N-terminal extensions. This N-terminal extension does not contain a known sorting signal for either nuclear localization or other cellular localization as assayed with PSORT (http://psort.hgc.jp/).

\( \text{KlHEM13} \) is a single-copy-number gene in \( \text{K. lactis} \), and Cpo activity depends on its expression (Fig. 1B and C). Two tss pools (tss-1 and tss-2) have been mapped to the 5’/H11032 region of \( \text{KlHEM13} \) (Fig. 2). The relative positions of the tss pools with reference to the TICs indicate that initiation of transcription from tss-1 would produce mRNA molecules containing TIC-1 and TIC-2, while initiation from tss-2 would produce mRNA molecules containing only TIC-2. Two proteins with different functions, l-\( \text{KlCpo} \) and s-\( \text{KlCpo} \), are produced from this single gene. The s-\( \text{KlCpo} \) protein has Cpo activity and complements the \( \text{Klhem13} \) deletion, unlike l-\( \text{KlCpo} \) (Fig. 4). Some evidence suggests that l-\( \text{KlCpo} \) (the only form present in the membranous fraction containing the nuclei) has a regulatory role in terms of transcriptional regulation of \( \text{KlHEM13} \) during hypoxia (Fig. 5). The utilization of multiple tss, frequently very close to one another and forming tss pools, is a frequent event in yeasts, but its functional significance is not fully understood (11). The alternative use of tss has been related to the production of protein isoforms with the same function but different cellular locations. Genes containing more than one in-frame TIC in their 5′ regions may give rise to isoforms with different subcellular localizations due to heterogeneous N-terminal ends, depending on which of the AUG codons is used to initiate translation. Often, such genes have alternative transcription start sites, leading to the synthesis of two different mRNAs. For example, the genes coding for the yeast aminoacyl-tRNA synthetases (21) or glutathione reductase (22) use this mechanism. Generally, however, this has not been related to the utilization of alternative promoters or to the production of proteins with various functions, as reported here for \( \text{KlHEM13} \). To our knowledge, this is the first report of the use of alter-
native tss as a mechanism for producing functionally different proteins in budding yeasts. However, this phenomenon has been described in the fission yeast Schizosaccharomyces pombe, although in that system one of the two transcripts is poorly translated, and its function remains to be elucidated (25).

Alternative selection of tss as a mechanism for producing different proteins has been described in plants and mammals (6, 15). Examples of regulatory mechanisms based on tss selection have been reported in higher eukaryotes. Start site selection within many murine core promoters differs among tissues, and dynamic usage of tss in these promoters is associated with CpG islands, promoter structures, and imprinting (15). In the KlHEM13 promoter, transcription from the tss-2 pool results in s-KlCpo production under aerobic and hypoxic conditions (Fig. 2A and 3A). This is in accordance with the finding that only s-KlCpo is able to complement the heme deficiency caused by hem13 deletion in K. lactis (Fig. 4A). Cpo activity is necessary for cell viability during aerobicosis and hypoxia, and transcription from tss-2 would ensure cell survival under these conditions. The tss at position −69 in the tss-1 pool is specific for hypoxia and disappears in the presence of hemin (Fig. 2A). This feature argues in favor of l-KlCpo having a function that is more relevant during hypoxia. This hypothesis is supported by the finding that during hypoxia, a specific complex is formed that contains the tagged protein and a promoter region of KlHEM13 (delimited from −434 to −200 in our experiments [Fig. 5B and C]), which is also implied in the control of KlHEM13 expression under this condition (Fig. 5D). However, analysis of the l-KlCpo sequence does not reveal any known DNA-binding domains. It is possible that in vivo, the interaction of l-KlCpo with DNA may be anchored by another, as yet unknown regulatory factor that binds directly to the KlHEM13 promoter during hypoxia. However, the data presented here do not preclude a direct interaction.

A functional explanation for the repression exerted by l-KlCpo on the KlHEM13 promoter could be that this regulatory mechanism is directed to avoiding toxicity caused by an excess of Cpo activity that could eventually cause accumulation of the product, protoporphyrinogen IX, and consequently a bottleneck in subsequent steps of the heme biosynthetic pathway. Indeed, this compound accumulates when protoporphyrinogen oxidase, the next enzyme in the biosynthesis of heme, has decreased activity, and this metabolic failure produces variegate porphyria in humans (7).

An intriguing question concerns why the l-KlCpo form lacks catalytic activity. There are precedents, proteins in which a terminal extension blocks the catalytic site. For instance, proteases lack enzymatic activity until they are processed. Of course, processing is not the only mechanism suitable for opening the catalytic site, and conformational changes produced by interaction with other protein partners or regulatory molecules have to be considered as well. The crystal structure of S. cerevisiae Cpo has been resolved (23), and on the basis of sequence homology, a similar structure is predicted for s-KlCpo (data not shown). However, homology modeling of the N-terminal extension present in l-KlCpo is not possible, because there are no homologous models for this region. Currently, research is under way to purify and crystallize the two KlCpo forms, l-KlCpo and s-KlCpo, in order to determine their structures and investigate correlations between structure and function.

With regard to the evolutionary origin and selective pressure for conservation of the various functions identified in the l-KlCpo and s-KlCpo proteins, it can be speculated that the “first” function positively selected in K. lactis could be the high catalytic Cpo activity, essential for life and carried by s-KlCpo. During hypoxia, KlHEM13 transcription is significantly increased, and accumulation of protoporphyrinogen IX could become toxic. In this context, a new compatible function that decreased the expression of the catalytically active form would be favorable for selection, and a new function for l-KlCpo could be established (Fig. 6). It is interesting to speculate about the mechanism by which this new function could arise. It is
likely that initially the gene was transcribed predominantly from the tss-2 pool, where the major transcription initiation sites still remain. Therefore, the short transcript that was translated to the catalytically active s-KlCpo would be actively produced. The new heme-regulated use of the transcriptional start site at position /H11002/69 in the tss-1 pool could increase the expression of the long transcript during hypoxia. To understand evolutionary positive selection for the coexistence of these two functions, l-KlCpo expression can be considered to have three advantages. First, it does not exclude s-KlCpo translation. Second, it lacks catalytic activity. Third, it is able to fine-tune hypoxic KlHEM13 gene expression, preventing the protoporphyrinogen IX excess that could be associated with overproduction of s-KlCpo.

Also of interest is the relationship between these two protein variants in K. lactis, l-KlCpo and s-KlCpo, and the phenomenon of moonlighting. Applying the moonlighting criteria (8, 12) to the coexistence of l-KlCpo and s-KlCpo in K. lactis, we find two coincidences: (i) they derive from a unique gene (without gene
fusion or alternative splicing) but perform two different functions, and (ii) the two functions are independent in the sense that inactivation of one of the functions, as happens with the catalytic activity in l-KICpO, does not affect the regulatory function. Still, we cannot strictly apply the term “moonlighting” to l-KICpO and s-KICpO, because we are able to fractionate them by PAGE, but we may consider them “moonlighting-like.”

In conclusion, two proteins with different N termini and different functions are derived from the single-copy-number KHEM13 gene by a mechanism based on tss selection and the use of two TICs.

ACKNOWLEDGMENTS

We thank Jean Michel Camadro (Institut Jacques Monod, Paris, France) for collaboration with the Cpo assays carried out with the K. lactis strains CBSâKU80 and CBSâKU80-hem13 in his laboratory. We thank M. Wesolowski-Louvel (Université Claude Bernard, Villeurbanne, France) for strain MW190-B and Yde Steensma (Delft University of Technology, Delft, The Netherlands) for strain CBSâKU80. MALDI-TOF/TOF analyses were performed at the “Unidad de Proteómica-Nodo asociado a ProteoRed, INIBIC-Complejo Hospitalario Universitario de A Coruña,” A Coruña, Spain. We are grateful to Cristina Ruiz-Romero and Jesús Mateos Martin for valuable assistance with these techniques. We thank BioMedES for English language editing.

This research was supported by grants BFU2006-03961/BMC and BFU2009-08854 from MICINN (Spain), cofinanced by FEDER (CEE). General support of the laboratory from 2008 to 2011 was funded by Xunta de Galicia (Consolidación C.E.O.U.2008/008), cofinanced by FEDER. A.V.V.’s salary was funded by “Lucas-Labrada” funds from Xunta de Galicia (Consolidacio ´n C.E.O.U.2008/008), cofinanced by FEDER. A.V.V.’s salary was funded by the “Lucas-Labrada program-2008” from Xunta de Galicia.

REFERENCES

17. Labbe, P. 1997. Purification and properties of coproporphyrinogen III oxi-
22. Outten, C. E., and V. C. Culotta. 2004. Alternative start sites in the Saccha-
romyces cerevisiae GLR1 gene are responsible for mitochondrial and cytoso-
26. Tizón, B., A. M. Rodríguez-Torres, and M. E. Cerdán. 1999. Disruption of six novel Saccharomyces cerevisiae genes reveals that YGL126c is necessary for growth in non-fermentable carbon sources, YGL126c for growth at low or high temperatures and YGL128c is implicated in the biosynthesis of methi-
ogous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast 10:1793–1808.