Appl Microbiol Biotechnol (2001) 55:734–741
DOI 10.1007/s002530100605

ORIGINAL PAPER

H.A. Kang · W.-K. Hong · J.-H. Sohn · E.-S. Choi
S.K. Rhee

Molecular characterization of the actin-encoding gene and the use of its promoter for a dominant selection system in the methylotrophic yeast *Hansenula polymorpha*

Received: 12 August 2000 / Received revision: 30 November 2000 / Accepted: 8 December 2000 / Published online: 7 April 2001

Abstract The actin gene (*ACT*) from the methylotrophic yeast *Hansenula polymorpha* was cloned and its structural feature was characterized. In contrast to the actin genes of other ascomycetous yeasts, which have only one large intron, the *H. polymorpha ACT* gene was found to be split by two introns. The *H. polymorpha ACT* introns were correctly processed in the heterologous host *Saccharomyces cerevisiae* despite appreciable differences in the splice site sequences. The promoter region of *H. polymorpha ACT* displayed two CCAAT motifs and two TATA-like sequences in a configuration similar to that observed in the *S. cerevisiae* actin promoter. A set of deleted *H. polymorpha ACT* promoters was exploited to direct expression of the bacterial hygromycin B resistance (*hph*) gene as a dominant selectable marker in the transformation of *H. polymorpha*. The resistance level of *H. polymorpha* transformants to the antibiotic was shown to be dependent on the integration copy number of the *hph* cassette. The selectivity of the hygromycin B resistance marker for transformants of higher copy number was remarkably increased with the deletion of the upstream TATA-like sequence, but not with the removal of either CCAAT motif, from the *H. polymorpha* promoter. The dosage-dependent selection system developed in this study should be useful for genetic manipulation of *H. polymorpha* as an industrial strain to produce recombinant proteins.

Introduction

Actin is one of the most highly conserved proteins in eukaryotic organisms. As a major component of microfilaments, actin is required in various cellular processes, such as motility, regulation of cell growth, and cell differentiation (Pollard 1990). In higher eukaryotic organisms, several actin genes (*ACT*) are expressed, although the different actins are highly similar in primary structure (Vandekerckhove and Weber 1978a). However, in lower eukaryotes, including slime mold (Vandekerckhove and Weber 1978b), several filamentous fungi (Dudler 1990; Matheucci et al. 1995), and various yeasts (Deshler et al. 1989; Losberger and Ernst 1989; Wery et al. 1996), only a single actin-encoding gene has been reported. Most actin genes from diverse species have been reported to contain introns. This has even been reported in several yeasts, in which intron-containing genes are exceptions rather than the rule (Woolford 1989). Incontrast to the extensive conservation of coding sequences among actin genes, a diversity of exon/intron organization in these genes has been observed. Thus the comparison of position and structure of actin introns has recently been used as a useful molecular biological method to understand the evolutionary process of actin genes in several organisms (Bagavathi and Malathi 1996; Wery et al. 1996).

Hansenula polymorpha is an ascosporogenic yeast capable ofusing methanol as its sole carbon and energy source. The biochemistry and physiology of methanol metabolism and peroxisome biogenesis have been the subject of extensive research on this yeast (Titorenko et al. 1993). During the recent past, along with another methylotrophic yeast, *Pichia pastoris*, *H. polymorpha* has also begun to draw attention as an attractive host system for the high-level expression of recombinant proteins (Hollenberg and Gellissen 1997). Transformations in *H. polymorpha* usually result in a variety of strains harboring various copy numbers of the heterologous DNA integrated into the host's genome in a head-to-tail arrangement (Gatzke et al. 1995). In *H. polymorpha*,even plasmids bearing an autonomous replication sequence exhibit a high frequency of heterologous integration in variable copy numbers. The feasibility of a stable multimeric integration of expression cassettes makes *H. polymorpha* an ideal host, especially for the coexpression of different proteins in a fixed ratio (Janowicz et al. 1991; Gellissen et al. 1996). The growing

H.A. Kang · W.-K. Hong · J.-H. Sohn · E.-S. Choi · S.K. Rhee (✉)
Microbial and Bioprocess Engineering Laboratory,
Korea Research Institute of Bioscience and Biotechnology,
Yusong P.O. Box 115, Taejon 305-600, Korea
e-mail: rheesk@mail.kribb.re.kr
Tel.: +82-42-8604450, Fax: +82-42-8604594

interest in this yeast for basic research and applications has demanded the development of genetic systems to analyze this organism at the molecular level. At present, however, only a few genes, except those involved in methanol metabolism and peroxisome biogenesis, have been cloned; and none of the intron-containing genes have yet been reported for this organism. Here we report the organization and structure of the *H. polymorpha* actin gene (*HpACT*), which contains two introns. In addition, we use the 5′-upstream region of *HpACT* as a homologous promoter, to direct expression of the *Escherichia coli* hygromycin B phosphotransferase gene in *H. polymorpha*.

Materials and methods

Strains and culture conditions

Yeast strains used in this study were *H. polymorpha* DL1-L (*leu2*), a derivative of *H. polymorpha* DL-1 (ATCC26012), and *Saccharomyces cerevisiae* 2805 (*MATa pep4::HIS3 prb-Δ1.6R can1 his3-20 ura3-52*), obtained from R.B. Wickner (NIH, Washington, D.C.). Yeast cells were cultivated in YPD medium (1% yeast extract, 2% peptone, and 2% glucose). Transformation and stabilization of *H. polymorpha* transformants were carried out as described by Sohn et al. (1999a), except that all LEU^+ transformants were cultivated with YNB broth (0.67% yeast nitrogen base without amino acids and 2% glucose) instead of YPD broth during the stabilization procedure. Selection for resistance to the antibiotic hygromycin B was carried out with YPD medium supplemented with different concentrations of hygromycin B, ranging from 0.15 g/l to 4.00 g/l (Sigma, St. Louis, Mo.).

Nucleic acid manipulations

Yeast chromosomal DNA was prepared according to the method of Holm et al. (1986) and yeast total RNA was isolated as described by Elion and Warner (1984). The DNA probes for Southern and Northern blot analysis were labeled, either with digoxigenin (DIG), using DIG-labeling kits (Roche Diagnostics, Mannheim, Germany), or with [α-^{32}P]ATP (Ambion, Austin, Tex.). The integration copy number of transformantswas determined by the quantitative reading of the Southern densitogram, using an image-analysis system as described by Sohn et al. (1999a). The reverse transcriptase-polymerase chain reaction (RT-PCR) was carried out using SuperScriptII RT (Gibco BRL, Grand Island, N.Y.). The two primers used for RT-PCR were GSP1 (5′-GGTTGGAATCATCTTAATTGAG-3′) and GSP2 (5′-CGAAGGAAACACGGCTCTTGGAG-3′), which correspond to the 5′ untranslated region (5′UTR) and the exon III of *HpACT*, respectively.

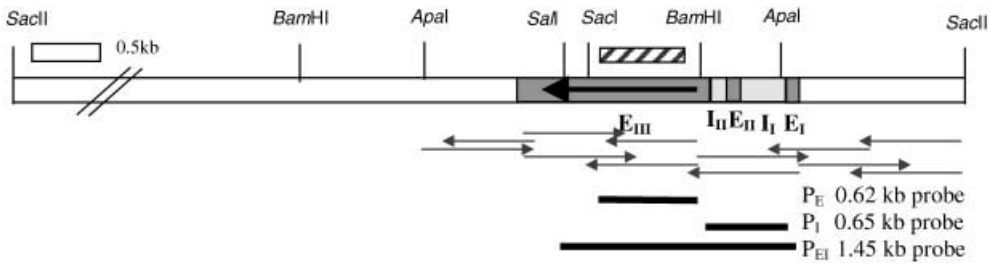
Fig. 1 Partial restriction map of the *Hansenula polymorpha ACT* gene. The 7.8-kb *Sac*II DNA fragment containing the whole coding region (~1.6 kb) of *HpACT* is shown, with the partial restriction enzyme sites. Sequencing strategy of the 3.0-kb *Apa*I/*Sac*II fragment is indicated by *arrows*. The 0.6-kb PCR fragment used in the cloning procedure of *HpACT* is indicated by a *hatched box*. The *dark-stippled areas* represent the exons (E_I, E_{II}, and E_{III}) and the *light-stippled areas* indicate the introns (I_I and I_{II}) of the *ACT* gene. The 0.62-kb exon-specific sequence probe (P_E), the 0.65-kb intron-specific sequence probe (P_I), and the 1.45-kb exon/intron-containing probe (P_{EI}), which were used in analyzing the *ACT* transcript, are represented by *closed bars*

DNA sequencing and analysis

DNA sequencing was performed by the dideoxy chain termination reaction using an ABI Prism dye terminator cycle sequencing ready-reaction kit (Perkin-Elmer, Foster City, Calif.). Comparison of nucleotide and deduced amino acid sequences was carried out through the BLAST and FASTA network server of the National Center for Biotechnology Information Database. The *ACT* genes were from *Candida albicans* (X16377; Losberger and Ernst 1989), *Kluyveromyces lactis* (M25826; Deshler et al. 1989), *S. cerevisiae* (J01310; Gallwitz and Sures 1980), *Schizosaccharomyces pombe* (Y00447; Mertins and Gallwitz 1987a), *Phaffia rhodozyma* (X89898;Wery et al. 1996), *Physarum polycephalum* (X07792; Adam et al. 1991), *Mus musculus* (M26689; Kim et al. 1989), and human (M20543; Taylor et al. 1988). The nucleotide sequence of *HpACT* has been submitted to the GenBank Nucleotide Sequence Database under Accession Number AF085278.

Construction of hygromycin B resistance cassette using the *HpACT* promoter

The *HpACT* promoters of 853, 242, 147, and 90 bp were amplified by PCR from the 7.8-kb *Sac*II fragment of *HpACT* (Fig. 1). The oligonucleotides used as sense primers (in each, the *Not*I site is underlined) were: HACT-F853 (5′-CCGGGCGGCCGCGGCCAACTCGTCTATCTTG-3′), HACT-F242 (5′-CCGGGCGGCCGCACGCAGTGCAGATGAGATTC-3′), HACT-F147 (5′-CCGGGCGGCCGCTCGCGAATATTCCAATCAGC-3′), and HACT-F90 (5′-CCGGGCGGCCGCTCTCAGTTTTTATAAAAGTT-3′). The oligonucleotide used as an antisense primer was HACT-R1 (5′-ggggaattcaggatcc*at*ttgctataactcaa-3′, with the *Bam*HI site underlined and the anticodon of the initiating ATG in italics). The PCR-amplified *HpACT* promoter (P_{HACT}) fragments contained a *Not*I site at the 5′ end and a *Bam*HI site just after the initiating ATG codon at the 3′ end to facilitate subsequent cloning procedures. A 0.24-kb N-terminal fragment of the *E. coli* hygromycin B phosphotransferase gene (*hph*) was amplified by PCR from pAN7-1 (Punt and van den Hondel 1992) using two primers, hph-F1 (5′-AATCGGATCCTGAACTCACCGCG-3′, with the *Bam*HI site underlined) and hph-R1 (5′-CCGCTCGAGGAATTCCCCAATGTCAATC-3′, with the *Eco*RI site underlined). The PCR-amplified *hph* N-terminal fragment contained a *Bam*HI site and an *Eco*RI site at the 5′ and 3′ ends, respectively. For the translational fusion with the P_{HACT} fragments, the *Bam*HI/*Eco*RI-digested N-terminal fragment of *hph* was ligated with the *Not*I/*Bam*HI-digested fragments of P_{HACT}. The resultant fragment, P_{HACT}-*hph*(N), was ligated with the 2.1-kb *Eco*RI/*Xba*I C-terminal fragment of *hph* fused with the terminator region of the *Aspergillus nidulans trpC* gene, derived from pAN7-1, generating the P_{HACT}-*hph* cassette (shown later in Results; see Fig. 5).



Results

Cloning of the *H. polymorpha ACT* gene

Alignment of the actin sequences from several yeasts and higher eukaryotes showed a high level of homology, not only at the protein level but also at the DNA level. To isolate the actin gene (*ACT*) in *H. polymorpha* by PCR techniques, we designed two degenerate primers, based on the DNA sequences of the exon II region of *S. cerevisiae ACT*. This region is highly conserved among the *ACT* genes of ascomycetous yeasts, to which *H. polymorpha* has been assigned (Barnett et al. 1990). Two primers, 5′-ATTGATAA(C/T)GG(A/T/C)TC(C/T)-GG(A/C)ATGTG-3′ and 5′-TCCTTGATGTC(A/C)CG-(A/G)ACAAT(C/T)TC-3′, were used to amplify a ~0.6-kb DNA fragment from the *H. polymorpha* chromosomal DNA. Sequencing of the PCR product revealed that the predicted amino acid sequence was almost identical to that of *S. cerevisiae ACT*, confirming that the PCR product was a partial fragment of *H. polymorpha ACT*. The 0.6-kb PCR fragment was used as a hybridization probe in Southern blot analysis of *H. polymorpha* chromosomal DNA digested with various enzymes to identify DNA fragments that may contain the whole *HpACT* gene. Only a single hybridization signal was detected in the digested DNA samples, suggesting a single actin gene in *H. polymorpha* (data not shown). The *Sac*II fragments of 7–8 kb were gel-eluted and ligated into the *Sac*II site of the bacterial plasmid pBluescript KS$^+$ (Stratagene, Kingsport, Tenn.). From the constructed library, positive clones hybridizing with the PCR product were isolated and sequenced. A 7.8-kb *Sac*II fragment was found to contain the whole coding region of *HpACT* (Fig. 1).

Sequence analysis and comparison of amino acids

The amino acid sequence, deduced from the nucleotide sequence of *HpACT*, showed that the gene contained a coding region for a protein of 376 amino acids (Fig. 2).

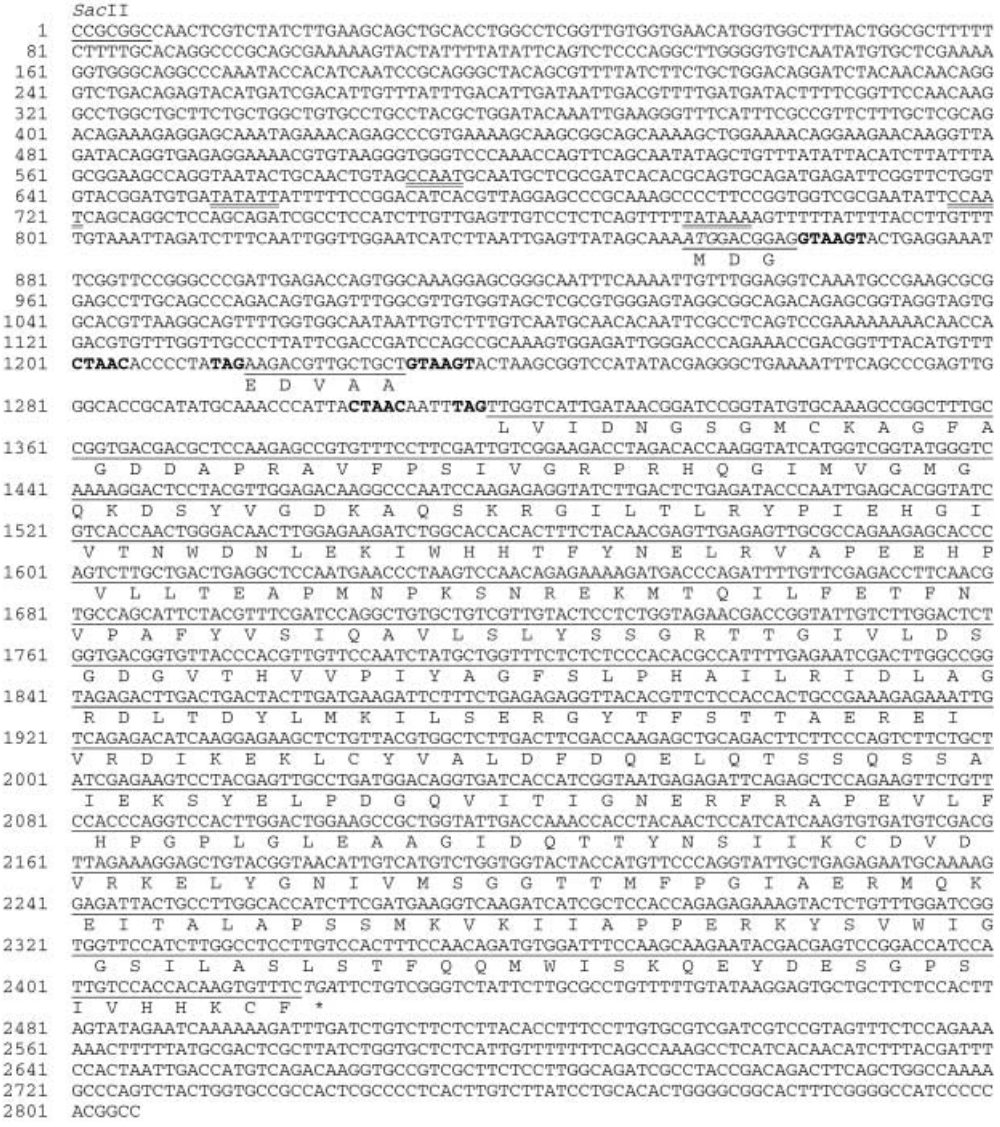


Fig. 2 Nucleotide sequence of the *H. polymorpha ACT* gene and its deduced amino acid sequences in single-letter code. The *Sac*II site is *underlined* and the *asterisk* denotes the stop codon of the *ACT* coding region. The amino acid sequences of exon I, exon II, and exon III are *shown below* their respective nucleotide sequences. Splice sites and branch sites are *highlighted in bold*. Putative CAAT motifs and TATA boxes are *double-underlined*

Interestingly, the coding region of ~1.6 kb appeared to be interrupted by two intervening sequences. Two pairs of the consensus sequences, corresponding to the 5′- and 3′-splice sites in fungal introns 5′-GTPuNGT-3′ and 5′-PyAG-3′ (Rambosek and Leach 1987), were found at the 5′-part of the coding region. The consensus sequences of the branch sites, 5′-CTPuAPy-3′ (Mertins and Gallwitz 1987b), were also found to be located 7–10 nucleotides upstream of the putative 3′-splice sites. These indicated that the first intron of 352 nucleotides was present within the triplet GAA coding for the fourth amino acid (glutamic acid) and that the second intron of 87 nucleotides was present between the eighth codon (alanine) and the ninth codon (leucine). Both introns appeared to be located near the 5′ terminus of *HpACT*, as observed in most of the intron-containing genes so far identified in yeasts (Woolford 1989).

A comparison of the deduced amino acid sequence with others showed the highest homology was with the actin of *S. cerevisiae* (94.6% identity). High amino acid sequence identity was also observed with those of other ascomycetous yeasts, *K. lactis* (93.6% identity) and *Candida albicans* (92.9% identity). In contrast, a relatively lower homology was shown with actins of the fission yeast *Sch. pombe* (85.5% identity) and the basidiomycetous yeast *Pha. rhodozyma* (83.3% identity), reflecting the relatively remote phylogenic relationship of these yeasts. Similarly low levels of homology were observed with actins of the slime mold *Phy. polycephalum* (84.0% identity) and higher eukaryotes, *M. musculus* smooth muscle γ-actin (83.9%) and human skeletal α-actin (79.8%). These results suggest that comparisons of actin sequences are more informative in assessing phylogenetic relationships between closely related organisms rather than between remotely related organisms. *H. polymorpha* actin (376 amino acids) is equal in length to the mammalian muscle-type γ actins (Kim et al. 1989). This is one residue longer than other yeast actins but is shorter than human skeletal α-actin (Taylor et al. 1988) and rabbit muscle actins (Harris et al. 1992).

Determination of splice junctions of *HpACT* mRNA

To determine the transcript size of *HpACT*, we analyzed *H. polymorpha* total RNA using a 1.45-kb exon/intron-containing sequence (P_{EI}), a 0.62-kb exon III-specific sequence (P_E), and a 0.65-kb intron-specific sequence (P_I) (Fig. 1) as probes in Northern blot analysis. The hybridization with probes P_{EI} or P_E revealed a ~1.5-kb transcript, which was thought to correspond to the spliced form of *HpACT* mRNA (Fig. 3A, lanes 2, 3). However, the probe P_I containing intron I and II sequences did not generate any hybridization signal, possibly due to extremely low levels of the precursor *ACT* mRNA and the released *ACT* intron (Fig. 3A, lane 1).

To confirm the putative splice junctions of *HpACT* mRNA, we carried out a RT-PCR experiment and isolated the N-terminus part of *HpACT* cDNA. With two primers designed using the regions at the 5′UTR and exon III of *HpACT* (Fig. 4A), a 134-bp PCR product was amplified from *H. polymorpha* total RNA, as was expected from linking exons I, II, and III (Fig. 4B, lane 2). As a control, the same RT-PCR experiment was also carried out using total chromosomal DNA of *H. polymorpha* as template, which generated a 573-bp PCR product containing additional sequences from introns I and II (Fig. 4B, lane 3). Four clones of the 134-bp PCR product

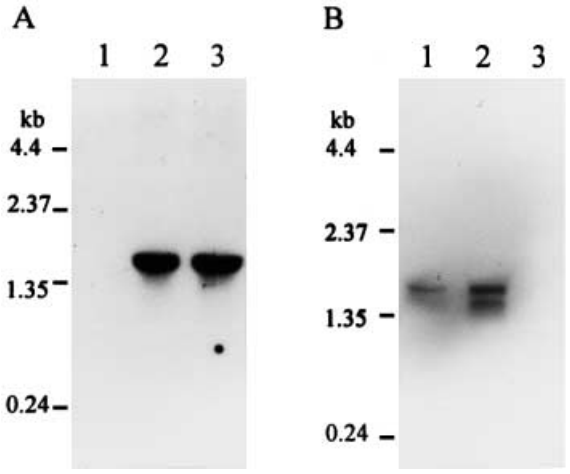


Fig. 3A,B Northern blot analysis of *H. polymorpha ACT* transcripts. **A** Analysis in *H. polymorpha*, using 20 µg of *H. polymorpha* total RNA fractionated on 1.8% formaldehyde gel and transferred onto a nylon membrane. Each lane was hybridized with a [^{32}P]ATP-labeled probe: 0.67-kb P_I (*lane 1*), 0.62-kb P_E (*lane 2*), and 1.45-kb P_{EI} (*lane 3*), respectively. RNA size markers are denoted *on the left*. **B** Analysis of the heterologous splicing in *Saccharomyces cerevisiae*. Total RNAs were prepared from *H. polymorpha* (*lane 1*), the recombinant *S. cerevisiae* harboring the plasmid HpACT-YEp352 (*lane 2*), and untransformed *S. cerevisiae* (*lane 3*). Hybridization of the membrane was carried out with the 0.62-kb P_E probe labeled with digoxigenin

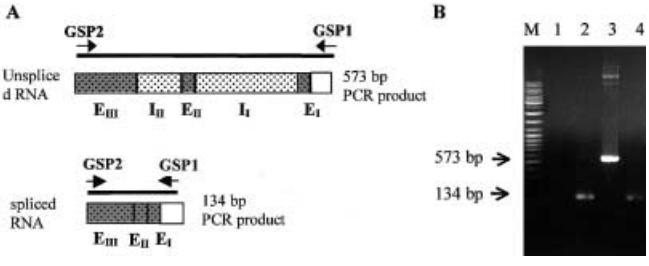


Fig. 4A,B Analysis of exon junctions of *H. polymorpha ACT* mRNA. **A** Diagram of the expected reverse transcriptase-PCR products from the unspliced and spliced *H. polymorpha ACT* mRNAs. The *dark- and light-stippled areas* represent the exons (E_I, E_{II}, and E_{III}) and the introns (I_I and I_{II}), respectively. The *empty box* represents the 5′UTR of *ACT* mRNA. **B** Photograph of PCR products. Each PCR was carried out with the primers GSP1 and GSP2, using different templates: *lane 1* no template, *lane 2 H. polymorpha* total RNA, *lane 3 H. polymorpha* chromosomal DNA, *lane 4* transformed *S. cerevisiae* total RNA, *lane M* DNA markers (10.0, 8.0, 6.0, 5.0, 4.0, 3.5, 3.0, 2.5, 2.0, 1.5, 1.2, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1 kb, MBI Fermentas, Lithuania)

Table 1 Intron consensus sequences of yeast actin genes. The nucleotide position is that of the intron within the gene. In the 5′ splice sequence, the exon/intron boundaries are indicated by a slash. Information in this table was obtained in part from Wery et al. (1996)

Species	Intron number	Intron size (bp)	Nucleotide position	5′ Splice sequence	Putative branch site	3′ Splice sequence
Candida albicans	1	656	9	GTG/GTATGT	TACTAAC	TAG/AAG
Hansenula polymorpha	1	352	11	GAG/GTAAGT	TTCTAAC	TAG/AAG
	2	87	377	GCT/GTAAGT	TACTAAC	TAG/TTG
Kluyveromyces lactis	1	777	7	CTG/GTATGT	TACTAAC	CAG/AGG
Saccharomyces bayanus	1	348	10	CAG/GTATGT	TACTAAC	TAG/AAG
S. cerevisiae	1	304	10	CAG/GTATGT	TACTAAC	TAG/AGG
Phaffia rhodozyma	1	338	3	CAT/GTACGT	TATTGAC	TAG/GGA
	2	98	363	CTC/GTATGT	TTCTTAT	TAG/GTA
	3	177	532	CCC/GTAAGT	TGCTGAT	CAG/CTC
	4	71	778	TGG/GTTCGT	TATTGAC	TAG/GGA
Consensus	–	–	–	GTPuNGT	CTPuAPy	PyAG

derived by the RT-PCR were sequenced and the result confirmed that the sequence corresponded exactly to the expected exon I–exon II–exon III linkage (data not shown).

Splicing of *HpACT* introns in *S. cerevisiae*

The introns of *Pha. rhodozyma ACT*, which show a noticeable difference in splice sequences from that of *S. cerevisiae ACT* (Table 1), cannot be spliced in *S. cerevisiae* (Wery et al. 1996). In contrast, the pre-mRNA of *K. lactis ACT*, which contains an intron resembling that of *S. cerevisiae ACT*, can efficiently be spliced in this heterologous yeast (Deshler et al. 1989). Therefore, we examined whether the two introns of *HpACT* could be correctly processed in the heterologous host *S. cerevisiae*. We transformed *S. cerevisiae* 2805 with the plasmid HpACT-YEp352, containing the 3-kb *Apa*I/*Sac*II fragment of *HpACT* in the *Sma*I site of YEp352 (Hill et al. 1986). The 3-kb *Apa*I/*Sac*II fragment was big enough to contain the promoter and terminator required for the *HpACT* pre-transcript to be expressed (Fig. 1). The total RNA from the *S. cerevisiae* transformed with HpACT-YEp352 was subjected to Northern blot analysis, probed with the 0.62-kb exon-specific sequence, P_E. A major signal of ~1.5 kb, the same size as the mature *ACT* transcript of *H. polymorpha*, was detected, indicating that *HpACT* mRNA was efficiently spliced in *S. cerevisiae* (Fig. 3B, lanes 1 vs. 2). No signal was detected in the untransformed *S. cerevisiae* (Fig. 3B, lane 3). We also obtained a 134-bp RT-PCR product from the transformed *S. cerevisiae*, the same size as the PCR product from *H. polymorpha* (Fig. 4B, lanes 2 vs. 4).

Heterologous expression of the *hph* gene using the *HpACT* promoter

It was noteworthy that the 5′-flanking region of *HpACT* displayed two CCAAT motifs and two TATA-like sequences in a configuration similar to that observed in the *S. cerevisiae* and *Thermomyces* actin genes (Munholland et al. 1990). Two CCAAT motifs were located at positions –264 and –137, respectively, relative to the AUG codon of *HpACT*, while two putative TATA boxes were observed at positions –201 and –80, respectively (Fig. 2). The *ACT* promoters have frequently been exploited to direct heterologous expression of antibiotic resistance genes as a dominant selectable marker in several fungi, based on their relatively invariable expression, regardless of the carbon sources used in the culture medium (Matheucci et al. 1995; Cox et al. 1996). We found that *H. polymorpha* was sensitive to the aminoglycoside antibiotic hygromycin B in concentrations similar to those reported for *S. cerevisiae* (Gritz and Davies 1983) and *Yarrowia lipolytica* (Cordero Otero and Gaillardin 1996). In the initial experiment to express *E. coli hph* (encoding hygromycin B phosphotransferase) as a dominant selectable marker in *H. polymorpha*, the 0.85-kb upstream sequence of *HpACT* containing the translation start site was fused in frame with the 5′ end of *hph* (Fig. 5A). The resultant $P_{HACT850}$-*hph* cassette was associated with the 1.6-kb DNA fragment containing *HLEU2* and TEL188 as a selectable marker and an autonomous replication sequence, respectively, generating an integration vector pHACT850-HyL (Fig. 5B). After transformation of *H. polymorpha* DL-1L (*leu2*) with pHACT850-HyL, transformed cells were allowed to recover for 5 h in liquid YPD and then plated, either onto YPD medium supplemented with 0.25 g hygromycin B/l or onto synthetic minimal medium. Hygromycin B-resistant transformants were obtained at a frequency comparable to that observed for LEU^+ transformants (1×10^3 transformants/μg of DNA vs. 2×10^3 transformants/μg of DNA). The result indicates that the $P_{HACT850}$-*hph* cassette, using the 853-bp *HpACT* promoter, conferred hygromycin B resistance sufficient to allow the direct selection of hygromycin B-resistant transformants in *H. polymorpha*.

Previous studies on a few fungal species suggested that the hygromycin B resistance level of transformants was a measure of the integration copy number of the hy-

Table 2 Comparison of the level of hygromycin B resistance in transformants harboring pHACT-HyL vectors. The number of LEU^+ cells resistant to hygromycin B is the number remaining from 10^6 cells of stabilized LEU^+ transformants plated on 1% yeast extract, 2% peptone, and 2% glucose, containing different concentrations of hygromycin B and incubated for 3 days. pCTEL188 (Sohn et al. 1999b) was used as a control plasmid without the P_{HACT}-*hph* cassette

Plasmid	Number of LEU^+ cells resistant to hygromycin B (hygromycin B given as g/l)						
	0.0	0.25	0.5	1.0	2.0	3.0	4.0
pCTEL188	$1{,}000\times10^3$	–	–	–	–	–	–
pHACT850-HyL	$1{,}000\times10^3$	300×10^3	60×10^3	8×10^3	2×10^3	3×10^2	25
pHACT240-HyL	$1{,}000\times10^3$	250×10^3	70×10^3	10×10^3	3×10^3	2×10^2	17
pHACT150-HyL	$1{,}000\times10^3$	30×10^3	2×10^3	4×10^2	30	2	–
pHACT90-HyL	$1{,}000\times10^3$	20×10^3	1×10^3	3×10^2	25	1	–

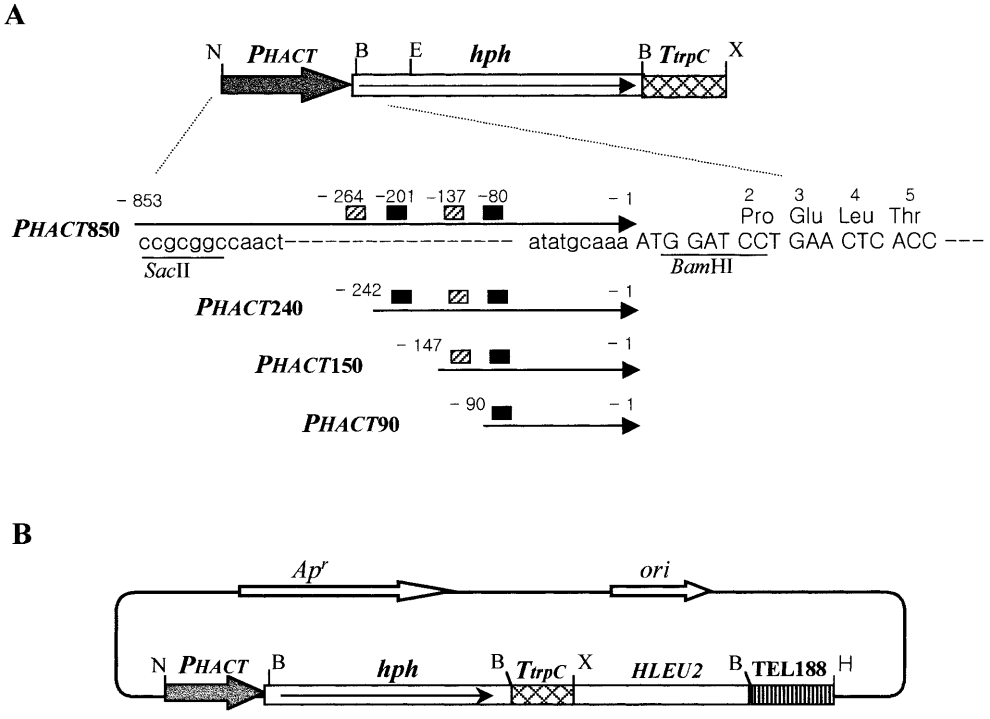


Fig. 5A,B Schematic representation of vector constructs containing P_{HACT}-*hph* cassette for transforming *H. polymorpha*. **A** P_{HACT}-*hph* cassettes under the control of the *HpACT* promoter. A set of deleted *HpACT* promoters containing the initiation codon ATG was fused in frame to the *Bam*HI site at the 5′ end of the bacterial hygromycin B resistance gene (*hph*) derived from pAN7-1 (Punt and van den Hondel 1992). P_{HACT} promoter region of the *H. polymorpha ACT* gene, T_{trpC} terminator region of the *Aspergillus nidulans trpC* gene. Putative CAAT motif and TATA element are represented as a *hatched box* and a *filled box*, respectively. **B** Generic plasmid base containing the P_{HACT}-*hph* cassette. Integration vector pHACT-HyL series were constructed by associating the *Not*I/*Xba*I-digested P_{HACT}-*hph* cassettes with the 1.6-kb *Xba*I/*Hin*dIII DNA fragment containing *HLEU2* and TEL188, derived from pCTEL188 (Sohn et al. 1999b), in the backbone of pBluescript KS^+. *B Bam*HI, *E Eco*RI, *H Hin*dIII, *N Not*I, *X Xba*I

gromycin B resistance cassette, which might provide a way to select for multicopy transformants (Punt and van den Hondel 1992). To check the correlation between resistance level and integration copy number of pHACT850-HyL in *H. polymorpha*, the stabilized LEU^+ transformants were plated onto YPD containing different concentrations of hygromycin B (0–4 g/l). The number of colonies growing on each plate was counted after 48 h of incubation (Table 2). Whereas the cell growth of transformants without the *hph* cassette was completely inhibited by 0.25 g hygromycin B/l, quite a few colonies from the transformants harboring pHACT850-HyL showed up on plates containing up to 4 g hygromycin B/l. For Southern blot analysis, five individual colonies among the transformants of pHACT850-HyL were picked from each plate at concentrations of 0–2 g hygromycin B/l. As shown in Fig. 6A, a single genomic copy of the *LEU2* gene fragment and the transforming *LEU2* gene fragment derived from pHACT850-HyL were detected at ~2.8 kb and ~1.9 kb, respectively. With increasing hygromycin B concentration, the relative intensity of the transforming *LEU2* fragment to that of the genomic *LEU2* fragment became stronger. The result of copy number analysis, based on the Southern densitogram, revealed that the copy number of pHACT850-HyL increased up to ten with the increase in hygromycin B concentration.

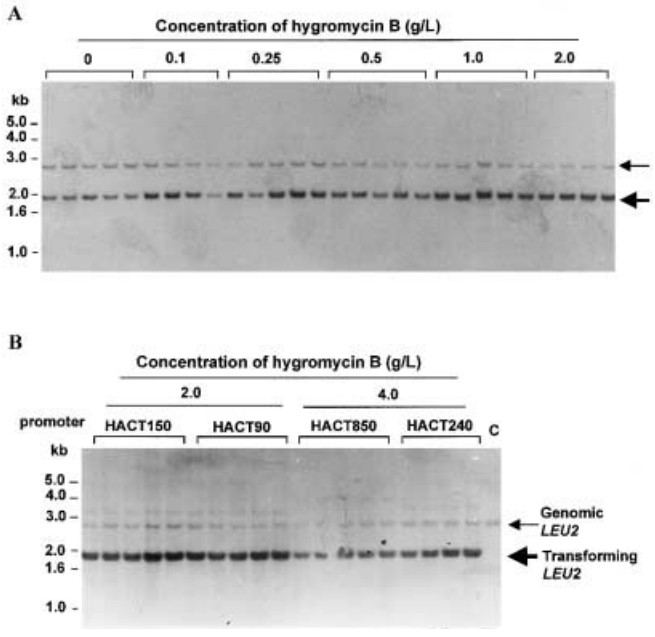


Fig. 6A,B Southern blot analysis of the number of copies integrated in transformants. Several independent colonies from transformation with a series of pHACT-HyL vectors were picked from plates containing different concentrations of hygromycin B. **A** Transformants with pHACT850-HyL. All genomic DNAs were digested with *Bam*HI and probed with the 1.2-kb digoxigenin-labeled *HLEU2* fragment, derived from the *Bam*HI/*Xba*I digestion of pCTEL188 (Sohn et al. 1999b). **B** Transformants with pHACT-HyL vectors containing deleted *HpACT* promoters. *HACT150* and *HACT90*, transformants of pHACT150-HyL and pHACT90-HyL selected at 2 g hygromycin B/l, respectively; *HACT850* and *HACT240*, transformants of pHACT850-HyL and pHACT240-HyL selected at 4 g hygromycin B/l, respectively; *C* untransformed cell. Colorimetric detection of the membrane was carried out by reaction with anti-(digoxigenin alkaline phosphatase; Roche Diagnostics, Mannheim, Germany)

Effect of promoter deletion on hygromycin B resistance

In an effort to increase the selectivity of the hygromycin B resistance marker for transformants of higher copy number, a set of deleted *HpACT* promoters was used to drive the expression of the *hph* gene (Fig. 5). After transformation of *H. polymorpha* DL-1L with pHACT240-HyL, pHACT150-HyL, or pHACT90-HyL, which contain the *hph* cassette under the control of the shortened *HpACT* promoters, the stabilized LEU^+ transformants were plated onto YPD containing different concentrations of hygromycin B (Table 2). Deletion of the *HpACT* promoter from –853 to –242 nucleotides, in which the putative upstream CAAT motif was removed, did not show appreciable influence on the level of hygromycin B resistance. However, further deletion of the promoter down to –147 nucleotides, whichdeleted the upstream TATA-like sequence, resulted in a dramatic decrease in the number of colonies grown on the hygromycin B plates. Transformants harboring pHACT150-HyL could hardly survive on the plates containing 3 g hygromycin B/l. Deletion up to the –90 nucleotide removed the downstream CAAT motif but caused only a negligible decrease in the resistance level. Southern blot analysis of the transformants of pHACT150-HyL and pHACT90-HyL, selected from 2 g hygromycin B/l, revealed that the number of copies integrated was about 25. This was about two-fold higher than the average copy number detected in the transformants of pHACT850-HyL and pHACT240-HyL selected at 4 g hygromycin B/l (Fig. 6B). The results clearly show that the dosage-dependent selection system with these weak resistance cassettes dramatically improved the selection procedure for the multiple integrants of *H. polymorpha*.

Discussion

In the present study, we showed that the *ACT* gene of *H. polymorpha*, the first gene in this organism to be reported to contain an intron, is split by two introns at the 5′-terminal part of the coding region. All yeast actins identified so far are encoded by intron-containing genes. One exception is the fission yeast *Sch. pombe* actin, which is encoded by an intronless gene (Mertins and Gallwitz 1987a). A comparative study on the splice site consensus in the actin genes of various fungi, including yeasts, showed there were significant differences in intron architecture, which reflected the phylogenetic relationships of these organisms (Wery et al. 1996). When compared to the identified actin introns of several yeasts, the *H. polymorpha ACT* introns showed a closer resemblance to those of ascomycetous yeasts including *S. cerevisiae*, *S. bayanus*, *K. lactis*, and *C. albicans* than with those of the basidiomycetous yeast *Pha. rhodozyma* (Table 1). This is in accordance with the earlier classification of *H. polymorpha* within the ascomycetes and with the efficient splicing of *HpACT* introns in *S. cerevisiae*. However, some appreciable differences of intron structure are present between *H. polymorpha* and other ascomycetous yeasts. The 5′-splice sites of the two introns of *H. polymorpha ACT* are GTAAGT, which is one base different from the GTATGT of other ascomycetous yeasts. The putative branch site of the first intron also shows a one-base difference from the TACTAAC of other yeasts. The most noticeable difference is the number and size of introns; in that *H. polymorpha ACT* has two introns, 352 and 87 nucleotides in length, whereas the other ascomycetous yeasts have only one large intron in their *ACT* genes, which ranges in size from 304 bp to 777 bp. The presence of introns is rare in most yeasts and in most cases is restricted to one intron, if any (Woolford 1989). Thus, it will be interesting to see whether the presence of multiple introns is a general feature of intron-containing genes in *H. polymorpha*.

The *ACT* promoters, generally regarded as a constitutive promoter, have been usefully exploited to direct heterologous expression of foreign genes in many organisms. Using the derivatives of the *HpACT* promoter, we have developed a set of P_{HACT}-*hph* cassettes to confer hygromycin B resistance as a dominant selectable marker in the transformation of *H. polymorpha.* The *HpACT* promoter

was found to contain two conserved CCAAT motifs and two TATA-like sequences. Our results indicate that the upstream TATA-like sequence was an important element for the *HpACT* promoter activity, while two CCAAT motifs contributed very little to the promo-ter activity, as previously reported in *S. cerevisiae* (Munholland et al. 1990). The hygromycin B selection system developed in this study should provide a versatile tool for the genetic manipulation of *H. polymorpha*. Moreover, the expression levels of the *E. coli hph* gene from the P_{HACT}-*hph* cassettes were dependent on the integration copy number of the cassettes, rendering the control of integration copy number feasible, simply by using different concentrations of antibiotic. Thus, the dosage-dependent selection system using the P_{HACT}-*hph* cassette would be useful in selecting-for multiple integrants with various copy numbers. Combined with a G418 selection system (Sohn et al. 1999a), the hygromycin B selection system in this study might facilitate the procedure for cointegrating heterologous genes in anticipated fixed copy numbers into the chromosome of *H. polymorpha* to produce different proteins simultaneously in stoichiometric ratios.

Acknowledgement This work has been funded by the Ministry of Science and Technology of Korea (KG1523). The authors deeply appreciate its financial support.

References

Adam L, Laroche A, Barden A, Lemieux G, Pallotta D (1991) An unusual actin-encoding gene in *Physarum polycephalum*. Gene 106:79–86

Bagavathi S, Malathi R (1996) Introns and protein revolution – an analysis of the exon/intron organization of actin genes. FEBS Lett 392:63–65

Barnett JA, Payne RW, Yarrow D (1990) Yeasts: characteristics and identification, 2nd edn. Cambridge University Press, Cambridge

Cordero Otero R, Gaillardin C (1996) Efficient selection of hygromycin B-resistant *Yarrowia lipolytica* transformants. Appl Microbiol Biotechnol 46:143–148

Cox GM, Toffaletti DL, Perfect JR (1996) Dominant selection system for use in *Cryptococcus neoformans*. J Med Vet Microbiol 34:385–391

Deshler JO, Larson GP, Rossi JJ (1989) *Kluyveromyces lactis* maintains *Saccharomyces cerevisiae* intron-encoded splicing single. Mol Cell Biol 9:2208–2213

Dudler R (1990) The single-copy actin gene of *Phytophthora megasperma* encodes a protein considerably diverged from any other known actin. Plant Mol Biol 14:415–422

Elion EA, Warner JR (1984) The major promoter element of rRNA transcription in yeast lies 2 kb upstream. Cell 39:663–673

Gallwitz D, Sures I (1980) Structure of a split yeast gene: complete nucleotide of the actin gene in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 77:2546–2550

Gatzke R, Weydemann U, Janowics ZA, Hollenberg CP (1995) Stable multicopy integration of vector sequences in *Hansenula polymorpha*. Appl Microbiol Biotechnol 43:844–849

Gellissen G, Piontek M, Dahlems U, Jenzelewski V, Gavagan JE, DiCosimo R, Anton DL, Janowicz ZA (1996) Recombinant *Hansenula polymorpha* as a biocatalyst: coexpression of the spinach glycolate oxidase (GO) and the *S. cerevisiae* catalase T (*CTT1*) gene. Appl Microbiol Biotechnol 46:46–54

Gritz L, Davies J (1983) Plasmid-encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. Gene 25:179–188

Harris DE, Warshaw DM, Periasamy M (1992) Nucleotide sequences of the rabbit alpha-smooth-muscle and beta-non-muscle actin mRNA. Gene 112:265–266

Hill JE, Myers AM, Koerner TJ, Tzagoloff A (1986) Yeast/*E. coli* shuttle vectors with multiple unique restriction sites. Yeast 2:163–167

Hollenberg CP, Gellissen G (1997) Production of recombinant proteins by methylotrophic yeasts. Curr Opin Biotechnol 7:525–530

Holm C, Meeks-Wagner DW, Fangman LW, Botstein D (1986) A rapid efficient method for isolating DNA from yeast. Gene 42:169–173

Janowicz ZA, Melber K, Merckelbach A, Jacobs E, Harford N, Comberbach M, Hollenberg CP (1991) Simultaneous expression of the S and L surface antigens of hepatitis B and formation of mixed particles in the methylotrophic yeast, *Hansenula polymorpha*. Yeast 7:431–443

Kim E, Waters SH, Hake LE, Hecht NB (1989) Identification and developmental expression of a smooth-muscle gamma-actin in postmeiotic male germ cells of mice. Mol Cell Biol 9: 1875–1881

Losberger C, Ernst JF (1989) Sequence of the *Candida albicans* gene encoding actin. Nucleic Acids Res 17:9488

Matheucci EJ, Henrique-Silva F, El-Gogary S, Rossini CHB, Leite A, Escobar Vera J, Carle Urioste JC, Crivellaro O, El-Dory H (1995) Structure, organization and promoter expression of the actin-encoding gene in *Trichoderma reesei*. Gene 161:103–106

Mertins P, Gallwitz D (1987a) A singly intronless actin gene in the fission yeast *Schizosaccharomyces pombe*: nucleotide sequence and transcripts formed in homologous and heterologous yeast. Nucleic Acids Res 15:7369–7379

Mertins P, Gallwitz D (1987b) Nuclear pre-mRNA splicing in the fission yeast *Schizosaccharomyces pombe* strictly requires an intron-contained, conserved sequence element. EMBO J 6: 1757–1763

Munholland J, Kelly JK, Wildeman AG (1990) DNA sequences required for yeast actin gene transcription do not include conserved CCAAT motifs. Nucleic Acids Res 18:6061–6068

Pollard TD (1990) Actin. Curr Opin Cell Biol 2:33–40

Punt PJ, Hondel CAMJJ van den (1992) Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. Methods Enzymol 216:447–457

Rambosek J, Leach J (1987) Recombinant DNA in filamentous fungi: progress and prospects. Crit Rev Biotechnol 6:357–393

Sohn JH, Choi ES, Kang HA, Rhee JS, Agaphonov MO, Ter-Avanesyan MD, Rhee SK (1999a) A dominant selection system designed for copy-number-controlled gene integration in *Hansenula polymorpha* DL-1. Appl Microbiol Biotechnol 51:800–807

Sohn JH, Choi ES, Kang HA, Rhee JS, Rhee SK (1999b) A family of telomere-associated autonomous replicating sequences and their functions in targeted recombination in *Hansenula polymorpha* DL-1. J Bacteriol 181:1005–1013

Taylor A, Erba HP, Muscat GE, Kedes L (1988) Nucleotide sequence and expression of the human skeletal alpha-actin gene: evolution of functional regulatory domains. Genomics 3:323–336

Titorenko VI, Waterham HR, Cregg JM, Harder W, Veenhuis M (1993) Peroxisome biogenesis in the yeast *Hansenula polymorpha* is controlled by a complex set of interacting gene products. Proc Natl Acad Sci USA 90:7470–7474

Vandekerckhove J, Weber K (1978a) At least six different actins are expressed in a higher mammal: an analysis based on the amino acid sequence of the amino-terminal tryptic peptide. J Mol Biol 126:783–802

Vandekerckhove J, Weber K (1978b) The amino acid sequence of *Physarum* actin. Nature 276:720–721

Wery J, Dalderup MJM, Linde JT, Boekhout T, Ooyen AJ van (1996) Structural and phylogenetic analysis of the actin gene from the yeast *Phaffia rhodozyma*. Yeast 12:641–651

Woolford JL (1989) Nuclear pre-mRNA splicing in yeast. Yeast 5:439–457