JOURNAL OF BACTERIOLOGY, Sept. 1993, p. 5548–5558
0021-9193/93/175548-11$02.00/0

Vol. 175, No. 17

Cloning, Primary Structure, and Regulation of the *HIS7* Gene Encoding a Bifunctional Glutamine Amidotransferase:Cyclase from *Saccharomyces cerevisiae*

MARKUS KUENZLER, TIZIANO BALMELLI, CHRISTOPH M. EGLI, GERHARD PARAVICINI,† AND GERHARD H. BRAUS*

Mikrobiologisches Institut, Eidgenössische Technische Hochschule, CH-8092 Zürich, Switzerland

Received 12 April 1993/Accepted 30 June 1993

The *Saccharomyces cerevisiae HIS7* gene was cloned by its location immediately downstream of the previously isolated and characterized *ARO4* gene. The two genes have the same orientation with a distance of only 416 bp between the two open reading frames. The yeast *HIS7* gene represents the first isolated eukaryotic gene encoding the enzymatic activities which catalyze the fifth and sixth step in histidine biosynthesis. The open reading frame of the *HIS7* gene has a length of 1,656 bp resulting in a gene product of 552 amino acids with a calculated molecular weight of 61,082. Two findings implicate a bifunctional nature of the *HIS7* gene product. First, the N-terminal and C-terminal segments of the deduced HIS7 amino acid sequence show significant homology to prokaryotic monofunctional glutamine amidotransferases and cyclases, respectively, involved in histidine biosynthesis. Second, the yeast *HIS7* gene is able to suppress His auxotrophy of corresponding *Escherichia coli hisH* and *hisF* mutants. *HIS7* gene expression is regulated by the general control system of amino acid biosynthesis. GCN4-dependent and GCN4-independent (basal) transcription use different initiator elements in the *HIS7* promoter.

Histidine is synthesized in an invariable series of 11 enzymatic reactions from ATP and phosphoribosyl-pyrophosphate (PRPP) in all histidine-autotrophic organisms studied so far (Fig. 1A). Enzymatic regulation of the unbranched pathway is achieved by feedback inhibition of the first step of the pathway by its final product, histidine, in both prokaryotic and eukaryotic microorganisms studied (9, 26). This biochemical invariability faces a considerable diversity in organization, structure, and regulation of the genes coding for the various enzymatic activities even within a biological kingdom.

In the best-studied organisms, *Escherichia coli* and *Salmonella typhimurium*, the 11 enzymatic activities are encoded by eight genes organized in a single operon [*hisGD CBHAF(IE)*] (Fig. 1) (11, 54). Three of the eight genes code for bifunctional enzymes [*hisD*, *hisB*, and *his(IE)*]. This situation, however, is not typical of prokaryotes and not even of eubacteria. In *Lactococcus lactis*, nine *his* genes exist, of which eight are clustered in an operon and one is located elsewhere on the chromosome (14). In *Streptomyces coelicolor*, eight genes map at three loci, one of them grouping six genes in an operon and the other two containing single genes (25, 30), whereas in *Bacillus subtilis* the eight genes are organized in two loci with seven genes and one gene, respectively (44). In the nitrogen-fixing eubacterium *Azospirillum brasilense*, at least four *his* genes are clustered (5, 17), and in *Staphylococcus aureus*, six *his* genes are clustered (42). In some of these organisms, not only the organization but also the structure of individual genes differs from the situation in enteric eubacteria. The two enzymatic activities encoded by the *hisB* gene in *E. coli* and *S. typhimurium* reside on separate genes in *L. lactis*, *S. coelicolor*, and *A. brasilense*. In addition, in *S. coelicolor* the *his(IE)* activities are separated as well. Physically separated *his(IE)* activities are also found in the methanogenic archaebacterium *Methanococcus vannielii* (6).

In fungi, the genes encoding the enzymatic activities of various biosynthetic pathways are scattered throughout the genome. In *S. cerevisiae*, the genetic information for the histidine biosynthetic enzymes is encoded by seven genes, which are located on six different chromosomes (*HIS1-7*) (Fig. 1) (10). The structure of the genes differs from that of enterobacteria in that the enzymatic activities of *hisD* and *his(IE)* are combined to a multifunctional enzyme catalyzing four steps in the histidine biosynthetic pathway encoded by the *HIS4* gene (16). Such a multifunctional enzyme exists as well in *Neurospora crassa* (29) and *Candida albicans* (2). In addition, as in *L. lactis*, *A. brasilense*, and *S. coelicolor* the *hisB* activities of *E. coli* and *S. typhimurium* are encoded by two independent genes in *S. cerevisiae* (*HIS2* and *HIS3*) (45, 50). The only plant gene cloned so far is a cDNA from the cabbage *Brassica oleracea* corresponding to the *hisD* gene of enterobacteria and encoding a bifunctional histidinol dehydrogenase (36), indicating that the organization and structure of the genes involved in histidine biosynthesis are also variable within the eukaryotic kingdom.

Coordinate regulation of the histidine-biosynthetic genes strongly depends on the gene organization in the corresponding organism. In *E. coli* and *S. typhimurium*, where all *his* genes are clustered in a single operon, coregulation is achieved by attenuation control and positive metabolic regulation of the operon (54). In *S. cerevisiae*, the scattered genes are part of a complex regulatory network which couples the transcriptional derepression of at least 30 structural genes involved in multiple-amino-acid biosynthetic pathways under environmental conditions of amino acid starvation (23). The final step in this general control system is the binding of the transcriptional activator protein GCN4

* Corresponding author. Electronic mail address: BRAUS@aeolus.ethz.ch.

† Present address: GLAXO Institute for Molecular Biology, Chemin des Aulx, CH-1228 Plan-les-Ouates, Geneva, Switzerland.

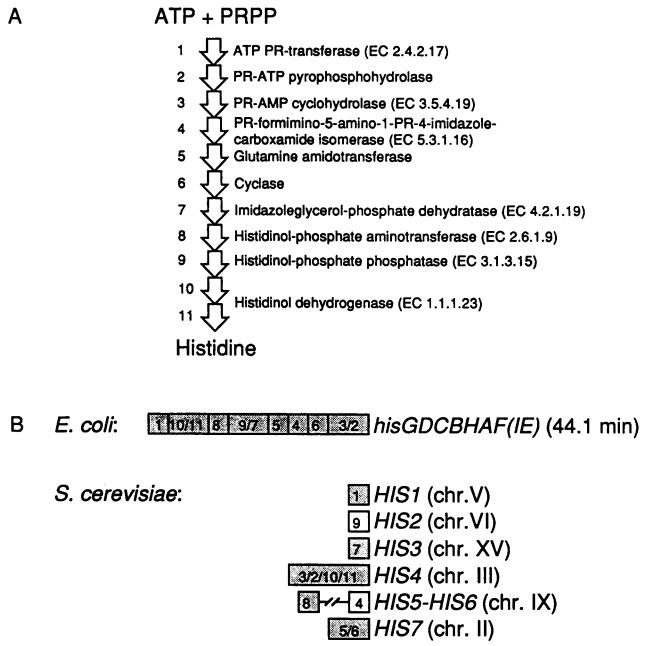


FIG. 1. Gene-enzyme relationships in histidine biosynthesis in *E. coli* and *S. cerevisiae*. (A) Schematic representation of the biosynthetic steps from ATP and PRPP to histidine with enzyme designations and assigned Enzyme Commission (EC) numbers. (B) Organization of the genetic information for the histidine biosynthetic enzymes in *E. coli* and *S. cerevisiae*. Genes are represented by boxes which are shaded for genes that are cloned and sequenced. Encoded enzymatic activities are indicated by numbers referring to panel A. Multiple numbers separated by slashes symbolize multifunctional enzymes. The chromosomal locations are given as map positions for *E. coli* (min) and as chromosome numbers for *S. cerevisiae* (chr.). The *S. cerevisiae HIS5* and *HIS6* genes are located on the same arm of chromosome IX.

to the promoters of the target genes resulting in elevated transcription of these genes.

In this paper, we describe the cloning and characterization of the *HIS7* gene, which codes for the enzyme catalyzing the fifth and sixth step of the histidine biosynthetic pathway in *S. cerevisiae*. The derived amino acid sequence and heterologous complementation of corresponding *E. coli* mutants suggests that the *HIS7* gene product is a bifunctional enzyme with an N-terminal glutamine amidotransferase and a C-terminal cyclase domain. Furthermore, we find that transcription of the *HIS7* gene is regulated by the general control system of amino acid biosynthesis and that the start sites for GCN4-dependent and GCN4-independent (basal) transcription are different.

MATERIALS AND METHODS

Strains and culture conditions. All strains used in this study are listed in Table 1. Yeast strains were all derivatives of the *S. cerevisiae* laboratory strain S288C (*MATα gal2 SUC2 mal CUP1*).

Cultivation of *S. cerevisiae* was performed at 30°C in either YEPD complete medium (46) or MV minimal medium (33). Appropriate supplements were added to the medium in recommended amounts (46). LB complete medium and M9 minimal medium for *E. coli* are described by Sambrook et al. (48). LB medium containing ampicillin (50 mg/liter) was used to select for transformants. *E. coli* was cultivated at 37°C.

Crossing of *S. cerevisiae*. Crossing of compatible yeast strains was performed as described previously (46). Selection for diploids after mating was done on MV minimal medium.

DNA techniques and sequencing. Enzymatic manipulation and cloning of DNA were performed as described by Sambrook et al. (48). *E. coli* MC1061 (12) was used for the propagation of plasmid DNA. DNA sequences were determined for both strands by the chain termination method (49) and the M13 subcloning technique (31). Oligonucleotide primers were purchased from Microsynth (Windisch, Switzerland). The M13 host JM101 (31) and the M13-based vectors M13mp18 and M13mp19 (56) were used for the isolation of single-stranded template DNA.

PCR. The polymerase chain reaction (PCR) technique for the amplification of cloned DNA fragments by using sequence-specific oligonucleotides was described previously (47). In this work, the technique was exploited for the production of a DNA fragment used for S1 nuclease mapping of the *HIS7* mRNA 3′ end. As standard reactions using Super *Taq* polymerase (P. H. Stehelin & Cie AG, Basal, Switzerland), 30 cycles of 1 min at 94°C, 1 min at 50°C, and 1 min at 72°C were performed in a Biometra Trioblock thermocycler (Biometra, Göttingen, Germany).

Yeast genomic DNA library. The yeast genomic DNA library contains DNA of strain YPH1 (*MATa ura3-52 lys2-801 ade2-101 GAL⁺ SUC⁺*) partially digested with *Sau*3AI in a YCp50 derivative in which the yeast *URA3* gene is replaced by the yeast *LEU2* gene.

Construction of the Δ(*aro4-his7*)::*URA3* disruption strain. *S. cerevisiae* RH1447 carrying a disrupted *ARO4-HIS7* locus was constructed as follows. In the course of subcloning of the *ARO4* gene, a chromosomal 3.5-kb *Xba*I-*Bam*HI fragment ranging from an *Xba*I site located approximately 0.9 kb upstream of the 5′ end of the region shown in Fig. 2A to the indicated *Bam*HI site was cloned into pGEM7Zf(+) (Promega, Madison, Wis.), yielding plasmid pME638. From this plasmid, a 2.4-kb *Acc*I fragment comprising the complete *ARO4* gene and the 5′ end of the *HIS7* gene (see Fig. 2A) was isolated and replaced by the chromosomal 1.1-kb *URA3* fragment in the same orientation as the substituted genes, resulting in plasmid pME642. Transformation of *S. cerevisiae* RH1377 with the 2.2-kb *Xba*I-*Bam*HI fragment from plasmid pME642 and selection for a Ura⁺ phenotype in the presence of supplementing amounts of histidine resulted in strain RH1447. The strain was examined for its His⁻ and concomitant Aro⁻ phenotypes in the presence of 5 mM phenylalanine and by Southern blot analysis.

Construction of strains with an integrated translational *HIS7-lacZ* fusion. The translational *HIS7-lacZ* fusion was constructed based on plasmid pNM482 (32). A 0.6-kb *Hpa*I-*Acc*I fragment containing the complete *ARO4-HIS7* intergenic region and the N-terminal 56 amino acids of the *HIS7* open reading frame (see Fig. 2A and 3) was inserted into pNM482 restricted with *Sma*I and *Acc*I. From this plasmid the *HIS7-lacZ* fusion gene was isolated as a 4-kb *Eco*RI-*Csp*45I fragment and inserted into vector pGEM7Zf(+) containing the 0.5-kb *Hin*dIII-*Bam*HI 3′ end of the yeast *ADH1* gene (7) to yield plasmid pME688. Therefore, this plasmid contains a 4.5-kb *Eco*RI-*Bam*HI *HIS7-lacZ* translational fusion cassette.

For the construction of an integrative *HIS7-lacZ* fusion cassette, the 1.9-kb *Bam*HI-*Hin*dIII *HIS7* fragment (same fragment as *Bam*HI fragment from pME692) was cloned as a

TABLE 1. Strains and plasmids

Species and strain or plasmid	Relevant characteristics	Reference or source
E. coli		
MC1061	Δ(*lacIPOZYA*)*X74 galU galK* StrA^r *hsdR* Δ(*ara-leu*)	12
JM101	Δ(*lac-pro*) *thi supE* F′(*traD36 proAB lacI^{q}ZΔM15*)	31
W3110	Wild type	*E. coli* Genetic Stock Center[a]
UTH6	λ$^-$ *hisA323*	*E. coli* Genetic Stock Center
UTH860	*ara-14 glnV44 galK2* λ$^-$ *rpsL 145 malT1*(λ^r) *xylA5 mtl-1 hisF860* λ^r	*E. coli* Genetic Stock Center
UTH1767	*malA1*(λ^r) *xyl-5 mtl-1 rpsL145 hisH1767* λ$^-$	*E. coli* Genetic Stock Center
S. cerevisiae		
RH1377	*MAT*α Δ*ura3*	ETH collection[b]
RH1447	*MAT*α Δ*ura3* Δ(*aro4-his7*)::*URA3*	This work
C20-2C	*MAT***a** *his7 ade2 ade4 ura1 lys2 tyr1 arg4 leu1 trp5 gal*	Yeast Genetic Stock Center[c]
RH1631	*MAT***a** *ura3-52*	ETH collection
RH1632	*MAT*α *ura3-52 gcd2-1*	ETH collection
F194	*MAT***a** *ura3-52 gal2 gcn4-103*	22
RH1371	*MAT***a** *aro3-2* Δ*ura3 gcd2-1*	ETH collection
RH1372	*MAT***a** *aro3-2* Δ*ura3*	ETH collection
RH1381	*MAT*α *aro3-2 ura3-52 gcn4-101*	ETH collection
RH1614	*MAT***a** *aro3-2* Δ*ura3* Δ*his7*::*lacZ gcd2-1*	This work
RH1615	*MAT***a** *aro3-2* Δ*ura3* Δ*his7*::*lacZ*	This work
RH1616	*MAT*α *aro3-2 ura3-52* Δ*his7*::*lacZ gcn4-101*	This work
Plasmids		
pME638	pGEM7Zf(+)[d] containing a 3.5-kb *Xba*I-*Bam*HI *ARO4-HIS7* fragment	ETH collection
pME642	pME638 with a Δ(*aro4-his7*)::*URA3* disruption	This work
pME688	pGEM7Zf(+) containing a 4.5-kb *Eco*RI-*Bam*HI *HIS7-lacZ* cassette	This work
pME692	pGEM7Zf(+) containing a 1.9-kb *Bam*HI-*Hin*dIII *HIS7* fragment	This work
pME693	pME688 with a 1.9-kb *Bam*HI *HIS7* fragment from pME692	This work
pME694	pGEM7Zf(+) containing a 1.9-kb *Sph*I-*Bam*HI *ARO4-HIS7* fragment	This work
pME696	pME694 with a 6.1-kb *Bsm*I-*Nsi*I Δ*his7*::*lacZ* fragment from pME693	This work
pME979	pGEM7Zf(+) containing a 2.4-kb *Eco*RV-*Hin*dIII *HIS7* fragment	This work

[a] Department of Biology, Yale University, New Haven, Conn.
[b] ETH, Eidgenössische Technische Hochschule.
[c] Department of Biophysics and Medical Physics, University of California, Berkeley.
[d] Promega, Madison, Wis.

homologous downstream region behind the *ADH1* 3′ end region in pME688, resulting in plasmid pME693. To ensure proper integrative recombination of the fusion, its homologous upstream region was subsequently enlarged to the *Sph*I site by ligating a 6.1-kb *Bsm*I-*Nsi*I fragment from the latter plasmid into plasmid pME694 containing the 1.9-kb *Sph*I-*Bam*HI fragment of the *ARO4-HIS7* locus (see Fig. 2A) on pGEM7Zf(+) to yield plasmid pME696.

A two-step procedure was used for the integration of the translational *HIS7-lacZ* fusion at the original *HIS7* locus on the yeast chromosome, resulting in a Δ*his7*::*lacZ* genotype. In the first step, the 1.1-kb *Kpn*I-*Acc*I fragment of the chromosome was replaced by the *URA3* gene in strains RH1371, RH1372, and RH1381 as described for the construction of RH1447. The resulting strains had both His$^-$ and an Aro$^-$ phenotypes, with the latter due to an *aro3-2* mutation. Transformation of the disruption strains with plasmid pME696 restricted with *Xba*I and selection for an Aro$^+$ phenotype in the presence of supplementing amounts of histidine and uracil resulted in strains RH1614, RH1615, and RH1616 carrying a translational *HIS7-lacZ* fusion instead of the original *HIS7* locus and an intact *ARO4* gene. Strains were examined for Ura$^-$ and His$^-$ phenotypes and by Southern blot analysis.

Poly(A)$^+$ RNA isolation. Yeast RNA enriched for polyadenylated RNA species was isolated as described previously (20). Isolated RNA was stored in 40% (vol/vol) isopropanol–120 mM sodium acetate at −20°C.

Primer extension analysis. The primer extension method to determine RNA 5′ ends was performed as described by Kassavetis and Geiduschek (27). For the mapping of the *HIS7* transcript 5′ ends, 50 μg of poly(A)$^+$ RNA of each strain was hybridized against an excess of a 5′-^{32}P-end-labelled 51-bp primer complementary to nucleotide positions +14 to +64 relative to the *HIS7* translational start site. Annealed primers were elongated with avian myeloblastosis virus reverse transcriptase. Elongation products were separated on a 6% polyacrylamide standard sequencing gel together with a T ladder generated by using the same primer as for the primer extension reactions.

S1 nuclease mapping. The S1 nuclease protection method for mapping RNA 5′ and 3′ ends was performed as described by Furter et al. (20). For *HIS7* mRNA 5′ end mappings, 30 μg of poly(A)$^+$ RNA of each strain was hybridized against an excess of a *Kpn*I-*Bam*HI fragment (ranging from nucleotide positions −769 to +415) which was ^{32}P labelled at the 3′ end of the antisense strand. The resulting hybrid molecules were digested with S1 nuclease, and the protected DNA strands were separated on a 6% polyacrylamide standard sequencing gel. As a size standard, ^{32}P-labelled pBR322 plasmid DNA restricted with *Hpa*II was used. Mapping of the *HIS7* transcript 3′ ends employed a fragment ranging from nucleotide positions +1503 to +2145 and ^{32}P labelled at the 5′ end of the antisense strand. This fragment was generated by a standard PCR reaction using two primers, ranging from positions +1467 to +1486 (20 bp) and comple-

mentary to positions +2116 to +2145 (30 bp) respectively, followed by subsequent cleavage with *Nar*I and filling in of the 5′ protruding end with [α-^{32}P]dCTP. Nuclease S1 digestion products were separated on a 6% polyacrylamide standard sequencing gel together with a sequence ladder generated by using the PCR primer complementary to nucleotide positions +2116 to +2145. Torula yeast RNA (30 μg) was used as a negative control.

Northern (RNA) analysis. Poly(A)$^+$ RNA (30 μg) of each strain was separated on a formaldehyde agarose gel, electroblotted onto a nylon membrane, and hybridized against DNA fragments which were ^{32}P labelled by using the oligolabelling technique described by Feinberg and Vogelstein (18). Probes used were made from a chromosomal 1.1-kb *Hin*dIII *URA3* fragment, a chromosomal 0.7-kb *Hpa*I *ARO4* fragment, and a chromosomal 0.9-kb *Bam*HI-*Xba*I *HIS7* fragment. The *URA3* transcript was chosen as an internal standard for the amount of RNA, as this gene is not regulated by the GCN4 protein.

β-Galactosidase assay. β-Galactosidase activities were determined by using permeabilized yeast cells and the fluorogenic substrate 4-methylumbelliferyl-β-D-galactoside (Fluka Chemie AG, Buchs, Switzerland). Yeast was cultivated in 5 ml of MV minimal medium supplemented with histidine and uracil to an optical density at 546 nm of between 1 and 4. Typically, cells from 0.5 ml of yeast culture were washed once with water and resuspended in 1 ml of reaction buffer (25 mM Tris-HCl [pH 7.5], 125 mM NaCl, 2 mM $MgCl_2$, 12 mM 2-mercaptoethanol). The cells were permeabilized by vortexing for 10 s after the addition of 50 μl of CH_2Cl_2 and 0.1% (wt/vol) sodium dodecyl sulfate. Then, 40 μl of permeabilized cells was incubated with 160 μl of reaction buffer containing 0.3 mM 4-methylumbelliferyl-β-D-galactoside for 30 min at 37°C. The reaction was stopped after 30 min by adding 50 μl of 25% (wt/vol) trichloroacetic acid. The cells were spun down, and the fluorescence of the supernatant was determined in an at least 1/4 dilution in glycine/carbonate reagent (133 mM glycine, 83 mM Na_2CO_3) with a Hoefer model TKO 100 fluorometer (Hoefer Scientific Instruments, San Francisco, Calif.). The concentration of product formed during the reaction was determined based on a standard curve in a range from 0 to 40 μM 4-methylumbelliferone (MUF) in reaction buffer. Product concentrations were normalized to the reaction time and the optical density of the culture. One unit of β-galactosidase activity is defined as 1 nmol of MUF h^{-1} ml^{-1} optical density at 546 nm^{-1}. The given values are means of at least three independent cultures. The standard error of the mean was less than 25%.

Sequence data analysis. Sequence data were analyzed with the Genetics Computer Group Sequence Analysis Software (15). Multisequence alignments were produced by the program PILEUP, and pairwise alignments and identity or similarity value calculations were done by using the program GAP.

Nucleotide sequence accession numbers. The nucleotide sequence presented in this paper has been assigned GenBank/EMBL accession numbers X61107 (*ARO4*) and X69815 (*HIS7*).

RESULTS

Cloning and sequencing of the *HIS7* gene. The *ARO4* gene of *S. cerevisiae*, encoding the tyrosine-inhibitable desoxyarabino-heptulosonate-phosphate synthase (DAHPS), was previously isolated and assigned to chromosome II (28). A disruption of *ARO4* replacing a chromosomal 2.4-kb *Acc*I

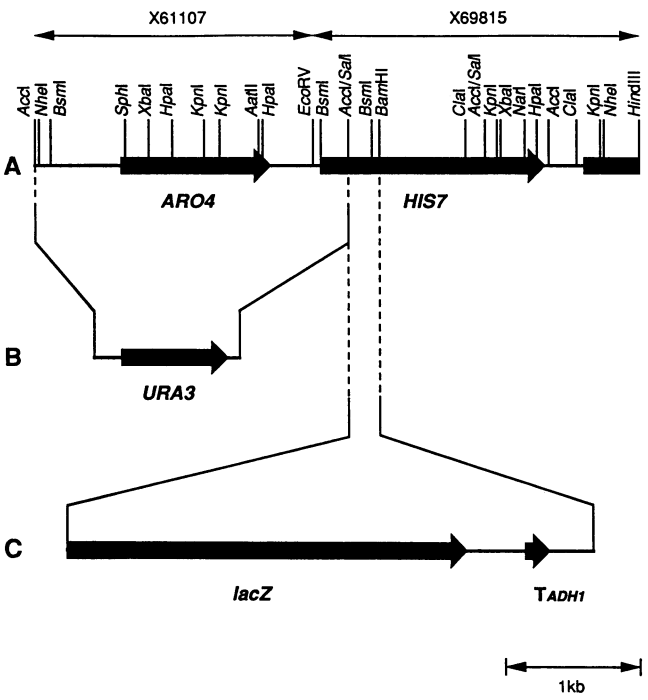


FIG. 2. Restriction map of the *ARO4-HIS7* locus in different *S. cerevisiae* strains. Assigned GenBank/EMBL sequence accession numbers for the two genes are indicated above. (A) Wild-type situation. (B) Disrupted locus in strain RH1447. (C) Translational fusion of the *HIS7* gene to the *E. coli lacZ* gene with the 3′ end of the yeast *ADH1* gene (T_{ADH1}) integrated at the original chromosomal locus in strains RH1614, RH1615, and RH1616.

fragment by the yeast *URA3* gene (Fig. 2B) resulted in an additional His auxotrophy. The fact that the *HIS7* gene is located on the same chromosome as the *ARO4* gene (34) led us to the presumption that a concomitant disruption of the *HIS7* gene could have caused the additional phenotype. To test this, we introduced the above-described gene disruption in *MAT*α strain RH1377 to yield RH1447 (Fig. 2B), and crossed both wild-type and disruption strain with a *MAT***a** *his7* mutant (strain C20-2C) from the Yeast Genetic Stock Center. Wild-type strain RH1377, but not disruption strain RH1447, was able to complement the *his7* mutation of strain C20-2C. This indicated that the *HIS7* gene was located adjacent to the *ARO4* gene and codisrupted with the *ARO4* gene. The relative location of the genes was determined by Northern analysis. A 1.8-kb poly(A)$^+$ RNA, which was not present in a disruption strain, could be detected by using the 0.5-kb *Eco*RV-*Bam*HI fragment shown in Fig. 2A as a probe (data not shown). The *HIS7* gene was therefore located immediately downstream of the *ARO4* gene. Additional evidence for the location of *HIS7* adjacent to *ARO4* was obtained by cloning of the complete *HIS7* gene by functional complementation of disruption strain RH1447 with a yeast genomic DNA library on a yeast centromeric plasmid. One of the transformants contained a 10.5-kb insert of yeast DNA. Subcloning localized the complementing activity on a chromosomal 6.2-kb *Hin*dIII fragment. This fragment was able to confer growth to RH1447 in the presence of 5 mM phenylalanine in the medium. Under these conditions the isoenzyme of the *ARO4* gene product, the phenylalanine-inhibitable, *ARO3*-encoded DAHPS, is fully inhibited and growth depends on an intact *ARO4* gene. This indicated that

the 6.2-kb *Hin*dIII fragment contained both the *ARO4* and the *HIS7* gene.

A comparison of the restriction map of the fragment with the one of the *ARO4* locus revealed a 2.4-kb *Eco*RV-*Hin*dIII fragment of chromosomal DNA located immediately downstream of the previously sequenced *ARO4* *Acc*I-*Eco*RV fragment (Fig. 2A) (28). The nucleotide sequence of the *Eco*RV-*Hin*dIII fragment contained a single open reading frame of 1,656 bp (Fig. 3). The *HIS7* gene product is thus predicted to consist of 552 amino acids with a calculated molecular weight of 61,082.

Bifunctional nature of the *HIS7* gene product. The *HIS7* gene was previously assigned to both the fifth and sixth step in histidine biosynthesis in *S. cerevisiae*, converting phosphoribosyl-formimino-5-amino-1-phosphoribosyl-4-imidazole-carboxamide to imidazoleglycerol-phosphate (19, 26). This assignment was based on the analysis of accumulation products caused by mutational blocks in the biosynthetic pathway. In prokaryotic microorganisms, by contrast, these reactions are carried out by two monofunctional enzymes, a glutamine amidotransferase (fifth step) and a cyclase (sixth step). In *E. coli* and *S. typhimurium* these enzymes are encoded by two different genes, the *hisH* gene (glutamine amidotransferase) and the *hisF* gene (cyclase). In both organisms the two genes are part of a single histidine operon and separated by the *hisA* gene, which codes for phosphoribosyl-formimino-5-amino-1-phosphoribosyl-4-imidazole-carboxamide catalyzing the fourth step of the histidine biosynthetic pathway (Fig. 1).

The deduced amino acid sequence for the *HIS7* gene was aligned to the prokaryotic glutamine amidotransferase and cyclase sequences currently available in the GenBank/EMBL data base (Fig. 4). Alignments revealed significant homology of the N-terminal segment of the *HIS7*-derived amino acid sequence (amino acids 1 to 213) with the various *hisH*-derived amino acid sequences and of the C-terminal segment (amino acids 235 to 552) with the various *hisF*-derived amino acid sequences. Thus the primary structure of the *HIS7* gene is consistent with its product being a bifunctional enzyme with a N-terminal glutamine amidotransferase and a C-terminal cyclase domain.

Many yeast genes like *HIS2* (45) and *HIS3* (50) have been functionally expressed in *E. coli*. In order to confirm the dual function of the *HIS7* gene product on a functional level, we tested the *HIS7* gene for its ability to complement different *E. coli his* mutants. Therefore we transformed *E. coli* K-12-derived strains W3110 (wild-type strain), UTH6 (*hisA* mutant), UTH860 (*hisF* mutant), and UTH1767 (*hisH* mutant) with plasmid pME979 carrying the 2.4-kb *Eco*RV-*Hin*dIII *HIS7* fragment on vector pGEM7Zf(+) (Promega, Madison, Wis.) and the empty vector as a negative control. Selection for transformants was done on LB complete medium containing ampicillin. The transformants were tested for the His phenotype by streaking them on M9 minimal glucose agar and incubating the plates for 3 days at 37°C. The untransformed strains were plated on M9 minimal glucose agar and M9 minimal glucose agar supplemented with 20 mg of histidine per liter as a control. Plasmid pME979 containing the yeast *HIS7* gene was able to suppress His auxotrophy of *E. coli* K-12 *hisH* and *hisF* mutants, whereas the *hisA* mutant could not be complemented (Fig. 5). Thus, the yeast *HIS7* gene product can functionally replace both the *hisH*-encoded glutamine amidotransferase and the *hisF*-encoded cyclase activity in *E. coli*.

In summary, both structural and functional findings reveal that the *S. cerevisiae* *HIS7* gene encodes a bifunctional enzyme homologous to prokaryotic *hisH* and *hisF* gene products.

***HIS7* regulation by the general control system of amino acid biosynthesis.** In *S. cerevisiae* many genes involved in amino acid biosynthetic pathways are coordinately regulated by the general control system of amino acid biosynthesis. As the final step of a regulatory cascade under the environmental conditions of amino acid starvation, this system activates transcription of target genes by binding of the protein GCN4 to distinct recognition elements in the promoters of the corresponding genes (23). For five histidine-biosynthetic gene products (*HIS1* to *HIS5*), a regulation by this system was demonstrated on the enzymatic level (26). After the isolation of the corresponding genes the regulation was confirmed on the transcriptional level for *HIS1*, *HIS3*, *HIS4*, and *HIS5* (16, 24, 39, 51). The *HIS7* gene could not be tested for a regulation by the general control system because of the lack of a convenient enzyme assay for the *HIS7* gene product. The isolation of the *HIS7* gene enabled us to perform an analysis of *HIS7* transcription regulation. Two independent assays were used to demonstrate regulation of *HIS7* transcription by the general control system of amino acid biosynthesis. Both assays made use of regulatory mutants in the general control system and a wild-type strain as a control. Strains carrying a *gcd2-1* mutation express GCN4 at a constitutively high level and therefore mimic the situation of amino acid starvation (38), whereas in *gcn4-103* and *gcn4-101* mutants no functional GCN4 is present (22). Wild-type strains exhibit intermediate GCN4 levels (22). In the first method, *HIS7* mRNA levels were determined relative to those of *ARO4* and *URA3* in *S. cerevisiae* RH1632 (*gcd2-1*), RH1631 (wild type), and F194 (*gcn4-103*) by Northern analysis (Fig. 6A). In a second approach, β-galactosidase activities of strains RH1614 (*gcd2-1*), RH1615 (wild type), and RH1616 (*gcn4-101*) carrying a translational *HIS7-lacZ* fusion integrated at the *HIS7* locus were measured (Fig. 6B). The experiments revealed an up to sixfold derepression of the *HIS7* gene in a *gcd2-1* mutant compared with that in a *gcn4* mutant. Thus, the *HIS7* gene is regulated by GCN4.

Transcript 5′ end mapping of GCN4-regulated genes under both repressing and derepressing conditions has revealed two types of transcription initiation patterns. In the *HIS3* (44) and the *TRP4* (20, 35) genes, the start sites for GCN4-dependent transcription differ from those for GCN4-independent (basal) transcription, whereas other genes like *HIS1* (24), *HIS4* (16), *HIS5* (39), *ARO3* (41), and *ARO4* (28) show the same pattern under both conditions. For the *HIS7* gene the situation was analyzed by mapping the mRNA 5′ ends in the general control regulatory mutants RH1371 (*gcd2-1*) and RH1381 (*gcn4-101*) and the wild-type strain RH1372 by both primer extension and S1 nuclease protection analysis (Fig. 7A). Three major *HIS7* mRNA 5′ ends could be mapped at positions −96, −88/89, and −60/64. The pattern of transcription start sites used was the same in all three genetic backgrounds, but the intensity of the signals at position −60/−64 relative to those at −88/89 and −96 was at least fivefold stronger in a *gcd2-1* background than in a *gcn4-101* background. Thus GCN4-dependent transcription uses preferentially the downstream located initiator elements at position −60/−64. The *HIS7* gene belongs therefore to the same transcription initiation type as *HIS3* and *TRP4*. The *HIS7* transcript 3′ ends were determined by the S1 nuclease protection method. Three ends located at positions +1723, +1739, and +1747 could be mapped in both a *gcd2-1* and a *gcn4-101* background (Fig. 7B). Thus the 3′ untranslated region of the *HIS7* gene has a length of 64 to 88 bp taking the

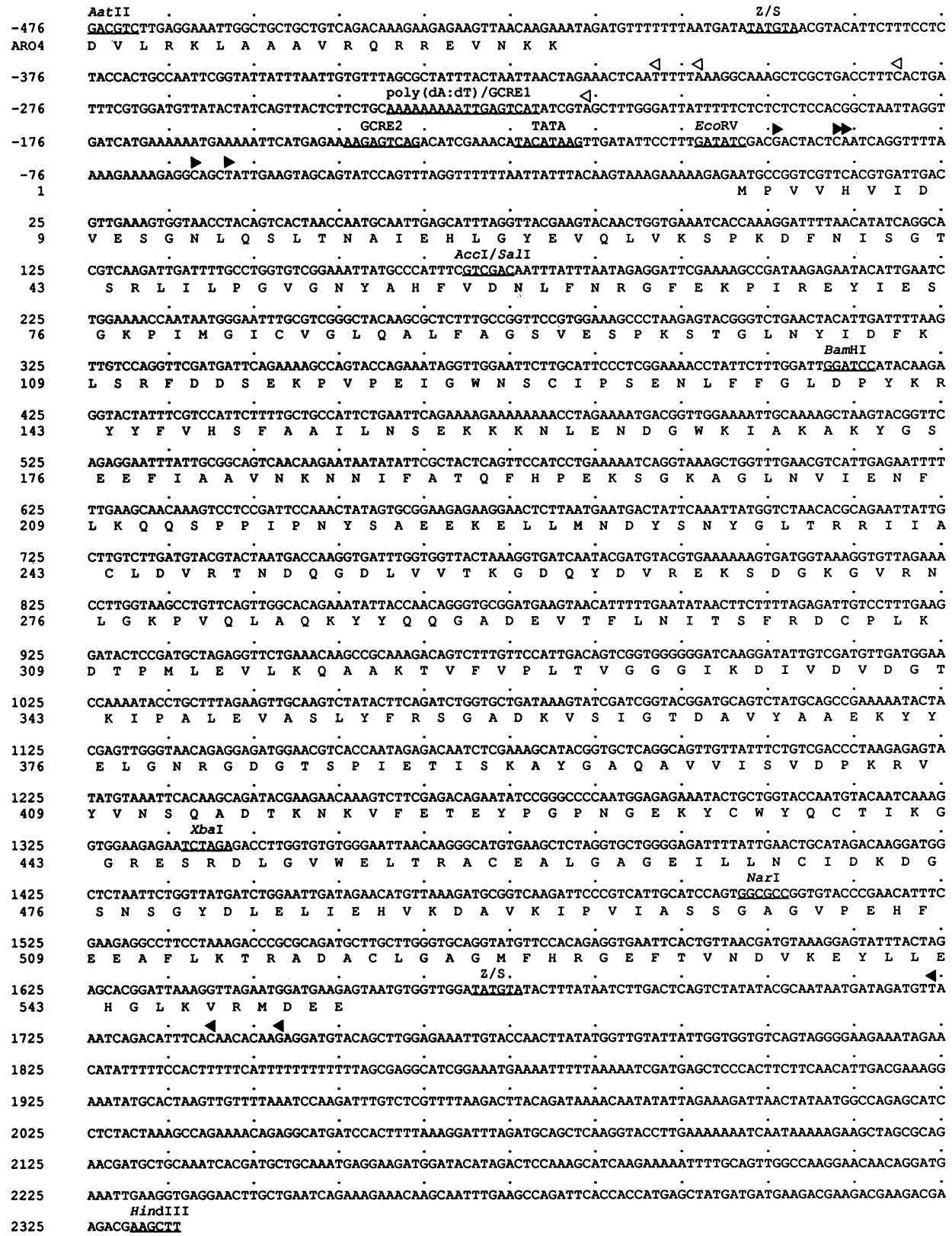


FIG. 3. Nucleotide sequence of the *HIS7* gene and deduced amino acid sequence for the encoded bifunctional glutamine amidotransferase: cyclase. Nucleotide numbering refers to the A (+1) of the first ATG in the open reading frame. The presented sequence comprises the whole *ARO4-HIS7* intergenic region and the last 19 codons of the *ARO4* open reading frame (ARO4). Mapped *HIS7* transcript 5′ and 3′ ends are indicated by solid arrowheads; *ARO4* transcript 3′ ends (28) are indicated by open arrowheads. Relevant restriction sites, the Zaret/Sherman consensus sequence for transcript 3′ end formation (Z/S), a poly(dA-dT) stretch, two putative GCN4 recognition elements (GCRE), and a putative TATA element in the *HIS7* promoter region, are indicated and underlined.

A

```
ScHIS7                MPVVHVID VESGNLQSLT NAIEHLGYEV ....QLVKSP  34
EcHisH                 MNVVIID TGCANLNSVK SAIARHGY.. ....EPKVSR  31
StHisH                 MNVVIID T.CANLSSVK SAGAPR.L.. ....HPGGQR  29
LlHisH                 MKKIVII D YNIGNLQSVQ AAFLRLGQ.. ....ETVISR  32
AbHisH                 METVALID YGSGNLRSPA KALERAAAGC HASFQVLVTS  38
StcHisH  LTAAVPAGAT GRAKKVVVFD YGFGNVRSAE RALARAGA.. ....DVEITR  44

ScHIS7   KDFNISGTSR LILPGVGNYA HFVDNLFNRG FEKPIREYIE SGKPIMGICV  84
EcHisH   DPDVVLLADK LFLPGVGTAQ AAMDQVRERE LFDLIK...A CTQPVLGICL  78
StHisH   EAEIVLRADK LFLPGVGTAQ AAMDQVRERE LIDLIK...A CTQPVLGICL  76
LlHisH   DLEEIRKADA LILPGVGAFP TAMNNLKKFN LIELIQERAA AGIPILGICL  82
AbHisH   DADAVRKADR VVLPAVGAFA DCKRGLSEVP ..........  ..........  68
StcHisH  DYDKAMNADG LLVPGVGAFA ACMEGLKAAR GDWIVDRRLS GGRPVMGICV  94

ScHIS7   GLQALFAGSV E.SPKSTGLN YIDFKLSRFD DSE..KPVPE IGWNSCIP..  129
EcHisH   GMQLLGRRS EESNGVDILG IIDEDVPKMT DFG..LPLPH MGWNRVYPQA  125
StHisH   GMQLLGRRS EETRGVDILN IIEQDVPKMT DFG..LPIPH MGWNRVYPQR  123
LlHisH   GMQVLFEKGY EIE.ERQGLG LLKGEVIPIK TNE..K.IPH MGWNQLNLAK  128
AbHisH   GMQLMAEGGR EYGVT.EGLG WIKGEVVKLE PADPTLKIPH MGWNELDIRR  117
StcHisH  GMQILFSRGI EHDVEAEGLD EWPGTVGPLE A.D..V.VPH MGWNTVEAPA  140

ScHIS7   SENLFFGLDP YKRYYFVHSF AAILNSEKKK NLENDGWKIA KAKYGSEEFI  179
EcHisH   GNRLFQGIED GAYFYFVHSY AMP.......  ...VNPWTIA QCNYGEPFTA  165
StHisH   GNRLFQGIED GAYFYFVHSY AMP.......  ...VNPWTIA QCNYGEPFTA  163
LlHisH   TSPTTHYLSG NDEVYFVHSY Q....ATCPD DELIA.....  YTTYGEVKIP  169
AbHisH   EHPVLAGLRE RAHAYFVHSY RFAVERPED. .......VIA SADYGGPF.A  158
StcHisH  DSQLFAGLDA DARFYFVHSY AVHEWTQESH NPLIAEPRVT WSTHGKPFVA  190

ScHIS7   AAVNKNNIFA TQFHPEKSGK AGLNVIENFL KQQSPPIPNY SAEEKELLMN  229
EcHisH   A.VQKDNFYG VQFHPERSGA AGAKLLKNFL EM                     196
StHisH   A.VQKDNFFG VQFHPERSGA AGAQLLKNFL EM                     194
LlHisH   AIVGKNNVIG CQFHPEKSGE IGRKILKAFL EEI                    202
AbHisH   AVVGRDNLVG TQFHPEKSDE TGLALVANFL TWRV                   192
StcHisH  A.VENGALWA TQFHPEKSGD AGAQLLTNWI ETL                    222

ScHIS7   DYSNYGLTKR IIACLDVRTN DQGDLVVIKG DQMDVREKSD GKGVRNLGKP  279
EcHisF         MLAKR IIPCLDVRDG Q.....VVKG VQF.......  .RNHEIIGDI  32
StHisF         MLAKR IIPCLDVRDG Q.....VVKG VQF.......  .RNHEIIGDI  32
LlHisF         MLTKR IIPCLDIKNG K.....VVKG INF.......  .VGLREIGDP  32
AbHisF         MLKMR VIPCLDVKDG R.....VVKG VNF.......  .VDLIDAGDP  32

ScHIS7   VQLAQKYYQQ GADEVIFLNI TSFRQCPLKD TPMLEVLKQA AKTVFVPLTV  329
EcHisF   VPLAKRVAEE GADELVFYDI TASSDGRVVD K...SWVSRV AEVIDIPFCV  79
StHisF   VPLAKRVADE GADELVFYDI TASSDGRVVD K...SWVARV AEVIDIPFCV  79
LlHisF   VELAKIMEEQ CADEIVFLDI TASFEEREII G...ELIGRA ARELSIPLTV  79
AbHisF   VEQARVVDRE GADELTFLDI TASHENRDTI Y...DVVRRT AEQVFMPLTV  79

ScHIS7   GGGIKDIVDV DGTKIPALEV ASLYFRSGAD KVSIGTDAVY AAEKYYELGN  379
EcHisF   AGGIKSLEDL ..........  AAKILSFGAD KISINSPALA DPTLITRLAD  118
StHisF   AGGIRSIDDL ..........  AAKILSFGAD KISINSPALA DPTLITRLAD  118
LlHisF   GGGIRSIDDL ..........  FRRILARGAD KVSVNSAAIE NPELIRQAAN  118
AbHisF   GGGVRTVDDL ..........  IRKLLLAGAD KVSINTAAIH RPEFVQEAAE  118

ScHIS7   RGDGTSPIET ISKAYGAGAV VISVDPKRVY VNSQADTKNK VFETEYPGPN  429
EcHisF   R.........  ....FGVQQI VVGIDTW...  ..........  .........Y  133
StHisF   R.........  ....FGVQQI VVGIDTW...  ..........  .........F  133
LlHisF   E.........  ....FGVQQV VVAIDAK...  ..........  .........K  133
AbHisF   K.........  ....FGAQQI VVAIDAK...  ..........  .........Q  133

ScHIS7   GEKYCWVQCT IKGGRESR.. ..DLGVWELT RACEALGAGE ILLNCIDKDG  475
EcHisF   DAETGKYHVN QYTGDESRTR VTQWETLDWV QEVQKRGAGE IVLNMMNQDG  183
StHisF   DDATGKYHVN QYTGDENRTR VTQWETLDWV QEVQQRGAGE IVLNMMNQDG  183
LlHisF   RADHRGVDVY IKGGREN... .AGLDLVDWA KKCERLGAGE ILLTSMDKDG  179
AbHisF   VEPGR.WEIF THGGRKA... .TGIDAIEWA KRMESYGAGE ILLTSMDRDG  178

ScHIS7   SNSGYDLELI EHVKDAVKIP VIASSGAGVP EHFEEAFLKT RADACLGAGM  525
EcHisF   VRNGYDLEQL KKVREVCHVP LIASGGAGTM EHFLEAFRDA DVDGALAASV  233
StHisF   VRNGYDLTQL KKVRDVCRVP LIASGGAGTM EHFLEAFRDA DVDGALAASV  233
LlHisF   TKTGYDLEML NDVCTAVNIP VVASGGCGKI SDIVEVFQNT RSDAALEASL  229
AbHisF   TKSGFDLALT RKVADGLRIP VIASGGVGTL DHLVEGIREG HATAVLAASI  228

ScHIS7   FHRGEFTVND VKEYLLEHGL KVRMDEE          552
EcHisF   FHKQIINIGE LKAYIATQGV EIRIC            258
StHisF   FHKQIINIGE LKAYILAGQGV EIRIC           258
LlHisF   FHYGEEQLMK LKTNL                       244
AbHisF   FHFGTYTIGQ AKAALAEAGI PVRPARMAEA AHG 261
```

B

Organism	Glutamine-amidotransferase (HisH)	Cyclase (HisF)
Escherichia coli	62.4 / 39.2 %	61.9 / 36.6 %
Salmonella typhimurium	57.3 / 35.9 %	61.1 / 36.2 %
Lactococcus lactis	54.3 / 36.0 %	65.4 / 44.9 %
Azospirillum brasilense	57.1 / 36.3 %	64.3 / 40.4 %
Streptomyces coelicolor	56.3 / 33.5 %	-/-

FIG. 4. Comparison of deduced amino acid sequences for different prokaryotic monofunctional glutamine amidotransferases (HisH) and cyclases (HisF) with the *HIS7* gene product from *S. cerevisiae*. (A) Multisequence alignment. The various sequences (*Ec*: *E. coli*; *St*: *S. typhimurium*; *Ll*: *L. lactis*; *Ab*: *A. brasilense*; *Stc*: *S. coelicolor*) were obtained from the GenBank/EMBL data base and aligned with the deduced HIS7 amino acid sequence (top line). *S. coelicolor* genes are designated as suggested by Limauro et al. (30). Residues similar in all compared sequences are boxed. (B) Pairwise comparisons. Identities and similarities (%) of the N-terminal segment of the HIS7 sequence (amino acids 1 to 213) with the various glutamine amidotransferase sequences (HisH) and of the C-terminal segment (amino acids 235 to 552) with the various cyclase sequences (HisF) were calculated on the basis of pairwise alignments.

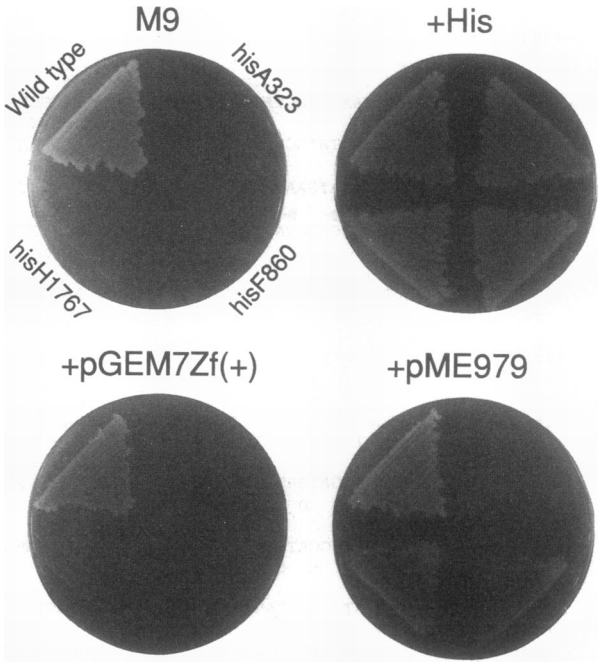


FIG. 5. Suppression of His auxotrophy in *E. coli* by the yeast *HIS7* gene. *E. coli* K-12 derivatives W3110 (wild type), UTH6 (*hisA323*), UTH860 (*hisF860*) and UTH1767 (*hisH1767*) harboring either no plasmid (M9, +His), the empty vector [pGEM7Zf(+) (Promega, Madison, Wis.)] or pGEM7Zf(+) bearing a 2.4-kb *Eco*RV-*Hin*dIII *HIS7* fragment (pME979) were streaked onto M9 minimal glucose agar [M9, +pGEM7Zf(+), +pME979] or M9 minimal glucose agar supplemented with 20 mg of histidine (+His) per liter and incubated for 3 days at 37°C. Complementation of a *hisA* mutation was tested as the *hisA* gene is located between the genes *hisH* and *hisF* in the *E. coli* histidine operon (Fig. 1).

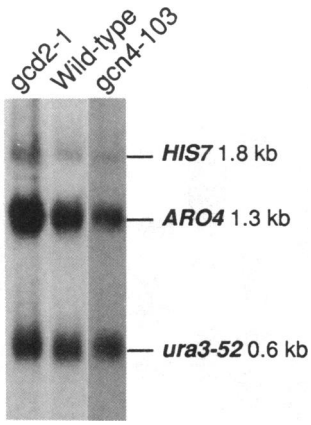
translational stop codon at position +1657 into account. According to the transcript end-mapping experiments, the *HIS7* transcripts have a length of approximately 1.8 kb, which corresponds to the length determined by Northern analysis (Fig. 6A).

DISCUSSION

In this report, we provide primary structure, complementing activity in *E. coli*, and regulation of the *S. cerevisiae HIS7* gene. The gene codes for the first eukaryotic glutamine amidotransferase and cyclase catalyzing the fifth and sixth step in the histidine biosynthetic pathway. The primary structure and function of the two enzymatic activities are shown to be conserved from *E. coli* to *S. cerevisiae*, whereas structure, organization, and regulation of the corresponding genes differ considerably in the two organisms (Fig. 1).

In *E. coli* the two enzymatic activities are encoded by two separate cistrons, *hisH* (glutamine amidotransferase) and *hisF* (cyclase), organized in a single operon. The two cistrons are separated by the *hisA* cistron, which codes for

A



B

Strain	Relevant genotype	ß-Galactosidase-activity (U)
RH1614	*gcd2-1*	143
RH1615	Wild-type	35
RH1616	*gcn4-101*	25

FIG. 6. *HIS7* expression analysis in *S. cerevisiae* strains expressing different amounts of GCN4 protein. (A) Northern analysis. Poly(A)$^+$ RNA of strains RH1632 (*gcd2-1*, high amount of GCN4 protein), RH1631 (wild type, intermediate amount of GCN4 protein), and F194 (*gcn4-103*, no GCN4 protein) was hybridized against DNA probes for the *HIS7*, *ARO4*, and *URA3* transcripts. Sizes of the various transcripts are indicated. The *URA3* transcript was chosen as a negative control and the *ARO4* transcript was chosen as a positive control for a regulation by the general control system of amino acid biosynthesis. (B) β-Galactosidase activity of integrated *HIS7-lacZ* fusions. Activities have been determined for strains indicated. Yeast strains harboring no *E. coli lacZ* gene did not show any detectable β-galactosidase activity (data not shown).

phosphoribosyl-formimino-5-amino-1-phosphoribosyl-4-imidazole-carboxamide isomerase (EC 5.3.1.16) catalyzing the fourth step of the pathway. In *S. cerevisiae* both enzymatic activities are fused on a single polypeptide chain. The *HIS7* gene codes for a bifunctional enzyme with an N-terminal glutamine amidotransferase and a C-terminal cyclase domain. Besides the HIS7 enzyme, there is only one other multifunctional enzyme in the histidine biosynthetic pathway in *S. cerevisiae* (Fig. 1). The *HIS4* gene product catalyzes four steps in the pathway and seems to be unique to fungi, as the four enzymatic activities are catalyzed by at least two distinct enzymes both in prokaryotes and plants (see introduction). The *HIS7* gene product probably represents an analogous situation, as in no prokaryote studied so far are these enzymatic steps fused on a single polypeptide chain (5, 11, 14, 17, 30) and furthermore in *N. crassa* the steps seem to be genetically coupled as well (1). By contrast, the physical uncoupling of the seventh and ninth step of the pathway in *S. cerevisiae* compared with the situation in enterobacteria is not unique to fungi, as the same situation is found both in eubacteria such as *L. lactis*, *S. coelicolor*, and *A. brasilense* (14, 17, 30) and in the methanogenic archaebacterium *M. vannielii* (6).

The *HIS7* gene was located on chromosome II immediately downstream of the previously isolated and characterized *ARO4* gene (28). The *ARO4* gene codes for the tyrosine-inhibitable DAHPS catalyzing the first step in the biosynthetic pathway of the aromatic amino acids. The two genes have the same orientation with a distance of only 416 bp between the two open reading frames. The mapped *ARO4* mRNA 3′ ends and *HIS7* mRNA 5′ ends are only 121 bp apart (Fig. 3) which leaves little space for signal sequences directing termination of *ARO4* transcription and initiation of *HIS7* transcription. Such a close packing of two independent genes is not untypical for *S. cerevisiae* (40). Nevertheless, it is remarkable that both genes are involved in amino acid biosynthetic pathways and coregulated by the general control system of amino acid biosynthesis in *S. cerevisiae*. Both the *HIS7* and the *ARO4* gene (28) are derepressed several-fold under amino acid starvation conditions. To our knowledge, there is no other example of two coregulated genes adjacent to each other in *S. cerevisiae*.

In the *HIS7* gene GCN4-dependent transcription uses only selected initiator elements used by GCN4-independent (basal) transcription, whereas in the *ARO4* gene the same elements are used under both repressing and derepressing conditions (28). In this respect the *HIS7* gene resembles the *HIS3* gene coding for imidazoleglycerol-phosphate dehydratase (51). Of the two mapped *HIS3* transcripts, only the one located more downstream is subject to GCN4 control. Constitutive and regulated *HIS3* transcription differ not only by their utilization of initiator elements but also by their required upstream promoter elements. Upstream elements in the *HIS3* promoter include a poly(dA-dT) stretch for constitutive transcription and two GCN4 recognition elements (GCRE) (21) for maximal induction by GCN4. In addition, two classes of TATA elements have been suggested, responsible for constitutive and GCN4-regulated *HIS3* transcription, respectively. In the *HIS4* gene the start site patterns for GCN4-dependent and GCN4-independent transcription are identical (16). The *HIS4* promoter contains not less than five GCREs for derepression by GCN4 (3). GCN4-independent transcription of the *HIS4* gene is controlled by the global activators BAS1 and BAS2 (4, 52) which have also been shown to regulate purine biosynthesis (13). In addition, BAS2 (also known as PHO2) is involved in the regulation of phosphate metabolism (53) and tryptophan biosynthesis (8). The *HIS4* TATA element (37) is required for correct mRNA start site selection by GCN4-dependent transcription but not by GCN4-independent transcription (43).

In the *HIS7* promoter region various putative upstream elements are present. Two possible GCREs, a one-mismatch consensus sequence (ATGACTCAA in inverse orientation) at position −228 (GCRE1) and a two-mismatch consensus sequence (CTGACTCTT in inverse orientation) at position −142 (GCRE2), are found (Fig. 3). The latter element contains the hexanucleotide sequence TGACTC followed by a T. This motif is found in several binding sites of BAS1 to DNA (13). As a further putative upstream element for

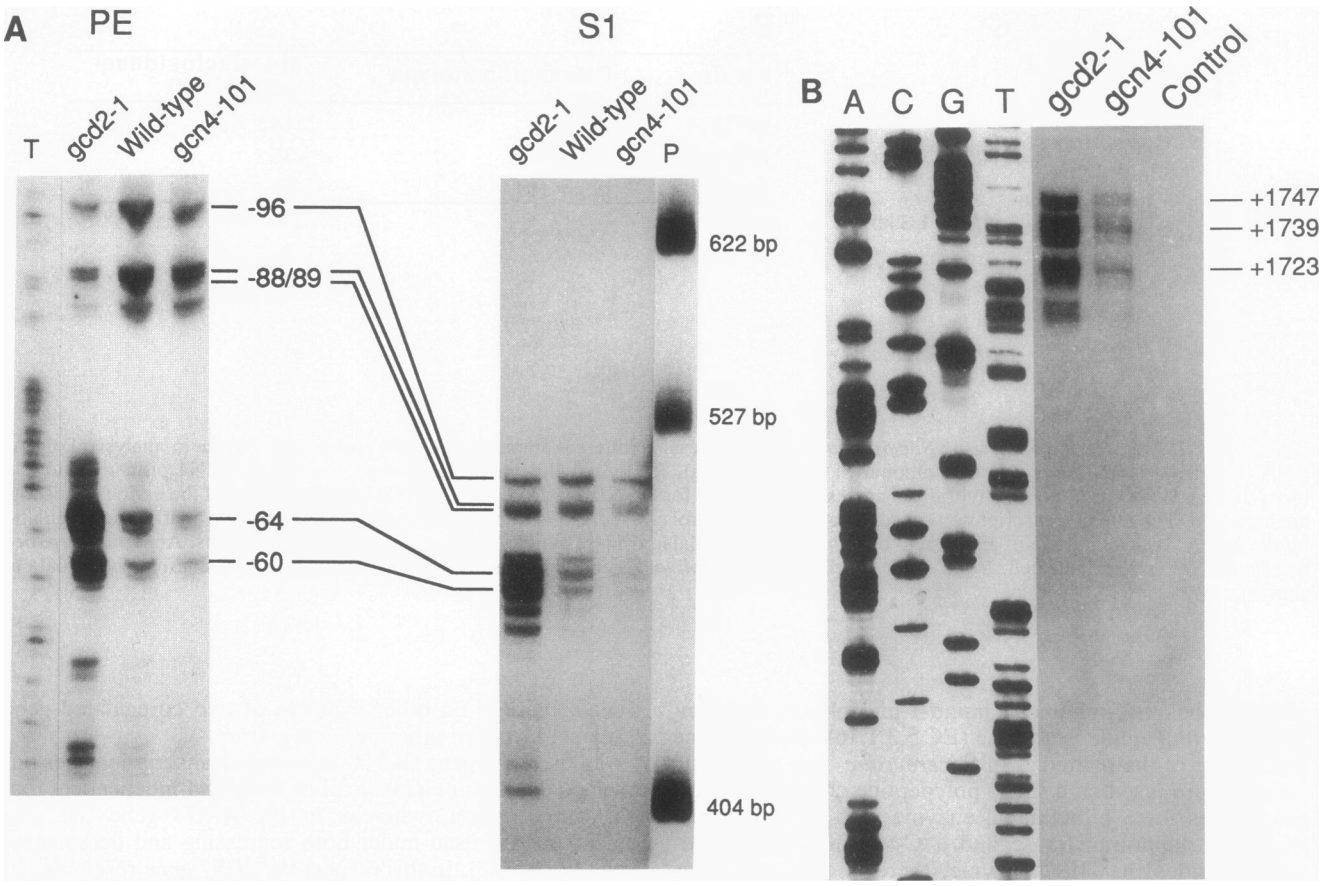


FIG. 7. Mapping of *HIS7* transcript ends in *S. cerevisiae* strains expressing different amounts of GCN4 protein. (A) *HIS7* mRNA 5′ end mapping. The 5′ ends of the *HIS7* transcript were determined in strains RH1371 (*gcd2-1*), RH1372 (wild type), and RH1381 (*gcn4-101*) both by primer extension analysis (PE) and S1 nuclease mapping (S1). A T ladder (T) produced with the same primer as for the primer extension reactions was used as a standard for the size of the elongation products. For S1 nuclease mapping, plasmid pBR322 DNA cut with *Hpa*II (P) served as a size marker. (B) *HIS7* mRNA 3′ end mapping. The 3′ ends of the *HIS7* transcript were determined in strain RH1371 (*gcd2-1*) and RH1381 (*gcn4-101*). The DNA fragment used for S1 nuclease mapping of the mRNA 3′ ends was produced by PCR technique. As a size standard a sequence ladder generated with one of the primers used in the PCR reaction was used. Torula yeast RNA (30 μg) was chosen as a negative control.

constitutive transcription, an 11-bp poly(dA-dT) stretch is located at position −236, immediately upstream of GCRE1. The only sequence element in the *HIS7* promoter region resembling a TATA element (55) is found as TACATAAG in normal orientation at position −122. In the 3′ untranslated regions of both the *HIS7* and *ARO4* genes perfect matches to the sequence TATGTA (57) are found as possible signal sequences for transcript 3′ end formation (Fig. 3).

Further experiments will concentrate on the identification and characterization of control elements in the *ARO4-HIS7* intergenic region involved in GCN4-independent (basal) or GCN4-dependent *HIS7* transcription and eventually in preventing transcriptional interference between the two adjacent and coregulated genes.

ACKNOWLEDGMENTS

We thank Ralf Hütter for his continued interest and support. We are especially grateful to Phil Hieter for providing the yeast genomic DNA library and Barbara Bachmann for providing the *E. coli his* mutants. We also appreciate Hans-Ueli Mösch for helpful discussions and Roney Graf and Brigitta Mehmann for critical reading of the manuscript.

This work was supported by the Swiss National Foundation (grant no. 31.29926.90) and the Swiss Federal Institute of Technology.

REFERENCES

1. **Ahmed, A., M. E. Case, and N. H. Giles.** 1964. The nature of complementation among mutants in the *histidine-3* region of *Neurospora crassa*. Brookhaven Symp. Biol. **17**:53–65.
2. **Altboum, Z., S. Gottlieb, G. A. Lebens, I. Polacheck, and E. Segal.** 1990. Isolation of the *Candida albicans* histidinol dehydrogenase (*HIS4*) gene and characterization of a histidine auxotroph. J. Bacteriol. **172**:3898–3904.
3. **Arndt, K., and G. R. Fink.** 1986. GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5′ TGACTC 3′ sequences. Proc. Natl. Acad. Sci. USA **83**:8516–8520.
4. **Arndt, K. T., C. A. Styles, and G. R. Fink.** 1987. Multiple global regulators control *HIS4* transcription in yeast. Science **237**:874–880.
5. **Bazzicalupo, M., R. Fani, E. Gallori, L. Turbanti, and M. Polsinelli.** 1987. Cloning of the histidine, pyrimidine and cysteine genes of *Azospirillum brasilense*: expression of pyrimidine and three clustered histidine genes in *Escherichia coli*. Mol. Gen. Genet. **206**:76–80.
6. **Beckler, G. S., and J. N. Reeve.** 1986. Conservation of primary structure in the *hisI* gene of the archaebacterium, *Methanococcus vannielii*, the eubacterium *Escherichia coli*, and the eucaryote *Saccharomyces cerevisiae*. Mol. Gen. Genet. **204**:133–140.
7. **Bennetzen, J. L., and B. D. Hall.** 1982. The primary structure of

the *Saccharomyces* gene for alcohol dehydrogenase I. J. Biol. Chem. **257**:3018–3025.
8. **Braus, G., H.-U. Mösch, K. Vogel, A. Hinnen, and R. Hütter.** 1989. Interpathway regulation of the *TRP4* gene of yeast. EMBO J. **8**:939–945.
9. **Brenner, M., and B. N. Ames.** 1971. The histidine operon and its regulation, p. 349–387. *In* H. J. Vogel (ed.), Metabolic pathways, vol. 5. Metabolic regulation. Academic Press, Inc., New York.
10. **Broach, J. R.** 1981. Genes of *Saccharomyces cerevisiae*, p. 653–727. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast *Saccharomyces*. Life cycle and inheritance. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Carlomagno, M. S., L. Chiariotti, P. Alifano, A. G. Nappo, and C. B. Bruni.** 1988. Structure and function of the *Salmonella typhimurium* and *Escherichia coli* K-12 histidine operons. J. Mol. Biol. **203**:585–606.
12. **Casadaban, M. J., A. Martinez-Arias, S. K. Shapira, and J. Chou.** 1983. β-Galactosidase gene fusions for analysing gene expression in *Escherichia coli* and yeast. Methods Enzymol. **100**:293–308.
13. **Daignan-Fornier, B., and G. R. Fink.** 1992. Coregulation of purine and histidine biosynthesis by the transcriptional activators BAS1 and BAS2. Proc. Natl. Acad. Sci. USA **89**:6746–6750.
14. **Delorme, C., S. D. Ehrlich, and P. Renault.** 1990. Histidine biosynthesis genes in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. **174**:6571–6579.
15. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
16. **Donahue, T. F., P. J. Farabaugh, and G. R. Fink.** 1982. The nucleotide sequence of the *HIS4* region of yeast. Gene **18**:47–59.
17. **Fani, R., M. Bazzicalupo, G. Damiani, A. Bianchi, C. Schipani, V. Sgaramella, and M. Polsinelli.** 1989. Cloning of histidine genes of *Azospirillum brasilense*: organization of the *ABFH* gene cluster and nucleotide sequence of the *hisB* gene. Mol. Gen. Genet. **216**:224–229.
18. **Feinberg, A. P., and B. Vogelstein.** 1984. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **137**:266–267.
19. **Fink, G. R.** 1964. Gene-enzyme relations in histidine biosynthesis in yeast. Science **146**:525–527.
20. **Furter, R., G. Paravicini, M. Aebi, G. Braus, F. Prantl, P. Niederberger, and R. Hütter.** 1986. The *TRP4* gene of *Saccharomyces cerevisiae*: isolation and structural analysis. Nucleic Acids Res. **14**:6357–6373.
21. **Hill, D. E., I. A. Hope, J. P. Macke, and K. Struhl.** 1986. Saturation mutagenesis of the yeast *HIS3* regulatory site: requirements for a transcriptional induction and for binding by GCN4 activator protein. Science **234**:451–457.
22. **Hinnebusch, A. G.** 1985. A hierarchy of *trans*-acting factors modulates translation of an activator of amino acid biosynthetic genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **5**:2349–2360.
23. **Hinnebusch, A. G.** 1988. Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. Microbiol. Rev. **52**:248–273.
24. **Hinnebusch, A. G., and G. R. Fink.** 1983. Repeated DNA sequences upstream from *HIS1* also occur at several other co-regulated genes in *Saccharomyces cerevisiae*. J. Biol. Chem. **258**:5238–5247.
25. **Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf.** 1985. Genetic manipulation of *Streptomyces*. A laboratory manual. John Innes Foundation, Norwich, United Kingdom.
26. **Jones, E. W., and G. R. Fink.** 1982. Regulation of amino acid and nucleotide biosynthesis in yeast, p. 181–299. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast *Saccharomyces*. Metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Kassavetis, G. A., and E. P. Geiduschek.** 1982. Bacteriophage T4 late promoters: mapping 5′ ends of T4 gene 23S mRNAs. EMBO J. **1**:107–114.
28. **Künzler, M., G. Paravicini, C. M. Egli, S. Irniger, and G. H. Braus.** 1991. Cloning, primary structure and regulation of the *ARO4* gene, encoding the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae*. Gene **113**:67–74.
29. **Legerton, T. L., and C. Yanofsky.** 1985. Cloning and characterization of the multifunctional *his-3* gene of *Neurospora crassa*. Gene **39**:129–140.
30. **Limauro, D., A. Avitabile, C. Cappellano, A. M. Puglia, and C. B. Bruni.** 1990. Cloning and characterization of the histidine biosynthetic gene cluster of *Streptomyces coelicolor* A3(2). Gene **90**:31–41.
31. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
32. **Minton, N. P.** 1984. Improved plasmid vectors for the isolation of translational *lac* gene fusions. Gene **31**:269–273.
33. **Miozzari, G., P. Niederberger, and R. Hütter.** 1978. Tryptophan biosynthesis in *Saccharomyces cerevisiae*: control of the flux through the pathway. J. Bacteriol. **134**:48–59.
34. **Mortimer, R. K., C. R. Contopoulou, and J. S. King.** 1992. Genetic and physical maps of *Saccharomyces cerevisiae*, edition 11. Yeast **8**:817–902.
35. **Mösch, H.-U., R. Graf, and G. H. Braus.** 1992. Sequence-specific initiator elements focus initiation of transcription to distinct sites in the yeast *TRP4* promoter. EMBO J. **11**:4583–4590.
36. **Nagai, A., E. Ward, J. Beck, S. Tada, J.-Y. Chang, A. Scheidegger, and J. Ryals.** 1991. Structural and functional conservation of histidinol dehydrogenase between plants and microbes. Proc. Natl. Acad. Sci. USA **88**:4133–4137.
37. **Nagawa, F., and G. R. Fink.** 1985. The relationship between 'TATA' sequence and transcription initiation sites at the *HIS4* gene of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **82**:8557–8561.
38. **Niederberger, P., M. Aebi, and R. Hütter.** 1986. Identification and characterization of four new *GCD* genes in *Saccharomyces cerevisiae*. Curr. Genet. **10**:657–664.
39. **Nishiwaki, K., N. Hayashi, S. Irie, D.-H. Chung, S. Harashima, and Y. Oshima.** 1987. Structure of the yeast *HIS5* gene responsive to general control of amino acid biosynthesis. Mol. Gen. Genet. **208**:159–167.
40. **Oliver, S. G., et al.** 1992. The complete DNA sequence of yeast chromosome III. Nature (London) **357**:38–46.
41. **Paravicini, G., G. Braus, and R. Hütter.** 1988. Structure of the *ARO3* gene of *Saccharomyces cerevisiae*. Mol. Gen. Genet. **214**:165–169.
42. **Pattee, P. A., H. C. Lee, and J. P. Bannantine.** 1990. Genetic and physical mapping of the chromosome of *Staphylococcus aureus*, p. 42–56. *In* R. P. Novick (ed.), Molecular biology of the staphylococci. VCH Publishers, New York.
43. **Pellman, D., M. E. McLaughlin, and G. R. Fink.** 1990. TATA-dependent and TATA-independent transcription at the *HIS4* gene of yeast. Nature (London) **348**:82–85.
44. **Piggot, P. J., and J. A. Hoch.** 1985. Revised genetic linkage map of *Bacillus subtilis*. Microbiol. Rev. **49**:158–179.
45. **Ratzkin, B., and J. Carbon.** 1977. Functional expression of cloned yeast DNA in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **74**:487–491.
46. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics: a laboratory course manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
47. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–494.
48. **Sambrook, J., E. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

49. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
50. **Struhl, K., J. R. Cameron, and R. W. Davis.** 1976. Functional genetic expression of eukaryotic DNA in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **73**:1471–1475.
51. **Struhl, K., W. Chen, D. E. Hill, I. A. Hope, and M. A. Oettinger.** 1985. Constitutive and coordinately regulated transcription of yeast genes: promoter elements, positive and negative regulatory sites, and DNA binding proteins. Cold Spring Harbor Symp. Quant. Biol. **50**:489–503.
52. **Tice-Baldwin, K., G. R. Fink, and K. T. Arndt.** 1989. BAS1 has a Myb motif and activates *HIS4* transcription only in combination with BAS2. Science **246**:931–935.
53. **Vogel, K., W. Hörz, and A. Hinnen.** 1989. The two positively acting regulatory proteins PHO2 and PHO4 physically interact with *PHO5* upstream activation regions. Mol. Cell. Biol. **9**:2050–2057.
54. **Winkler, M. E.** 1987. Biosynthesis of histidine, p. 399–411. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
55. **Wobbe, C. R., and K. Struhl.** 1990. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. Mol. Cell. Biol. **10**:3859–3867.
56. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33**:103–109.
57. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28**:563–573.