Structure and function of the yeast UR3 gene: expression in Escherichia coli

(Recombinant DNA; DNA sequence polymorphisms; OMP decarboxylase; promoters; Saccharomyces cerevisiae; plasmid and phage vectors)

Mark Rose, Paula Grisafi and David Botstein

Department of Biology, 56-344, Massachusetts Institute of Technology, Cambridge, MA 02139 (U.S.A.) Tel. (617) 253-7240

(Received October 12th, 1983)
(Revised received February 2nd, 1984)
(Accepted February 7th, 1984)

SUMMARY

The expression of the cloned Saccharomyces cerevisiae UR3 gene in Escherichia coli on both plasmid and phage vectors was studied. Isolates of the gene from two different laboratory strains of yeast differ in their ability to be expressed in E. coli in the absence of external adjacent promoters of transcription. The DNA sequence of the two genes was determined and revealed several differences in the DNA flanking the structural gene. One base change alters the "Purbox" of an E. coli promoter present in the yeast sequences. Three amber alleles of the yeast gene were also cloned from yeast. Two of the alleles could be suppressed in E. coli by a tRNA suppressor mutation. One of the amber alleles was determined to be a mutation in the seventh codon of the structural gene, thereby establishing the reading frame and extent of the coding sequence. The initiation codon of the reading frame encoding the UR3 structural gene is preceded by two other ATG codons in a different reading frame 61 and 79 bp away. The nearer ATG begins an open reading frame that overlaps the structural gene sequence by 17 bp. With the DNA sequence of the UR3 gene many of the common yeast vector plasmids are now completely known at the level of DNA sequence.

INTRODUCTION

The enzyme OMP decarboxylase (orotidine-5'-phosphate carboxyl-lyase, EC 4.1.1.23) is specified in S. cerevisiae (yeast) by the UR3 gene. Expression is regulated at the level of transcription (Bach et al., 1979). Increased intracellular levels of dihydro-orotate lead to increased levels of the enzyme (Lacroix, 1968) and the UR3 mRNA (Bach et al., 1979). A fragment of yeast DNA which expresses this enzyme in E. coli was isolated by complementation of the corresponding E. coli purF mutation (Bach et al., 1979). The mode of expression of the yeast gene in E. coli was obscure, partially due to the lack of knowledge of plasmid promoters and partially due to unexpected changes in the level of functional enzyme expressed under different growth conditions and on different plasmids.

Abbreviations: bp, base pairs; kb, 1000 bp; LB, see MATERIALS AND METHODS, section b; m.o.i., multiplicity of infection; OMP, orotidine-5'-phosphate.
To clarify the means by which the gene is expressed in E. coli, we have recloned the gene onto a phage λ vector. λ vectors have relatively well-characterized promoters and stable copy number in lysogens. We chose to reclone the gene directly from yeast for two reasons. First, the gene was originally isolated on DNA shears fragments (Bach et al., 1979) which contained relatively little yeast DNA sequence flanking the region which expresses OMP decarboxylase. Second, the source of the DNA that had been used for the original isolation of the gene was the genetically heterogeneous and poorly-characterized diploid strain + D4 (Penes et al., 1978). The genetics of the UR3 gene had, on the other hand, been carefully studied in a series of isogenic strains derived from strain FL100. Therefore, we recloned the UR3 gene from strain FL100 as well as several ura3- nonsense mutants derived from FL100.

Analysis of expression in E. coli revealed fundamental differences between the UR3 genes obtained from FL100 and + D4. DNA sequence differences in the region upstream of the genes can explain the observed differences in their expression in E. coli.

**MATERIALS AND METHODS**

(a) Strains

1. **Bacterial strains**
   - The standard strain employed was DB566 (try2, lacZム, pyrF), Mu-1 hadR + hadM) described by Bach et al. (1979). DB6660 was derived from DB566 by P1 transduction to SupF from a strain carrying SII.

2. **Yeast strains**
   - FL100 (MATa, a wild-type) and the three amber mutants ura3-3, ura3-18 and ura3-25 which were derived by mutation from FL100 were all the generous gifts of Francois Lacroute.

(b) Bacteriophages

3. **λ-ﬁlTrads** (Davis et al., 1980) was provided by R. Davis. 2NNM16 (λclac5 str133 / mom21 cts sr143 rnr5 sr153) was obtained from N. Murray. 2NNm434 cts and 2NNm434 cts p4 were obtained from J. Weil (Frindt et al., 1971; Weil et al., 1972). 2NNm6 plclac5 cts H57 was obtained from M. Lichter.

4. **Plasmids**
   - "Clone 1" and "clone 2" are pMB9-derived plasmids carrying clear fragments of yeast DNA containing the UR3 gene. "Clone 6" is a pBR322-derived plasmid containing the 1.1-kb UR3 HindIII fragment. All of the preceding plasmids were described in Buch et al. (1979), pBR322 plasmid DNA was the generous gift of D. Shortle.

(b) Media

- E. coli cells were generally grown in LB (Miller, 1972) broth or on LB agar. Cells used for OMP-decarboxylase assays were grown in M9 salts (Miller, 1972) supplemented with 0.2% glucose or maltose and 1% casamino acids (Difco, technical grade, charcoal-filtered before use). Phage stocks were grown either as plate stocks (Davis et al., 1980) or in liquid (Blatzer et al., 1977). Ampicillin was used at a concentration of 100 μg/ml (Bristol Laboratories).

5. **Tetracycline (Sigma)** was used at a concentration of 15 μg/ml.

(c) Nucleic acid techniques

1. **Enzymes**
   - Restriction endonucleases, E. coli DNA polymerase I, and T4 DNA ligase were purchased from New England Biolabs and used according to the recommendations of the manufacturer. Micrococcus luteus DNA polymerase was purchased from Miles. Bacterial alkaline phosphatase and T4 polynucleotide kinase were purchased from Boehringer-Mannheim.

2. 5′- end labeling
   - We routinely used 1–2 units of M. luteus DNA polymerase in a 40-μl reaction containing 70 mM Tris-HCl pH 8.0, 7 mM MgCl2, 1 mM β-mercaptoethanol, 25 pmol [α-32P]dNTP and 5–10 pmol of DNA ends.

3. 3′- end labeling and DNA sequencing
   - Conditions for phosphatase, phosphorolyzing DNA ends and sequencing were essentially those of Maxam and Gilbert (1980). Nick-translated probe was prepared according to the method of Hohn and Steinhart et al. (1981).

4. **Plaque filter hybridization and nick translation**
   - Hybridizations were done essentially as described by Benton and Davis (1977). Nick-translated probe was prepared according to the method of Hohn and Steinhart et al. (1981).

(d) DNA preparation

1. **λ DNA**
   - Large-scale DNA was prepared by the cesium chloride density gradient method. Small-scale DNA was isolated by the phenol extraction method.

2. **Yeast DNA**
   - High M, yeast DNA was isolated by a modification of the procedure described by Peter et al. (1970).

3. **Plasmid DNA**
   - Large-scale plasmid DNA was prepared by the cesium chloride density gradient method of Hohn and Steinhart et al. (1981).

4. **DNA transport and packaging**
   - E. coli cells were transformed according to the method of Hanahan et al. (1981). Transfection of the DNA/protein mixture was performed by the method of Hohn and Steinhart et al. (1981).

5. **OMP-decarboxylase**
   - OMP-decarboxylase activity was determined in a test-tube assay in 1 ml of 0.1 M Tris-HCl and 1 mM dithiothreitol, the reaction mixture contained 100 μg of total protein, the assay mixture was incubated at 37°C for 30 min. The reaction was stopped by the addition of 1 ml of 10% trichloroacetic acid. The precipitate was centrifuged and 0.2 ml of the supernatant was assayed.
DNA was prepared by a modified method of Rigby et al. (1977) as described in Davis et al. (1980).

(e) DNA transformation, transformation and in vitro λ packaging

E. coli cells were transformed with plasmid DNA according to the method of Mandel and Higa (1970). Transformation with λ DNA was performed according to the modifications of the plasmid transformation procedure described by Davis et al. (1980). Packaging extracts were prepared and used according to the method of Iohn and Murray (1977) using the strains of Sternberg et al. (1977).

(f) OMP-decarboxylase assays

OMP decarboxylase was assayed in E. coli by modification of the method of Lieberman et al. (1955). 25 ml of exponentially growing cells were quick-chilled in an ice-salt bath, centrifuged and resuspended in 1 ml of 0.1 M Tris-HCl pH 8.0, 20% (v/v) glycerol and 1 mM dithiothreitol. Expression from lam promoters was assayed at various times after infection of the indicated E. coli strains at a multiplicity of infection of 2. Cells were broken by sonication and extracts were clarified by centrifugation for 5 min in an Eppendorf centrifuge. Activity was determined in 1 ml reaction mixtures containing 0.1 M Tris-HCl pH 8.0, 20 μg pyridoxal phosphate, 10 mM MgCl₂ and 0.1 mM OMP. Radioactive assays also contained 0.05 μCi of [¹⁴C]OMP. [¹⁴C]CO₂ was collected by absorption into 0.2 ml of 2 N NaOH contained in a plastic reservoir within a 20-ml reaction chamber (Kontes). The reaction was stopped by injection of 0.5 ml of 2 M H₂SO₄ and incubation at 37°C for 180 min. The reservoir was removed and dropped into 5 ml of Aquasol (New England Nuclear) and counted in the ¹⁴C window of a Beckman Scintillation counter. The activity thereby measured was standardized relative to the nonradioactive spectrophotometric assay in which a decrease in absorbance at 285 nm of 1.38 corresponds to a conversion of 1 μmol/ml of OMP to UMP. Protein concentration, when measured, was determined by the method of Bradford (1976) using bovine serum albumin as a standard and reagents purchased from Bio-Rad.

RESULTS

(a) Isolation of the ural gene

A pool of recombinant phages was constructed by ligation of EcoRI-digested yeast DNA into fdg7 (Davis, 1980) vector DNA. Approx. 1 × 10⁶ independent plaques of hybrid phages were obtained and screened by plaque filter hybridization (Benton and Davis, 1977) for homology to ura3 specific DNA using clone 6 (Fig. 1) as probe. Twelve plaques hybridized, of which nine were purified for further study. Two other phages were also obtained by lysogenic complementation of an E. coli protF strain using λ Jam6 pfac 58185 helper phage. Restriction enzyme analysis revealed that all of the phages contained a 13-kb EcoRI fragment which is the same size as the band observed in gel-transfer hybridization analysis of genomic DNA (not shown).

The position and orientation of the ura3 HpaIII fragment within the EcoRI fragments were determined by both heteroduplex and restriction enzyme
Fig. 1. Restriction maps of various cloned fragments of yeast DNA containing the URA3 gene. The upper map is of the 1.1 kb HindIII fragment. The 1.3-kb EcoRI fragment, shown in the middle panel, was cloned from strain FL100. "Close 1" is a pMB9 plasmid containing a single fragment of yeast DNA isolated by Bach et al. (1979) from strain DM4. The restriction sites are as follows: EcoRI, HindIII, BamHI, BglII, SstI, SalI, Smal, Csp1, PvuII, XmaIV, BglII, Smal. None of the restriction sites indicated on the HindIII fragment (Per, NotI and EcoRI) are not indicated on the other maps.

analysis these data are summarized in Fig. 1. Nineteen of the phages containing the HindIII fragment with the PerI site closer to the long arm (containing the "late" genes) of the λ vector are defined as orientation I phages (URA3-I). The two remaining phages (URA3-II) are contained the PerI site closer to the right arm of λ ("early" genes). Two of the phages in orientation I contained an extra HindIII site close to one end of the insert.

(b) Expression of the URA3 gene from phage promoters

To examine the ability of the cloned gene to be expressed in E. coli, cells of the perP mutant were infected with phages bearing URA3 in either orientation. OMP decarboxylase was assayed at several times after infection. Orientation I phages produce the enzyme early in infection, whereas orientation II phage produce the enzyme early and transiently (Fig. 2). This is consistent with expression of an unstable enzyme predominately from phage promoters. Orientation II phage expression would be from the early leftward facing promoter pL. Orientation I phage would be expressed from the late rightward facing promoter pL, as a result of circularization of the λ chromosome. The highest level of activity during the infection corresponds to less than 5% of the level of wild-type E. coli cells grown without uracil.

To determine the level of expression from the URA3 gene in the absence of the strong transcription provided by the phage promoters, it was necessary to construct stable lysogens of the phages. As the hybrid phages contain deletions of both the β and α genes required for lysogenization, we crossed the hybrid phages with phage NMB16 (J. Jambulac-

imm21c(18). Turbulent economy region). Like sequences recombinant for mcv by the red-berry (1976). Candi by restriction enzymes. Ligands of A are able to grow slowly whereas lysogeny pendent on the additional. Enzyme assays contained low but not detectable early detected in the URI of the level present of the maximal levels asymmetries of lytic that the URA3 genomotors for express.

To verify that the level of transcription Q-transactivator isomer of the prophage 1968. Henskowitz superinfected with contains the homologous prophage is specific to prophage induce a heterologous Q activation" showing the Q2 infected with a phage and either the Q region from a crypt can be seen in Fig 6B.

decarboxylase activity (U/mg/min/5% DCM solvent)

Fig. 2. Kinetics of expression of OMP decarboxylase from the URA3 gene carried in pGP75. (a) EDB806 cells (perP-) were grown to 3 x 10^6 cells per ml at 37°C and infected with URA3-I (O) or URA3-II (C) at an moi of 2. At the indicated times aliquots were chilled, sonicated, and frozen at -20°C. Sonicated extracts were prepared and assayed for OMP-decarboxylase activity as described in MATERIALS AND METHODS. Activities are expressed as rate of substrate reacted per min (100) of extract. Such extracts usually contain approx. 2 mg protein/ml. Measurements are the mean of triplicate determinations. (b) Schematic structure of URA3-I. Activity is expressed in terms of the transformant lysates as in Fig 1. (c) Schematic structure of URA3-II. Note that during lysogeny the λ chromosome is circular. Thus the rightward facing late promoter pL, which lies to the right of the O region (Fig. 3b) can transcribe the URA3 gene in late infection.

"
less than 5% of the wild type, with or without uracil, expression from the c1 and c2 promoters was too low to be detected. To eliminate the possibility that the phages as the source of the c1 and c2 promoters, we crossed the JM816 (Δ Jamda4-5:: kanr) strain with MJ29 (λ tnm26; c1) strain. Turbid (containing the JNMB16 immunity region), Lac− (containing the yeast UR43 sequences) recombinants were picked and screened for prototrophy by the red plaque test of Enquist and Weiss (1976). Candidate phages were further checked by restriction enzyme analysis for the correct structure. Lysozymes of UR43-I in the pyrF mutant were able to grow slowly in the absence of added uracil whereas lysogens of UR43-II were completely dependent on the addition of uracil to the growth media. Enzyme assays showed that UR43-I lysogen contained low but detectable levels of OMP decarboxylase whereas the UR43-II lysogen contained no detectable enzyme activity. The level of enzyme detected in the UR43-I lysogen was less than 0.5% of the level present in a pyrF strain and about 10% of the maximal level observed during infection. The enzyme assays of lytic and lysogenic expression suggest that the UR43 gene is dependent on the phage promoters for expression in E. coli.

To verify that the enzyme level is a function of the level of transcription from phage promoters, we used Q transcription to specifically activate the late operon of the prophage (Thomasson, 1966; Dautel et al., 1968; Hershkowitz and Singer, 1974). In lysogens superinfected with a heteroimmune phage, which contains the homologous Q gene, the late operon of the prophage is specifically activated in the absence of prophage induction. A heteroimmune phage with a heterologous Q gene does not cause this "transcription" showing that the activation is specifically due to the Q gene product. Lysozymes of UR43-I were infected with a 0 phage having phage 434 immunity and the eth QSR region or the heterologous region from a cryptic lambda prophage (p4). As can be seen in Fig. 3, the UR43-I lysogen, OMP decarboxylase synthesis is induced by superinfection of λsm434QSR. The heterologous phage is unable to induce the enzyme and the lysogen bearing the opposite orientation of UR43-I within the prophage are not induced. Thus the enzyme level is a function of transcription from the late promoter p2. These experiments imply that the gene as isolated from the yeast strain FL100 does not contain a functional E. coli promoter. In addition, they show that the UR43 gene must be oriented on UR4-I so that it is normally transcribed in the same direction as the late operon. Referring to the HindIII fragment illustrated in Fig. 1 this means that transcription is from the

---

**Fig. 3. Expression of OMP decarboxylase from the λ sm434QSR promotor in lysogens by OMP decarboxylase activity**

(a) Lysogen of UV43-I (□, ○) and UV43-II (●, ▽) were grown at 32°C to a density of 2 x 10⁶ cells/ml in the presence of uracil. Cells were infected with λsm434QSR (□, ○) and λsm434QSR (●, ▽) at an m.o.i. of 4. At the indicated times cells were chilled, concentrated 25-fold and stored frozen at -20°C. Sonicated extracts were prepared and assayed for OMP decarboxylase activity as in Fig. 2. (b) Schematic structure of lysogen of UR43-I. Note that integration changes the order of the DNA relative to the λ promotors. Flanking E. coli DNA is indicated by broken lines. Restriction sites are as in Fig. 1 and 2. (c) Schematic structure of lysogen of UR43-II.

---

Pcl site towards the Smal site. The same conclusion was reached by Hubert et al. (1980).

(c) Expression of the UR43 gene from plasmid promoters

Previously published results on the UR41 gene isolated from the yeast strain x D4 indicated that the HindIII fragment expressed OMP decarboxylase in either orientation at the HindIII site of pBR322, indicating that the gene is expressed in E. coli by a promoter present on the yeast fragment. Our results
indicated that it is unlikely that the URA3 gene from the strain FL100 contains such a promoter. To resolve this apparent contradiction, the URA3 HindIII fragment was isolated from URA3-1 and ligated into the HindIII site of pBR322. Ampicillin-resistant transformants of the pYF strain DB6566 were screened for uracil-independent growth and resistance to tetracycline. Of 45 tetracycline-sensitive colonies isolated, 24 were uracil-dependent. Plasmid DNA was isolated from six Ura" Tet" and four Ura" Tet" colonies. All ten plasmids contained the URA3 HindIII fragment. The orientation of the HindIII fragment correlated exactly with the ability to grow without added uracil. The orientation in which the insert confers uracil independence corresponds to transposition of the URA3 gene in a counter-clockwise fashion on pBR322. This is opposite to the direction of transcription of the tetracycline resistance gene whose promoter is disrupted by insertion at the HindIII site. This is consistent with the existence of a previously identified counter-clockwise promoter in the tetracycline resistance region (Stuber and Bujard, 1981). We repeated the subcloning of the HindIII fragments from the + D4 URA3 containing plasmid "clone 1" using the same batch of pBR322 DNA. Sixteen Ura" plasmids analyzed were equally divided between the two orientations of the URA3 fragment. No difference in growth rate in the absence of uracil was observed amongst transformants carrying different plasmids. Thus, the URA3 genes from yeast strains + D4 and FL100 differ in their ability to be expressed in the E. coli cell.

(d) DNA sequence of the URA3 gene

To determine the reason for the difference in expression of the genes in E. coli and the extent of divergence between them we determined the DNA sequence of both URA3 genes. Two different strategies were used to sequence the two genes. The URA3 HindIII fragment from + D4 was sequenced from the naturally occurring restriction sites within the gene using the chemical method of Maxam and Gilbert (1980). The scheme of fragments which were sequenced is shown in Fig. 4. The gene from FL100 was sequenced by making use of a series of plasmids each of which contained a unique BamHI site which had been artificially introduced at random sites in the URA3 gene (Rose and Botstein, 1983). The position of the BamHI sites used for the sequencing is shown in Fig. 4.

The DNA sequences of the two genes are shown in Fig. 5. A single long open reading frame of 801 bp starts with an ATG codon. The predicted protein is 267 amino acids long, which agrees well with the observed monomer M, of 27 500 (Brody and Wiest- leitner, 1976). As can be seen, the genes from the two yeast strains are very similar in DNA sequence. Ten differences are found (two deletions, one of 5 bp at position 14 and one of 1 bp within a run of T's starting at position 23, six transitions at positions 33, 49, 88, 146, 637, 1132, and two transversions at positions 22 and 704). Two of the mutations change a total of three restriction sites (at position 49 the presence of T in the + D4 sequence destroys HinfI and HpaII sites, at position 704 the FL100 sequence

Fig. 4. Strategy for DNA sequencing of the URA3 1.4-kb HindIII fragment. The HindIII fragment derived from the + D4 background was sequenced from the indicated restriction sites (lower set of arrows). The fragment obtained from the FL100 background (upper set of arrows) was largely sequenced from a set of plasmids containing derivatives of the gene with artificial BamHI sites introduced among positions within the gene. The positions of the BamHI sites are indicated by the tails of the upper set of arrows which have triangles. The length of the arrow is proportional to the amount of sequence obtained from any one labeled site. The method of labeling and therefore the DNA strand from which the sequence was obtained is indicated by the width of the tail of the arrows. Fragments which were 5' labeled by polynucleotide kinase are indicated by the thick tails above the arrow shaft. Fragments which were labeled at the 3' end by DNA polymerase are indicated by thick tails below the shaft. Totally thick tails indicate that both strands were sequenced. The hatched region indicates the URA3 coding region. Restriction sites are as in Fig. 1 and as follows: T, XhoI, A, AsuII, E, HindIII, D, HindIII, P, HpaII, S, SalI, A, PspI.
contains an Atb site that is not present in the +D4 sequence).

Much of the region to the 5' side of the coding region is composed of long stretches in which one strand is largely composed of purines or pyrimidines. This asymmetry switches strand at one place. This feature may be involved in the regulation or expression of the gene in yeast. In addition a fairly good match is present between the sequence TATATA beginning at position 129 and the Goldberg-Hogness box that has been postulated to be important in positioning the start of transcription in eukaryotes (Corden et al., 1980). Of some interest in regards to the expression of the gene in E. coli is the fairly good correspondence between the sequence starting at position 145 and the conserved -10 region of the E. coli promoter. In addition a fairly good match to the -35 region is found with the appropriate spacing at position 125. This is the best match to the consensus sequence that can be made in the region 5' to the coding sequence. One of the base differences between the two sequences is a change from A to G that would mutate one of the most conserved bases in the -10 region. Indeed mutation of this base has been observed to be a promoter down mutation in several systems (Rosenberg and Court, 1979). It is likely that the difference in expression in E. coli observed between the two genes is due to mutation of a fortuitous promoter sequence.

Two of the base changes between the two yeast genes lie within the coding sequence. One of the coding sequence mutations would result in the change of a serine residue to alanine. The second coding sequence change does not change the codon assignment. In addition to the difference in expression in E. coli that we observed, Lacroute, F. (personal communication) has observed antigenic differences between the OMP-decarboxylase proteins produced by the two yeast strains. This observation is explained by the predicted single amino acid difference between the two proteins.

The ATG at position 227 is not the first ATG found in the sequence. Two additional ATGs are present upstream at positions 142 and 160. The first ATG is followed immediately by an in frame termination codon in +D4 or two codons later in strain FL100. The second ATG begins an open reading frame which does not terminate until 28 codons downstream. This open reading frame overlaps the long open reading frame begun by the third ATG by 17 bp. The presence of ATG codons close upstream of eukaryotic genes is unusual (Kozak, 1980).

The DNA sequence downstream of the Ura3 coding sequence contains neither the AATAAA transcription termination signal of higher eukaryotes (Benoit et al., 1980) nor the yeast consensus termination sequence identified by Zurr et al. (1982). However, the sequence TTATTTTTA, identified by Henikoff et al. (1983) as essential for transcription termination of a Drosophila gene in yeast, is present downstream of URA3 at position 1219. Note the presence of a point mutation within this sequence in the FL100 version of the gene.

(e) Amber mutations of the URA3 gene

Three amber mutations of the URA3 gene isolated in the FL100 background were cloned onto 2μ7, essentially as described for the wild-type gene. Lysogens of orientation I plagues were constructed as before using the w' pyrF strain DB6656 and the isogenic sul2 pyrF strain DB6660. Lysogens of lura1-18 and lura1-3 were able to grow in the absence of uracil only in the strain carrying the sul2 suppressor. Lysogens of the third cloned amber mutation were Ura- regardless of the suppressor mutation. Thus two of the genes behave as amber mutations in E. coli as they did in yeast. The mutations aur1-22 presumably has a different suppressor spectrum from the other two mutations. Suppression of the two amber alleles was confirmed by subcloning the BamHI fragments containing the aur1-3 and aur1-17 alleles onto cytoplasmic plasmids.

One of the mutations, aur1-3, has been shown to map genetically near one end of the gene (Loissen and Lacroute, 1979), between two artifically introduced BamHI sites (Falco, 1983). This led to the prediction that the mutation would lie within a specific enzyme restriction fragment. The appropriate TaqI-Aur1 fragment was isolated and sequenced. The aur1-3 mutation was found to be an A to T transversion at position 245 (Fig. 5), which changes the seventh codon from lysine to arginine. The autoradiogram of this sequencing gel is shown in Fig. 6. This confirms that the protein starts to the 5' side of this mutation and not at the next available methionine at position 299.

DISCUSSION

The yeast URA3 different laboratory strain

Fig. 6. DNA sequencing of the S. cerevisiae aur1-3 mutant. A minus-strand tail labeled with [α-32P]dCTP was prepared and the 3' end labeled with TaqI-Aur1 as described in Methodology. The plasmid was digested with BamHI, electrophoresed on a 5% polyacrylamide sequencing gel, and autoradiographed. The positions of the autoradiographic bands were determined by computer analysis. The mutagenic nucleotide appears to be an A to T transversion at position 245.

The yeast URA3 different laboratory strain

The yeast URA3 different laboratory strain
The yeast URA3 gene was isolated from two different laboratory strains of *S. cerevisiae*. Expression of the gene from the yeast strain FL100 in *E. coli* is completely dependent on transcription from bacterial promoters in the surrounding vector sequences. The gene as isolated from the yeast strain +D4 contains sequences which are sufficiently similar to the bacterial promoter to express the gene in *E. coli* in the absence of vector promoters. The expression of the gene from this promoter is not strong; even on the multicopy plasmid pBR322 the enzyme level amounts to only about a third of the level of the analogous gene on the *E. coli* chromosome. Both yeast genes still possess functional eukaryotic promoters (Chevallier et al., 1980; Rose et al., 1981). Thus expression of a yeast gene in *E. coli* is not an inherent property of yeast promoters even for those genes which are fortuitously expressed in *E. coli*.

The DNA sequence of the 5' region of the two genes differs in several places. Adjacent to the structural gene isolated from +D4 is a sequence which is fairly homologous to the consensus sequence for the *E. coli* promoter. The idea that this sequence is the promoter responsible for expression in *E. coli* is supported by the observation that one of the most conserved bases of the -10 region is different in the gene which is not expressed in *E. coli*. The observation that this base change affects expression of the gene in *E. coli* but not in yeast suggests that the essential promoter elements are not the same in yeast and *E. coli*.

We have isolated the homologous DNA fragment from strains bearing amber alleles of the *URA3* gene. Some of these DNA fragments express functionalOMP decarboxylase only in *E. coli* strains which carry an amber suppressor. Direct DNA sequencing of one of the amber alleles revealed that the mutation is a change in the seventh codon of the structural gene. The *ura3-3* allele was originally assumed to lie near the amino terminus of the structural gene because it showed the most extreme instability of the mRNA as compared to a series of amber mutations mapping throughout the *URA3* gene (Losson and Lapidot, 1979).

The unexpected finding of out-of-frame *ATG* codons preceding the apparent initiator for the *URA3* coding sequence deserves mention, especially since one of these, if expressed, would constitute an overlapping gene (Barrell et al., 1976; Shaw et al., 1978). Gene fusion experiments suggest that the initiator codon for this small peptide is both transcribed and
translated; these results and their implications are discussed elsewhere (Rose and Botstein, 1983). The existence of numerous apparently silent DNA polymorphisms among two different laboratory strains of yeast is a surprising result. It has long been known that different strains commonly contain restriction site polymorphisms (Petes and Botstein, 1977). In this case some 10 different changes appear in about 1100 bases for an average divergence of about 1%. Most of changes are clustered in the 200 bp to the 5' side of the structural gene. Presumably most of this region contains little essential information and therefore is free to change. Two of the mutations are in the coding sequence and one of these causes a change in the primary structure of the protein. The probable presence of such changes all over the genome must be considered when using strains which are not known to be isogenic. Such base changes would lead to mismatches in the heteroduplex DNA formed during recombination. The mismatches, particularly those involving small deletions might have serious effects on the formation, branch migration and resolution of heteroduplexes as well as on the polarity and extent of mismatch correction. Results obtained from studying gene conversions of known mutations might be biased by the presence of these silent accompanying mutations. In recent years, the advent of yeast transformation (Hinnen et al., 1978) has led to the use of E. coli-colony shuttle vectors for the cloning of genes directly into yeast (Nasmyth and Reed, 1980; Botstein and Davis, 1982). The URA3 gene has been widely used as a selectable marker in both E. coli and yeast partly because of the small size of the HindIII fragment containing it. Two plasmids which have been of particular use are YEp24 and YIp5 (Botstein et al., 1979). These plasmids contain the URA3 gene derived from the + D4 strain. With the reporting here of the sequence of the URA3 gene all of the components of these vectors have been sequenced (Sutcliffe, 1979; Hartley and Donelson, 1986; Broach, 1982). This should be of some use to those who wish to restrict map yeast genes which have been isolated on them. A detailed list of restriction sites contained within the DNA fragment is presented in Table 1 along with a list of the sites which do not appear.

## Table 1

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Sites</th>
<th>Enzyme</th>
<th>Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcI</td>
<td>551, 1056</td>
<td>HindII</td>
<td>731, 867</td>
</tr>
<tr>
<td>AdhI</td>
<td>1, 235, 285, 1078, 1167</td>
<td>HindIV</td>
<td>71, 1111</td>
</tr>
<tr>
<td>AdhII</td>
<td>380, 661, 662</td>
<td>ApaI</td>
<td>831, 1107</td>
</tr>
<tr>
<td>AvrII</td>
<td>1106</td>
<td>Aval</td>
<td>948</td>
</tr>
<tr>
<td>Aval</td>
<td>380</td>
<td>BglII</td>
<td>953, 152, 236, 489, 494 838, 742, 995, 927, 1008</td>
</tr>
<tr>
<td>AvrIII</td>
<td>1050</td>
<td>BstNI</td>
<td>437, 680, 390, 963</td>
</tr>
<tr>
<td>BglI</td>
<td>953, 1011</td>
<td>NcoI</td>
<td>830, 1106</td>
</tr>
<tr>
<td>BglII</td>
<td>119, 720, 957</td>
<td>NcoI</td>
<td>432</td>
</tr>
<tr>
<td>EcoP1</td>
<td>893</td>
<td>PvuII</td>
<td>205</td>
</tr>
<tr>
<td>EcoRI*</td>
<td>61, 387, 503, 754, 985, 1111</td>
<td>RsaI</td>
<td>342, 475, 538, 726</td>
</tr>
<tr>
<td>EcoRII</td>
<td>605</td>
<td>SacI</td>
<td>906</td>
</tr>
<tr>
<td>EcoRV</td>
<td>413</td>
<td>SmaI</td>
<td>953, 1011</td>
</tr>
<tr>
<td>Fnu4HI</td>
<td>254, 278, 451, 632, 1014</td>
<td>SmaI</td>
<td>1306</td>
</tr>
<tr>
<td>HindI</td>
<td>661</td>
<td>SstI</td>
<td>661</td>
</tr>
<tr>
<td>HindII</td>
<td>602, 662, 1016</td>
<td>SpeI</td>
<td>230, 492</td>
</tr>
<tr>
<td>HpaI</td>
<td>857</td>
<td>XhoI</td>
<td>905</td>
</tr>
<tr>
<td>RsaI</td>
<td>119</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENTS

We wish to thank François Lacroute for sharing results prior to publication, Marion Carlson for helpful advice and Michael Lichten for stimulating discussion. This work was supported by grants from the American Cancer Society (MVF90) and the National Institutes of Health (GM21523 and GM18973). M.R. was supported by a National Science Foundation Fellowship and an N.I.H. training grant (GM07287).

REFERENCES


Lacroute, F.: Regulation of purine biosynthesis in Saccharo-

Liebman, I., Kornberg, A. and Simonson, E.S.: Eukaryotic
Communicated by G.R. Fink.

INTRODUCTION
Secretin is a 27-amino acid peptide consisting of two domains, a headpiece and a tailpiece, which are connected by a reverse-phase-heated, tryptic gels, or immunoblotting.

SECRETIN

**To whom reprint addressed at Laboratoires, Takada-Gun, Hoshi**

Abbreviations: bp, BstEII-ApaI; kb, kilobase; polypeptides are listed in the order of their molecular weight.

0378-1119/84/00503-001$02.00/0