Volume 14 Number 24 1986 Nucleic Acids Research

The *ILV5* gene of *Saccharomyces cerevisiae* is highly expressed

Jens G.Litske Petersen and Steen Holmberg

Department of Physiology, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark

Received 14 October 1986; Accepted 17 November 1986

ABSTRACT

The nucleotide sequence of the yeast ILV5 gene, which codes for the branched-chain amino acid biosynthesis enzyme acetohydroxyacid reductoisomerase, has been determined. The ILV5 coding region is 1,185 nucleotides, corresponding to a polypeptide with a molecular weight of 44,280. Transcription of the ILV5 mRNA initiates at position -81 upstream from the ATG translation start codon and terminates between 218 and 222 bases downstream from the stop codon. Consensus sequences have been identified for initiation and termination of transcription, and for general control of amino acid biosynthesis, as well as repression by leucine. The ILV5 gene is regulated slightly by general amino acid control. Codon usage of the ILV5 gene has the strong bias observed in yeast genes that are highly expressed. In agreement with this, the reductoisomerase monomer, with an apparent molecular weight of 40,000, has been identified in an SDS polyacrylamide gel pattern of total soluble yeast proteins as a gene dosage dependent band.

INTRODUCTION

Acetohydroxyacid reductoisomerase (EC.1.1.1.86) catalyses the second common step in the parallel biosynthesis of isoleucine and valine. In the presence of NADPH and magnesium ions it converts α-aceto-α-hydroxybutyrate to α,β-dihydroxyacid-β-methylvalerate and α-acetolactate to α,β-dihydroxyisovalerate. It has been purified to homogeneity from Salmonella typhimurium (1,2) and identified in several other organisms, including yeast (3,4,5,6,7). In Salmonella, the enzyme has a molecular weight of 220,000 and consists of four identical subunits (2,8). The synthesis of reductoisomerase in bacteria is induced by either substrate, and differs in this respect from the other isoleucine-valine enzymes which are multivalently repressed by isoleucine, valine and leucine (9,10). The structural gene for the reductoisomerase of Escherichia coli (ilvC) has been cloned (11). In Saccharomyces cerevisiae, the reductoisomerase is coded for by a single gene located on chromosome XII (12,13). The gene, designated ILV5, has recently been cloned (14). The level of reductoisomerase in the cell is regulated, since the simultaneous presence of all three branched-chain amino acids in the growth medium

appears to cause repression of the synthesis of the enzyme (15,16, 17). Reductoisomerase is located in the mitochondria together with other isoleucine-valine biosynthesis enzymes (18).

In order to gain further insight into the expression of the ILV5 gene and the structure of its regulatory elements, we have sequenced the gene and located the points of transcription initiation and termination. In addition, we show that the gene is highly expressed in wild-type yeast cells and is probably regulated through the general amino acid control system (19,20).

MATERIALS AND METHODS

Chemicals and enzymes

All chemicals were analytical grade. 5-fluoroorotic acid was supplied by SCM Specialty Chemicals, Gainesville, Florida. Enzymes were purchased from Boehringer (Mannheim, Germany), Bethesda Research Laboratories (Bethesda, Maryland, USA) and New England Biolabs (Beverly, Massachusetts, USA). The 15-mer 'universal sequencing primer' and the 16-mer 'hybridization primer' were from Boehringer. An 18-mer deoxyoligonucleotide used in primer extension was synthesized on an Applied Biosystems 380A DNA Synthesiser. E. coli tRNA (type XX) was from Sigma (St. Louis, Missouri, USA). Radionucleotides were from New England Nuclear (Massachusetts, USA).

DNA sequencing

Standard procedures using the single-stranded bacteriophage M13 vector mp19 and dideoxynucleotide sequencing were employed (21,22,23). Programs for nucleotide sequence data handling were kindly supplied by Dr. R. Staden, MRC, Cambridge, England.

Preparation of plasmids, DNA fragments and yeast RNA preparation

Methods for the preparation of plasmids and total yeast RNA have been described (24). Restriction fragments for use in M13 cloning and nuclease S1 mapping were isolated by agarose gel electrophoresis and electroelution onto a dialysis membrane (25), followed by purification in Elutip-d columns (Schleicher and Schüell).

Plasmids, strains, genetic constructions and media

Plasmid pE17-11/B1 (14) was kindly provided by J. Polaina. It consists of YRp17 (pBR322-TRP1 ARS1-URA3) (26) with the yeast ILV5 gene on a 3.4 kb fragment produced by partial digestion of yeast DNA with MboI. The fragment is inserted in the BamHI site (Figure 1A). Plasmid constructions followed standard procedures (27).

The S. cerevisiae strains were X2180-1A (MATa SUC2 mal mel gal2 CUP1),

C82-1613 (MATa ilv5-1 ura3 trp1 hom7) (14) and M1-2B (MATα ura3-52 trp1) (28). For increased expression of acetohydroxyacid reductoisomerase in yeast, strain M1-2B was transformed with plasmid pEIv5, yielding strain C86-P506. The plasmid consists of YEp24 (pBR322-2μ-URA3) (26) with the 3.4 kb yeast DNA fragment containing the ILV5 gene in the BamHI site. A yeast strain with an ilv5 deletion was constructed as follows: The 3.4 kb yeast insert with ILV5 was transferred from pE17-11/B1 to the integration vector YIp5 (pBR322-URA3) (26) and 1.4 kb of the insert, mainly ILV5 coding region, was removed by cleavage with KpnI, followed by religation. The resulting plasmid with the ilv5 deletion was integrated into yeast strain M1-2B, whereupon loop-out and loss of the ILV5 gene together with the vector was selected for (29). One of the resulting isolates, C86-2485 (MATα ura3-52 trp1 ilv5-Δ1), was used in this study. E. coli strains employed were HB101 (27) and JM101 (30).

Minimal ammonium medium (SD) for yeast was composed of 6.7 g Difco Bacto Yeast Nitrogen Base without Amino Acids and 20 g glucose, buffered with 10 g succinic acid and sodium hydroxide to pH 5.8, for 1 liter medium. Minimal medium with proline as nitrogen source (SDPro) was prepared in the same way, but using Yeast Nitrogen Base without Ammonium and Amino Acids and 2 g proline per liter medium. This medium was adjusted to pH 4.5. Nutrient supplements were added at the concentrations described earlier (13).

Northern analysis

Twenty μg total yeast RNA, isolated from strain X2180-1A after growth in SD medium, was separated by electrophoresis in a 1.4% agarose gel containing formaldehyde (27), followed by transfer to nitrocellulose filter (Millipore HAWP 00010). Approximately 1 μg of single-stranded DNA of M13 with a yeast DNA insert was the template for labeling with (α-^{32}P)-dATP using DNA polymerase I large fragment and either the 16-mer hybridization primer or the 15-mer sequencing primer. In the latter case the probe was denatured. Hybridization to the immobilized RNA was carried out in 50% formamide at 42°C according to Alwine et al. (31).

Primer extension

To map the 5' end of ILV5 mRNA by primer extension (32), 160 pmoles of a synthetic deoxynucleotide d(5'GGCTTGAGTTCTCAACAT3') (the first 6 codons of the sense strand of the ILV5 coding region) were labeled at the 5' end with 150 μCi (γ-^{32}P)-ATP (27) and purified by electrophoresis in a 16% polyacrylamide/urea sequencing gel to yield a total of $1.9 \cdot 10^7$ dpm of labeled oligonucleotide. Three aliquots of total yeast RNA and one aliquot of E. coli tRNA (40 μg each) were each mixed with $4 \cdot 10^6$ dpm labeled primer in 70 μl of

70 mM Tris-HCl, pH 8.3, 70 mM KCl, 14 mM $MgCl_2$ and heated at 70°C for 20 min. The samples were brought to a volume of 100 µl by the addition of DTT to a final concentration of 10 mM and to 50 µM for each of the four deoxynucleotide triphosphates. After 5 min preincubation at 23, 43 or 50°C for yeast RNA and 43°C for the control, the samples were further incubated at the same temperatures for 30 min with 100 U avian myeloblastosis virus reverse transcriptase (Boehringer). After ethanol precipitation, the dried samples were incubated with 5 µl 0.1 M NaOH, 5 mM EDTA at 30°C for 4 hours and analysed on a 8% sequencing gel followed by autoradiography. To determine the size of elongation products and to test the quality of the labeled primer, dideoxy sequencing reactions with the omission of (α-^{32}P)-dATP (personal communication by Henrik Dalbøge Andersen, Nordisk Gentofte, Denmark) were carried out after annealing $4 \cdot 10^5$ dpm primer to 1 µg of single-stranded M13mp19-Sc4 DNA (containing the mRNA-identical strand of most of the ILV5 gene and its promoter region, see Figure 1C).

Nuclease S1 protection mapping

Nuclease S1 mapping of mRNA ends was carried out with total yeast RNA and double-stranded probes according to Berk and Sharp (33) with the modifications described earlier (24). Mapping of the transcription start point employed a 399 bp DdeI restriction fragment (Figure 1B): Approximately 1 pmole was dephosphorylated with 0.5 U calf intestinal alkaline phosphatase (Boehringer) (27) and the 5' protruding ends labeled with (γ-^{32}P)-ATP and T4 DNA polynucleotide kinase. Mapping of transcription termination employed approximately 1 pmole of a 434 bp EcoRI-DdeI restriction fragment (Figure 1B): The mRNA complementary strand was 3' end-labeled at the EcoRI site with 50 µCi (α-^{32}P)-dATP and Klenow DNA polymerase I (27). The other strand was labeled at the DdeI site using 50 µCi (α-^{32}P)-dGTP and 0.1 mM each of non-radioactive dATP and dTTP.

Determination of reductoisomerase activity and α-acetohydroxyacid levels

Strain X2180-1A was grown exponentially at 30°C in 200 ml minimal medium for at least 16 hours. To starve for tryptophan, different amounts of 5-methyltryptophan (Sigma) were added (from a 10 mM stock solution in water) at a cell density of 10^6 per ml. The cells were harvested at 10^7 cells per ml, washed once with 0.5 mg/ml cycloheximide and permeabilized by Triton X-100 (34,35). Reductoisomerase activity was measured spectrophotometrically with α-acetolactate as the substrate (16). Ethyl-α-aceto-α-acetoxypropionate was kindly provided by E.L. Kline, Edinboro State College, Pennsylvania. Total protein was measured with the Bio-Rad Protein Assay (Bio-Rad Laboratories,

Richmond, California) after boiling of the cell samples for 5 min in 0.5 M NaOH. Bovine serum albumin (Sigma) was used as standard.

SDS polyacrylamide gel electrophoresis of yeast proteins

An extract of total soluble proteins from yeast was prepared by resuspending a cell pellet in 3-4 volumes of cold 10 mM Tris-HCl pH 8.0, 1 mM EDTA. The cells were broken with 0.45 mm$^{\phi}$ glass beads for 90 seconds at 2-4°C in a Braun Melsungen MSK2 Homogenizer. Insoluble material was removed by centrifugation for 10 min in an Eppendorf centrifuge. SDS polyacrylamide gel electrophoresis was carried out according to Laemmli (36) and proteins were stained with Coomassie Brilliant Blue R-250. Protein molecular weight standards (high range) were from Gibco BRL (Eppenstein, West Germany).

RESULTS

Nucleotide sequence of the ILV5 gene

The molecular cloning of the ILV5 gene on a 3.4 kb DNA fragment from a partial MboI restriction digest of yeast DNA into the BamHI site of the vec-

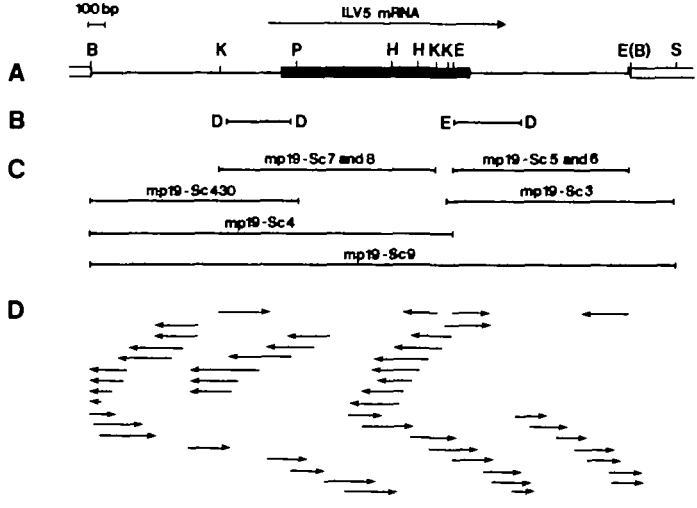

Figure 1. Restriction endonuclease map and sequence strategy of the ILV5 gene. A. Restriction endonuclease sites in the 3.4 kb yeast DNA fragment containing the ILV5 gene. The fragment was cloned from a partial MboI digest of yeast DNA into the BamHI site of the vector YRp17 (14). B=BamHI, K=KpnI, P=PvuII, H=HindIII, E=EcoRI, S=SalI. The ILV5 coding region is indicated by a solid line. B. Position of probes used in S1 nuclease mapping, D= DdeI. C. Five restriction fragments were subcloned into M13-mp19 (Sc3 through Sc9) for sequencing by the dideoxy chain termination method. Also shown is the M13 clone mp19-Sc430 used for Northern analysis. D. Series of deletions were made in the replicative forms of the M13 clones mp19-Sc4 and mp19-Sc9 for further sequencing Arrows indicate individual sequence readings.

```
                BamHI/MboI
-1191           G GATCCTGTTA AAACCTCTAG TGGAGTAGTA GATGTAATCA ATGAAGCGGA AGCCAAAAGA CCAGAGTAGA

                                                DdeI
-1120 GGCCTATAGA AGAAACTGCG ATACCTTTTG TGATGGCTAA ACAAACAGAC ATCTTTTTAT ATGTTTTTAC TTCTGTATAT

-1040 CGTGAAGTAG TAAGTGATAA GCGAATTTGG CTAAGAACGT TGTAAGTGAA CAAGGGACCT CTTTTGCCTT TCAAAAAAGG

                                                                HpaI
 -960 ATTAAATGGA GTTAATCATT GAGATTTAGT TTTCGTTAGA TTCTGTATCC CTAAATAACT CCCTTACCCG ACGGGAAGGC

 -880 ACAAAAGACT TGAATAATAG CAAACGGCCA GTAGCCAAGA CCAAATAATA CTAGAGTTAA CTGATGGTCT TAAACAGGCA

                      DdeI          AccI                                          DdeI
 -800 TTACGTGGTG AACTCCAAGA CCAATATACA AAATATCGAT AAGTTATTCT TGCCCACCAA TTTAAGGAGC CTACATCAGG

                                                DdeI
 -720 ACAGTAGTAC CATTCCTCAG AGAAGAGGTA TACATAACAA GAAAATCGCG TGAACACCTT ATATAACTTA GCCCGTTATT

 -640 GAGCTAAAAA ACCTTGCAAA ATTTCCTATG AATAAGAATA CTTCAGACGT GATAAAAATT TACTTTCTAA CTCTTCTCAC

                                                DdeI
 -560 GCTGCCCCTA TCTGTTCTTC CGCTCTACCG TGAGAAATAA AGCATCGAGT ACGGCAGTTC GCTGTCACTG AACTAAAACA

                      KpnI                                      DdeI
 -480 ATAAGGCTAG TTCGAATGAT GAACTTGCTT GCTGTCAAAC TTCTGAGTTG CCGCTGATGT GACACTGTGA CAATAAATTC

 -400 AAACCGGTTA TAGCGGTCTC CTCCGGTACC GGTTCTGCCA CCTCCAATAG AGCTCAGTAG GAGTCAGAAC CTCTGCGGTG

 -320 GCTGTCAGTG ACTCATCCGC GTTTCGTAAG TTGTGCGCGT GCACATTTCG CCCGTTCCCG CTCATCTTGC AGCAGGCGGA

 -240 AATTTTCATC ACGCTGTAGG ACGCAAAAAA AAAATAATTA ATCGTACAAG AATCTTGGAA AAAAAATTGA AAAATTTTGT

 -160 ATAAAAGGGA TGACCTAACT TGACTCAATG GCTTTTACAC CCAGTATTTT CCCTTTCCTT GTTTGTTACA ATTATAGAAG

                                                MboI                    DdeI
  -80 CAAGACAAAA ACATATAGAC AACCTATTCC TAGGAGTTAT ATTTTTTTAC CCTACCAGCA ATATAAGTAA AAAATAAAAC

   +1 ATG TTG AGA ACT CAA GCC GCC AGA TTG ATC TGC AAC TCC CGT GTC ATC ACT GCT AAG AGA ACC
      Met Leu Arg Thr Gln Ala Ala Arg Leu Ile Cys Asn Ser Arg Val Ile Thr Ala Lys Arg Thr  21
                                                      PvuII                          MboI
  +64 TTT GCT TTG GCC ACC CGT GCT GCT GCT TAC AGC AGA CCA GCT GCC CGT TTC GTT AAG CCA ATG
      Phe Ala Leu Ala Thr Arg Ala Ala Ala Tyr Ser Arg Pro Ala Ala Arg Phe Val Lys Pro Met  42
                                                                      AccI
 +127 ATC ACT ACC CGT GGT TTG AAG CAA ATC AAC TTC GGT GGT ACT GTT GAA ACC GTC TAC GAA AGA
      Ile Thr Thr Arg Gly Leu Lys Gln Ile Asn Phe Gly Gly Thr Val Glu Thr Val Tyr Glu Arg  63
                                                                          MboI
 +190 GCT GAC TGG CCA AGA GAA AAG TTG TTG GAC TAC TTC AAG AAC GAC ACT TTT GCT TTG ATC GGT
      Ale Asp Trp Pro Arg Glu Lys Leu Leu Asp Tyr Phe Lys Asn Asp Thr Phe Ala Leu Ile Gly  84

 +253 TAC GGT TCC CAA GGT TAC GGT CAA GGT TTG AAC TTG AGA GAC AAC GGT TTG AAC GTT ATC ATT
      Tyr Gly Ser Gln Gly Tyr Gly Gln Gly Leu Asn Leu Arg Asp Asn Gly Leu Asn Val Ile Ile 105

 +316 GGT GTC CGT AAA GAT GGT GCT TCT TGG AAG GCT GCC ATC GAA GAC GGT TGG GTT CCA GGC AAG
      Gly Val Arg Lys Asp Gly Ala Ser Trp Lys Ala Ala Ile Glu Asp Gly Trp Val Pro Gly Lys 126

 +379 AAC TTG TTC ACT GTT GAA GAT GCT ATC AAG AGA GGT AGT TAC GTT ATG AAC TTG TTG TCC GAT
      Asn Leu Phe Thr Val Glu Asp Ala Ile Lys Arg Gly Ser Tyr Val Met Asn Leu Leu Ser Asp 147

 +442 GCC GCT CAA TCA GAA ACC TGG CCT GCT ATC AAG CCA TTG TTG ACC AAG GGT AAG ACT TTG TAC
      Ala Ala Gln Ser Glu Thr Trp Pro Ala Ile Lys Pro Leu Leu Thr Lys Gly Lys Thr Leu Tyr 168
                                                                                  DdeI
 +505 TTC TCC CAC GGT TTC TCC CCA GTC TTC AAG GAC TTG ACT CAC GTT GAA CCA CCA AAG GAC TTA
      Phe Ser His Gly Phe Ser Pro Val Phe Lys Asp Leu Thr His Val Glu Pro Pro Lys Asp Leu 189
                                                          MboI
 +568 GAT GTT ATC TTG GTT GCT CCA AAG GGT TCC GGT AGA ACT GTC AGA TCT TTG TTC AAG GAA GGT
      Asp Val Ile Leu Val Ala Pro Lys Gly Ser Gly Arg Thr Val Arg Ser Leu Phe Lys Glu Gly 210

 +631 CGT GGT ATT AAC TCT TCT TAC GCC GTC TGG AAC GAT GTC ACC GGT AAG GCT CAC GAA AAG GCC
      Arg Gly Ile Asn Ser Ser Tyr Ala Val Trp Asn Asp Val Thr Gly Lys Ala His Glu Lys Ala 231
       HindIII
 +694 CAA GCT TTG GCC GTT GCC ATT GGT TCC GGT TAC GTT TAC CAA ACC ACT TTC GAA AGA GAA GTC
      Gln Ala Leu Ala Val Ala Ile Gly Ser Gly Tyr Val Tyr Gln Thr Thr Phe Glu Arg Glu Val 252

 +757 AAC TCT GAC TTG TAC GGT GAA AGA GGT TGT TTA ATG GGT GGT ATC CAC GGT ATG TTC TTG GCT
      Asn Ser Asp Leu Tyr Gly Glu Arg Gly Cys Leu Met Gly Gly Ile His Gly Met Phe Leu Ala 273
```

```
                                                                HindIII
+820 CAA TAC GAC GTC TTG AGA GAA AAC GGT CAC TCC CCA TCT GAA GCT TTC AAC GAA ACC GTC GAA
     Gln Tyr Asp Val Leu Arg Glu Asn Gly His Ser Pro Ser Glu Ala Phe Asn Glu Thr Val Glu 294
                                  MboI
+883 GAA GCT ACC CAA TCT CTA TAC CCA TTG ATC GGT AAG TAC GGT ATG GAT TAC ATG TAC GAT GCT
     Glu Ala Thr Gln Ser Leu Tyr Pro Leu Ile Gly Lys Tyr Gly Met Asp Tyr Met Tyr Asp Ala 315
                                                 KpnI
+946 TGT TCC ACC ACC GCC AGA AGA GGT GCT TTG GAC TGG TAC CCA ATC TTC AAG AAT GCT TTG AAG
     Cys Ser Thr Thr Ala Arg Arg Gly Ala Leu Asp Trp Tyr Pro Ile Phe Lys Asn Ala Leu Lys 336
                                                      KpnI                MboI       EcoRI
+1009 CCT GTT TTC CAA GAC TTG TAC GAA TCT ACC AAG AAC GGT ACC GAA ACC AAG AGA TCT TTG GAA
      Pro Val Phe Gln Asp Leu Tyr Glu Ser Thr Lys Asn Gly Thr Glu Thr Lys Arg Ser Leu Glu 357

+1072 TTC AAC TCT CAA CCT GAC TAC AGA GAA AAG CTA GAA AAG GAA TTA GAC ACC ATC AGA AAC ATG
      Phe Asn Ser Gln Pro Asp Tyr Arg Glu Lys Leu Glu Lys Glu Leu Asp Thr Ile Arg Asn Met 378
                                                                    +1186
+1135 GAA ATC TGG AAG GTT GGT AAG GAA GTC AGA AAG TTG AGA CCA GAA AAC CAA TAAAGAGGAA
      Glu Ile Trp Lys Val Gly Lys Glu Val Arg Lys Leu Arg Pro Glu Asn Gln Stp              395

+1196 AATAATATCA AGTGCTGGAA ACTTTTTCTC TTGGAATTTT TGCAACATCA AGTCATAGTC AATTGAATTG ACCCAATTTC
        Stp
+1276 ACATTTAAGA TTTTTTTTTT TTCATCCGAC ATACATCTGT ACACTAGGAA GCCCTGTTTT TCTGAAGCAG CTTCAAATAT

+1356 ATATATTTTT TACATATTTA TTATGATTCA ATGAACAATC TAATTAAATC GAAAACAAGA ACCGAAACGC GAATAAATAA
                                                                          DdeI
+1436 TTTATTTAGA TGGTGACAAG TGTATAAGTC CTCATCGGGA CAGCTACGAT TTCTCTTTCG GTTTTGGCTG AGCTACTGGT

+1516 TGCTGTGACG CAGCGGCATT AGCGCGGCGT TATGAGCTAC CCTCGTGGCC TGAAAGATGG CGGGAATAAA GCGGAACTAA
                 DdeI
+1596 AAATTACTGA CTGAGCCATA TTGAGGTCAA TTTGTCAACT CGTCAAGTCA CGTTTGGTGG ACGGCCCCTT TCCAACGAAT

+1676 CGTATATACT AACATGCGCG CGCTTCCATA TATACACATA TACATATATA TATATATATA TATGTGTGCG TGTATGTGTA

+1756 CACCTGTATT TAATTTCCTT ACTCGCGGGT TTTTCTTTTT TCTCAATTCT TGGCTTCCTC TTTCTCGAGT ATATAATTTT
                                                                          +1898
+1836 TCAGGTAAAA TTTAGTACGA TAGTAAAATA CTTCTCGAAC TCGTCACATA TACGTGTACA TA ATG TCT GAA CCA
                                                 XbaI
+1910 GCT CAA AAG AAA CAA AAG GTT GCT AAC AAC TCT CTA GAA CAA TTG AAA GCC TCC GGC ACT GTC

+1973 GTT GTT GCC GAC ACT GGT GAT TTC GGC TCT ATT GCC AAG TTT CAA CCT CAA GAC TCC ACA ACT
                    MboI                                                  MboI
+2036 AAC CCA TCA TTG ATC TTG GCT GCT GCC AAG CAA CCA ACT TAC GCC AAG TTG ATC GAT GTT GCC

+2099 GTG GAA TAC GGT AAG AAG CAT GGT AAG ACC ACC GAA GAA CAA GTC GAA AAT GCT GTG GAC AGA
                    EcoRI          MboI
+2162 TTG TTA GTC GAA TTC GGT AAG GAG ATC  +2188
```

Figure 2. Nucleotide sequence of the ILV5 gene (mRNA-identical strand) and the deduced amino acid sequence of the ILV5 polypeptide. The A of the translation initiation codon is numbered +1 and the preceding G -1. The 1,185 nucleotide long open translational reading frame is terminated by a TAA stop codon at position +1186. At position +1898 is noted another open reading reading frame. The positions for transcription initiation and termination are marked with triangles. Several sequence elements of possible significance can be recognized outside the ILV5 coding region. A promoter-like sequence (TATAAAA) is underlined, and sequences which may signify possible control elements for transcription termination (61,62) are marked with overlining. Three TGACTC sequences (or its complement), implicated in general control of amino acid biosynthesis (48) are marked with wavy overlining. A putative leucine control sequence, TCCGGTACCGGTTC (cf.57), is present between positions -379 and -366.

tor YRp17 has been described by J. Polaina (14). A restriction map of the cloned fragment (in plasmid pE17/11-B1) is given in Figure 1A. Before the sequencing was initiated, the plasmid was cleaved with EcoRI, religated with

Figure 3. Northern analysis. RNA molecules transcribed from the ILV5 region were identified by hybridization of ^{32}P-labeled DNA to total yeast RNA separated in agarose gels. Lane 1: The probe (mp19-Sc430) contained labeled DNA from position +181 to -1191 of the mRNA-complementary strand of the ILV5 gene. Lane 2: Hybridization with mp19-Sc3 containing the mRNA-complementary strand from position +1044 (in the coding region) to position +2188. Messenger RNA sizes were determined from the positions of the large and small ribosomal RNAs.

T4 DNA ligase and used to transform Escherichia coli strain HB101. This resulted in a plasmid lacking the 1.05 kb EcoRI fragment of the DNA insert, but otherwise identical to pE17-11/B1. The deletion plasmid was unable to complement the ilv5 mutation of yeast strain C82-1613 (MATa ilv5-1 ura3 trp1 hom7). Most likely, the ILV5 coding region therefore spans the left EcoRI site in the fragment shown in Figure 1A.

Sequencing of the 3.4 kb yeast DNA fragment was carried out using the bacteriophage M13 and dideoxy chain termination method (21,22). Five fragments from single or double restriction enzyme digests of pE17/11-B1 were cloned into M13mp19 (single enzyme fragments in both orientations), giving clones mp19-Sc3 through mp19-Sc9, see Figure 1C. Two of the phage clones, mp19-Sc4 and mp19-Sc9, were used for creating two new sets of M13 clones from opposite strands by the method of Hong (23). Of the entire 3,379 base pair fragment, 71% was sequenced in the direction from right to left in Figure 1 and 91 % in the opposite direction. In regions where the sequence was obtained only in one direction, sequencing was carried out twice and almost always from different M13 phages. The sequence was established without ambiguity and found to contain one open translational reading frame of 1,185 nucleotides (Figure 2).

Transcription of the ILV5 gene

To determine the size of the ILV5 mRNA, ^{32}P-labeled single-stranded M13 probes with DNA from different parts of the sequenced region were hybridized to total yeast RNA blotted onto nitrocellulose filter after electrophoresis in a denaturing agarose gel (Figure 3). Probe mp19-Sc430 (Figure 1C) with the mRNA-complementary strand from position +181 to position -1191 labeled with

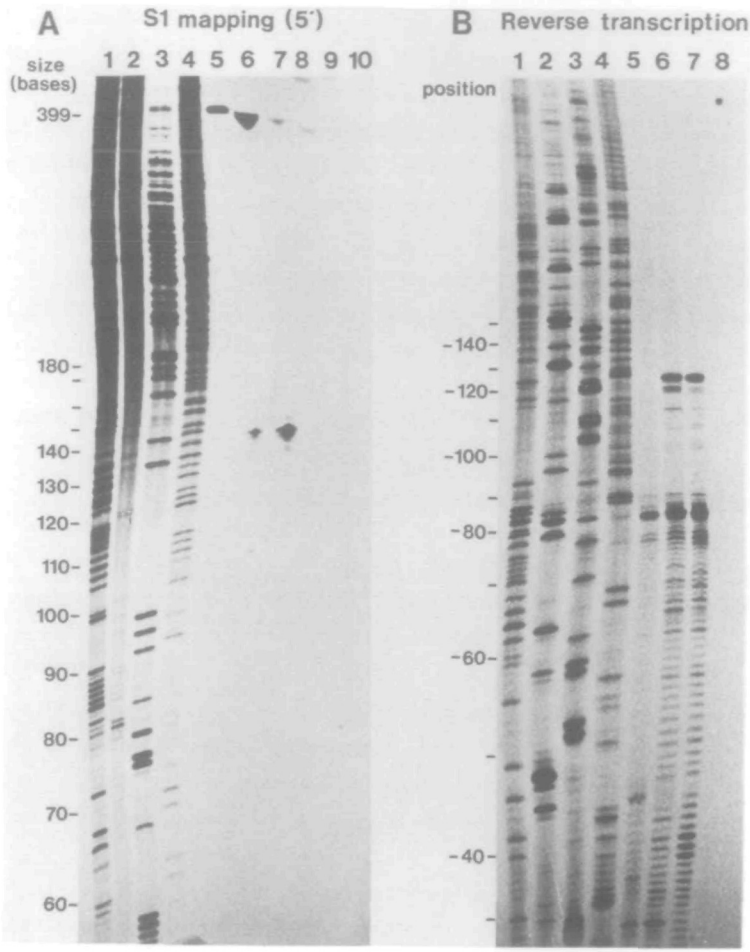

Figure 4. Mapping of the 5' terminus of the ILV5 mRNA. Figure 4A shows nuclease S1 mapping: Lanes 1-4 are T, C, G and A sequencing tracks serving as size markers. Lane 5 contains $3 \cdot 10^3$ dpm of 5' end-labeled 399 bp DdeI fragment (see Figure 1B). Lanes 6-9 show the result of hybridizing $3 \cdot 10^5$ dpm of probe to 36 µg of yeast RNA at 39, 43, 47 and 50°C, respectively, followed by nuclease S1 digestion. Lane 10: Hybridization to 40 µg E. coli tRNA at 43°C, followed by S1 digestion. The gel was exposed for 48 hours at -80°C with intensifying screen. Figure 4B shows primer extension mapping: Lanes 1-4 are the T, C, G and A dideoxy sequencing tracks of mp19-Sc4 using $2 \cdot 10^5$ dpm of the 5' end-labeled primer 5'-GGCTGGAGTTCTCAACAT. The positions of the sequence readings upstream from the ILV5 gene are indicated. Lanes 5-7 show the result of hybridizing $2 \cdot 10^6$ dpm of labeled primer to 40 µg total yeast RNA, followed by elongation with reverse transcriptase at 23, 43 and 50°C, respectively. Lane 8: Reverse transcription with E. coli tRNA. The gel was exposed for 24 days at room temperature.

^{32}P, hybridized to one RNA species with a size of about 1,500 ribonucleotides (lane 1). The same mRNA is hybridizing to mp19-Sc3, containing ^{32}P-labeling of the same strand from the KpnI site at position +1044 in the ILV5 coding region to position +2188, 1,003 nucleotides downstream from the TAA stop codon (lane 2). In addition, this probe also hybridizes to an RNA molecule of about 1,300 nucleotides, transcribed from a region containing the start of another open reading frame downstream from the ILV5 gene (Figure 2). There is no further detectable transcription from the sequenced region.

Northern analysis was also carried out with M13 clones which by sequencing were known to represent different parts outside the coding region. It was shown that transcription of the 1.5 kb ILV5 mRNA initiated downstream of position -150 and terminated between the 3' end of the coding region and 270 bp downstream from it (data not presented).

Nuclease S1 and primer extension mapping of the ILV5 mRNA

Nuclease S1 protection mapping (33) was carried out in order to localize precisely the ends of the ILV5 mRNA. To map the transcription start point, the 5' ends of the double-stranded DdeI restriction fragment from position -347 to +56 (Figure 1B) were labeled with ^{32}P and hybridized to total yeast RNA or E. coli tRNA in 80% formamide. The sizes of protected DNA fragments were then determined on a polyacrylamide gel using dideoxy sequencing of a known DNA sequence for size marking. The result is shown in Figure 4A. Hybridization to yeast RNA at 39 and 43°C (lanes 6 and 7) results in protection of 137-138 nucleotides. No DNA fragments are protected at higher temperatures or with E. coli tRNA. This places the 5' terminus of the ILV5 mRNA 137-138 bases upstream from nucleotide +56, i.e. at position -81 to -82 (see Figure 2). No other major transcription start points have been identified. The 5' end of the mRNA was also mapped by primer extension (32). A 5' end-labeled synthetic oligonucleotide ^{32}P-d(GGCTTGAGTTCTCAACAT), complementary to the sequence +1 to +18 in the ILV5 coding region, was annealed to yeast RNA and to E. coli tRNA. Subsequent reverse transcription was allowed at different temperatures to diminish any effect of folding of the 5' region on the elongation reaction. The sizes of elongation products were determined on a sequence gel by comparison to dideoxy sequencing reactions of the ILV5 upstream region using the labeled 18-mer as primer (Figure 4B, lanes 1-4). With yeast RNA, a doublet band is observed at all three temperatures at positions -80 and -81 upstream from the ILV5 coding region (lanes 5-7). In lanes 6 and 7 is also seen a band at position -124 to -125. Since a start point for transcription at this position was not observed in the S1 mapping

experiment described above, this is not regarded as an ILV5 mRNA 5' end. In conclusion, both S1 mapping and primer extension have located the 5' end of the ILV5 mRNA to position -81 ±1 upstream from the start of translation (Figure 2).

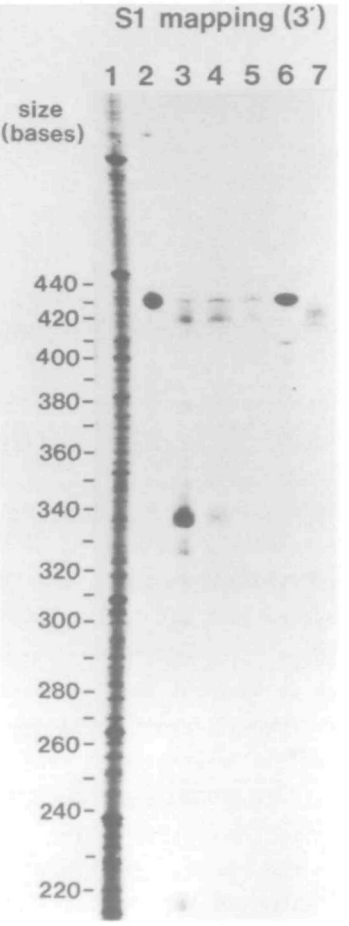

Figure 5. Nuclease S1 mapping of the 3' end of the ILV5 mRNA. Lane 1: The A-track of a known DNA sequence serving as size marker. Lane 2: A sample ($3 \cdot 10^3$ dpm) of the 434 bp EcoRI-DdeI fragment (see Figure 1) 3' end-labeled at the EcoRI site (positions +1071 and + 1072, i.e. mRNA-complementary strand). Lanes 3-4: $4 \cdot 10^5$ cpm of the same probe were hybridized to 36 µg of yeast RNA at 43 and 50°C, respectively, and treated with S1 nuclease. Lane 5: S1 experiment with 40 µg E. coli tRNA, 43°C. Lane 6: $3 \cdot 10^3$ cpm of the 434 bp fragment labeled at the DdeI site (position +1505 downstream from ILV5, i.e. RNA-identical strand). Lane 7: $4 \cdot 10^5$ dpm of the probe shown in lane 6 to 36 µg yeast RNA at 43°C and treating with S1 nuclease. The gel was exposed for 90 hours at room temperature.

Table I. Amino acid composition (Mol%) of the yeast ILV5 polypeptide and the purified acetohydroxyacid reductoisomerase of Salmonella typhimurium (calculated from Hofler et al.(2)).

Amino acid	ILV5 polypeptide	Salmonella reductoisomerase
Ala	8.4	10.9
Val	6.1	6.3
Leu	8.6	9.4
Ile	4.6	5.4
Pro	4.1	3.6
Phe	4.1	3.3
Trp	1.8	1.1
Met	2.0	3.3
Gly	8.6	9.4
Ser	5.6	4.6
Thr	6.6	4.2
Cys	0.8	1.0
Tyr	4.8	3.3
Asx	9.7	9.6
Glx	9.6	12.6
Lys	7.1	6.3
Arg	6.6	4.2
His	1.3	1.5
Total	100.4	100.0

The 3' end of the mRNA was mapped by S1 mapping using a double-stranded EcoRI-DdeI restriction fragment from position +1070 in the coding region to a position 322 nucleotides (+1505) downstream from the TAA stop codon as a probe. The mRNA-complementary strand was 3' end-labeled at the EcoRI site with (α-^{32}P)-dATP at positions +1072 and +1073. The opposite strand was labeled at the DdeI site by filling-out the G at position +1505 with (α-^{32}P)-dGTP. The result of this analysis is shown in Figure 5. After hybridazation, only the fragment labeled at the EcoRI site is protected (lanes 3 and 4). The size of the fragment is between 332 and 336 nucleotides, placing the polyadenylation site of the ILV5 mRNA this distance 3' from position +1072, i.e. between 218 and 222 downstream from the translation stop codon (+1403 to +1407) (see Figure 2).

From this analysis, the size of the ILV5 mRNA can be calculated to be between 1,483 and 1,490 ribonucleotides, in agreement with the size determined from Northern analysis (Figure 3).

Amino acid sequence of the ILV5 polypeptide

The amino acid sequence of the ILV5 polypeptide, as deduced from the DNA sequence, is shown in Figure 2. Initiation of translation is indicated to take place at the first AUG codon in the ILV5 mRNA. The polypeptide is composed of 395 amino acids with a molecular weight of 44,280. We have com-

Table II. Codon usage within the ILV5 gene.

First Letter	Second letter T	Second letter C	Second letter A	Second letter G	Third Letter
T	Phe 2	Ser 10	Tyr 0	Cys 2	T
	Phe 14	Ser 9	Tyr 19	Cys 1	C
	Leu 3	Ser 1	stp 0	stp 0	A
	Leu 29	Ser 0	stp 0	Trp 7	G
C	Leu 0	Pro 3	His 0	Arg 6	T
	Leu 0	Pro 0	His 5	Arg 0	C
	Leu 2	Pro 13	Gln 12	Arg 0	A
	Leu 0	Pro 0	Gln 0	Arg 0	G
A	Ile 3	Thr 10	Asn 1	Ser 1	T
	Ile 15	Thr 16	Asn 17	Ser 1	C
	Ile 0	Thr 0	Lys 1	Arg 20	A
	Met 8	Thr 0	Lys 27	Arg 0	G
G	Val 13	Ala 22	Asp 20	Gly 33	T
	Val 11	Ala 11	Asp 0	Gly 1	C
	Val 0	Ala 0	Glu 26	Gly 0	A
	Val 0	Ala 0	Glu 0	Gly 0	G

pared the amino acid composition of this polypeptide to that of purified reductoisomerase from Salmonella typhimurium (2) (Table I). The two proteins are very similar, as the difference in content between any amino acid is less than 2.5 Mol%. The molecular weight of the native reductoisomerase of S. typhimurium has been determined to be 220,000 (1,2). The enzyme is composed of 4 identical subunits, determined to be 57,000 by sedimentation equilibrium. In acrylamide gels, molecular weights ranging from 46,500 to 52,000 were found (2), i.e. similar to the calculated molecular weight of the ILV5 polypeptide. The size also agrees with the maximum molecular weight estimate of E. coli reductoisomerase of 50,000, calculated from the cloned ilvC gene (11).

The yeast reductoisomerase is found in the mitochondria together with at least three of the other isoleucine-valine enzymes (18). The amino-terminal region of the ILV5 polypeptide is basic, a feature that has been noted previously for proteins to be transported into the mitochondria (37). Of the first 59 amino acids, there are 8 arginines and 3 lysines, and no acidic residues. Another feature of the ILV5 polypeptide is a very hydrophilic carboxy-terminal region. Of the 33 carboxy-terminal amino acids, 52% are either acidic or basic (see Figure 2). In a computer analysis of the hydrophobicity along the ILV5 polypeptide chain according to Kyte and Doolittle (38), it appeared to lack membrane-associating regions.

Codon usage and expression of the ILV5 gene

Strong codon bias is usually seen for highly expressed genes in yeast

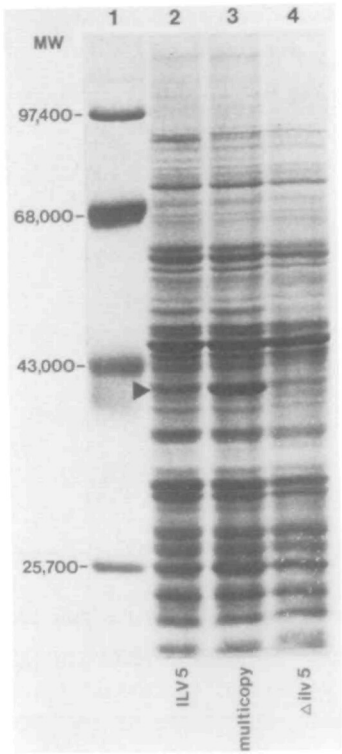

Figure 6. Identification of acetohydroxyacid reductoisomerase in an 8% SDS polyacrylamide gel of total soluble yeast proteins. Lane 1: Molecular weight markers (phosphorylase b 97,400, bovine serum albumin 68,000, ovalbumin 43,000, α-chymotrypsinogen 25,700). Lane 2: Strain M1-2B (wild-type), Lane 3: Strain C86-P506 (M1-2B/pBR322-2μ-URA3-ILV5), Lane 4: Strain C86-2485 (ilv5-Δ1). The position of reductoisomerase, apparent molecular weight 40,000, is indicated by an arrow.

(39). The codon usage is listed in Table II. Interestingly, the codon bias index (39) of the ILV5 polypeptide is 0.89, i.e. comparable to that of alcohol dehydrogenase isoenzyme I. Of the 61 possible triplets, only 37 are used for ILV5. This is in contrast to ILV1 and ILV2, which use 57 and 59 codons, respectively (40,41).

The strong codon bias for ILV5 prompted us to identify the yeast reductoisomerase in an SDS polyacrylamide gel pattern of soluble yeast proteins. For this purpose, three yeast strains were compared. The wild-type strain was M1-2B (MATα ura3-52 trp1). Strain C86-2485 harbors a chromosomal deletion of the ILV5 gene, ilv5-Δ1, extending from position -376 to +1044, i.e.

Table III. Derepression of reductoisomerase in strain X2180-1A

Medium	Generation time (h)	Reductoisomerase nmol/min/mg protein				mean
SD	2.0	2.27	2.31	2.58		2.4
SD + 5-MT(0.5 mM)	2.6	2.49	2.32			2.4
SDPro	5.0	2.61	3.14	2.37	2.50	2.7
SDPro + 5-MT (0.05 mM)	9	3.49	4.14			3.8
SDPro + 5-MT (0.20 mM)	10	3.19	3.23			3.2

it lacks the promoter region and most of the coding region. C86-P506 is strain M1-2B transformed with plasmid pEIv5 (pBR322-2μ-URA3-ILV5) for overproduction of reductoisomerase. Strains M1-2B and C86-2485 were grown in synthetic complete medium, while C86-P506 was grown in the same medium lacking uracil. The cells were harvested at late exponential phase and total soluble protein extracts were prepared. These were analysed on an 8% polyacrylamide gel with SDS. The result is shown in Figure 6. About 50 stained protein bands with apparent molecular weights above 20,000 can be distinguished. One polypeptide in the extract from the wild-type strain, with an apparent molecular weight of 40,000 (lane 2), is missing in the ilv5 deletion strain (lane 4), but is one of the most abundant bands in the ILV5 multi-copy strain (lane 3), and is thereby identified as the yeast reductoisomerase subunit.

Derepression of reductoisomerase by amino acid starvation

A number of genes coding for enzymes involved in amino acid biosynthesis in yeast are under general amino acid control (19,20). When cells are starved for any one of several amino acids, expression of the genes is induced 2- to 10-fold due to an increase in transcription (42,43,44,45,46). Studies of HIS3 and HIS4 have defined a repeated sequence element, TGACTC, upstream from transcriptional initiation, which is essential for positive regulation in response to amino acid starvation (44,47,48,49).

Three TGACTC sequence elements (or its complement) are present upstream of the transcription initiation point in the ILV5 gene, namely at -340, -312 and -140, see Figure 2. The presence of these sequences suggests that the ILV5 gene is under general control. In order to test this hypothesis, the wild-type strain X2180-1A was grown under different amino acid starvation conditions, and reductoisomerase levels determined in permeabilized cells. The result of this analysis is shown in Table III. No derepression of the enzyme is observed when cells are starved for tryptophan by the addition of 5-methyltryptophan to ammonium minimal medium. However, the enzyme level

increases slightly, by a factor of about 1.4, when 5-methyltryptophan (0.05 mM) is added to proline minimal medium. Our interpretation is that the ILV5 gene is under general control, but that severe starvation conditions are required for even a small derepression.

DISCUSSION

We have shown that the ILV5 gene of yeast contains one contiguous open translational reading frame of 1,185 base pairs. There is one point of transcriptional initiation, located 81 nucleotides upstream from the start of translation (Figure 2). The transcription start point is very close to the sequence CAAG (at position -80), which is present in the promoter region of several other yeast genes (24,50,51,52,53). In higher eukaryotes, an A-T rich region with the canonical sequence TATA$\frac{A}{T}$A$\frac{A}{T}$ (the TATA element) has been suggested to play a role in iniation of transcription. It is usually located 26 to 34 bases upstream from the start of transcription (54). TATA-like sequences are often present upstream of yeast genes, but the distance to the point of transcription varies from 35 to 180 bases among the genes studied (55). In the ILV5 gene, a perfect TATA element is present at position -161, i.e. 80 nucleotides upstream from the transcription start point.

The presence of three TGACTC containing sequences (at -140, -312 and -340) 5' of the transcription start point of the ILV5 gene, suggests that it is under general amino acid control (44,47,48,49). However, the reductoisomerase level increased only slightly when cells were grown under severe amino acid starvation conditions (Table III). It is possible that the high basal level of the enzyme masks efficient general control.

As for the other isoleucine-valine enzymes, the reductoisomerase activity is lowered by the simultaneous addition of isoleucine, valine and leucine to minimal growth medium (15,16,17). The molecular mechanism for this apparent pathway-specific regulation is not clear (see ref. 56 for discussion). In leucine biosynthesis, Martinez-Arias et al. (57) have identified a GC-rich palindromic upstream element containing the sequence GCCGGAACCGGCTT, which is important for the specific leucine repression of LEU2. Such sequences have also been noted in the upstream regions of LEU1 (58) and LEU4 (59). A related sequence, TCCGGTACCGGTTC, is present upstream of ILV5 (from position -379, see Figure 2), and we have noted a similar element, GCCGGTACCGGCTT (at position -467), in the 5' flanking region of the ILV2 gene (41). It is therefore tempting to suggest that leucine control acting on isoleucine-valine genes play a role in 'multivalent repression'. The requirement for isoleucine and

valine in this repression could be to abolish a derepression response under general amino acid control by adding leucine alone, a growth condition which has been shown to result in isoleucine-valine limitation (60).

The polyadenylation site of the ILV5 mRNA has been located between 218 and 222 nucleotides downstream from the translational stop codon (positions +1403 and +1407). From a comparison of five yeast genes, Bennetzen and Hall (61) have suggested that the sequence TAAATAA$^{A}_{G}$, located 25-40 bp upstream from the polyadenylation site, is a transcription termination signal in yeast. Such a sequence, with a 6/8 fit, is located at position +1353. In the same region (at position +1362), we have noted the sequence TTTTTACA, which differs by only one nucleotide from the element TTTTTATA, which by Henikoff et al. (62) has been suggested to be involved in transcription termination in yeast. A similar sequence was noted 38 bases upstream from the termination of transcription in the ILV1 gene (24) and is also present about 20 nucleotides upstream from the termination point for the ILV2 mRNA (41). Zaret and Sherman (63) have identified a tripartite structure to be important for signaling termination of the CYC1 mRNA. Such a structure is not readily recognized in the 3' untranslated region of the ILV5 gene.

We have identified the acetohydroxyacid reductoisomerase subunit in a stained SDS polyacrylamide gel pattern of total soluble proteins from wild-type yeast cells. This correlates well with the codon usage, which is biased in a manner similar to that of highly expressed genes (39). The bias is very similar to the ADC1 gene (alcohol dehydrogenase isoenzyme I), for which the mRNA constitutes 0.5-1% of total mRNA (39). Another feature of many highly expressed yeast genes is the presence of a C-T rich stretch in the spacer element between the TATA element and the point of transcription initiation (51). At the phosphoglycerate kinase gene, a CT-block of 20 bases is located 10 nucleotides upstream from the start of transcription (51). A sequence of 21 nucleotides, of which 19 are C or T (and 2 are G's), is located between positions -114 and -94, i.e. separated by 12 bases from the start of transcription in the ILV5 gene.

The enzymatic properties of reductoisomerase from yeast and other organisms (1,2,3,4,5,6,7,8), as well as the similarities in molecular weight and in amino acid composition between the ILV5 polypeptide and the purified enzyme of Salmonella typhimurium, suggest that the reductoisomerase is highly conserved among different organisms. From purification of the enzyme of S. typhimurium (2) and from Neurospora (5) it can be estimated that the enzyme constitutes about 0.5-2% of total soluble protein. Our results point

to a similar level in yeast. This is at least one order of magnitude higher than usually found for amino acid biosynthesis enzymes. Hofler et al. (2) have suggested that the high level of reductoisomerase in S. typhimurium could be necessary to compensate for the poor substrate turn-over rates of the enzyme. The same might be the case in yeast.

Many mitochondrial proteins are encoded in the nucleus, synthesized on free ribosomes in the cytoplasm (64) and transported into the mitochondria post-translationally. Usually, these proteins are synthesized as longer precursor polypeptides with amino-terminal presequences which are cleaved off upon import into mitochondria (37). It has been shown that cleavable presequences may contain signals, not only for mitochondrial targeting, but also for intramitochondrial sorting (65). The ILV5 polypeptide has a basic amino-terminal region typical of a matrix-targeting sequence. Among the first 57 amino acids, 11 are basic and none are acidic. The polypeptide does not contain any long stretch of uncharged amino acids resembling the 'stop-transfer sequence' seen in the presequence of intermembrane bound cytochrome c_1 or the 70 kD outer membrane polypeptide (65). We therefore suggest that reductoisomerase is located in the mitochondrial matrix. This may also be the case for threonine deaminase, the first enzyme in isoleucine biosynthesis, since the ILV1 polypeptide amino-terminal region is very similar (39). Concerning acetohydroxyacid synthase, Falco et al. (41) have noted a stretch of 10 uncharged amino acids in the basic amino-terminal region of the ILV2 polypeptide. This may not be sufficient to serve as a stop-transfer signal.

The calculated molecular weight of the ILV5 polypeptide is 44,280, i.e about 4,000 larger than the apparent molecular weight of reductoisomerase (Figure 6). Whether this difference is the result of cleavage of a presequence of about 40 amino acids is not known. In the protein pattern of the ILV5 over-producing yeast transformant (Figure 6, lane 3) there appears to be no protein band of about 44,000 molecular weight which could be the uncleaved precursor polypeptide. In the case of the ILV1 and ILV2 polypeptides, suggestive evidence for cleavable presequences have been presented (40,41).

ACKNOWLEDGEMENTS

Morten C. Kielland-Brandt, Torsten Nilsson-Tillgren and Claes Gjermansen are acknowledged for numerous stimulating discussions, their contribution to this work and for critically reading this manuscript. Diter von Wettstein

is thanked for his support, encouragement and careful reading of the manuscript. We thank Julio Polaina for the ILV5 gene, and Peter Niederberger, Zürich, for many discussions about general control of amino acid biosynthesis, and for suggesting the use of proline-containing medium in derepression studies. Gitte Bank and Kirsten Schultz are acknowledged for excellent technical assistance and Nina Rasmussen for drawing the figures. We are grateful to Ib Groth Clausen and Henrik Dalbøge Andersen, Nordisk Gentofte, Gentofte, Denmark for providing procedures for DNA sequencing with end-labeled primer. This work was supported by grants 16-3557.H-699 and 133/001-84154 from the Danish National Research Councils to D. von Wettstein.

REFERENCES

1. Arfin,S.M. and Umbarger,H.E. (1969) J. Biol. Chem. 244, 1118-1127
2. Hofler,J.G., Decedue,C.J., Luginbuhl,G.H., Reynolds,J.A. and Burns,R.O. (1975) J. Biol. Chem. 250, 877-882
3. Radhakrishnan,A.N., Wagner,R.P. and Snell,E.E. (1960) J. Biol. Chem. 235, 2322-2331
4. Umbarger,H.E., Brown,B. and Eyring,E.J. (1960) J. Biol. Chem. 235, 1425-1432
5. Kiritani,K. and Wagner,R.P. (1970) in Methods in Enzymology, Tabor,H. and Tabor,C.W. Eds., Vol 17A, pp. 745-750, Academic Press, New York and London
6. Satyanarayana,T. and Radhakrishnan,A.N. (1962) Biochim. Biophys. Acta 56, 197-199
7. Strassman,M., Shatton,J.B. and Weinhouse,S. (1960) J. Biol. Chem. 235, 700-705
8. Arfin,S. (1970) in Methods in Enzymology, Tabor,H. and Tabor,C.W. Eds., Vol 17A, pp. 751-755, Academic Press, New York and London
9. Ratzin,B., Arfin,S. and Umbarger,H.E. (1972) J. Bacteriol. 112, 131-141
10. Freundlich,M., Burns,R.O. and Umbarger,H.E. (1962) Proc. Nat. Acad. Sci. USA 48, 1804-1808
11. McCorkle,G.M., Leathers,T.D. and Umbarger,H.E. (1978) Proc. Nat Acad. Sci. USA 75, 89-93
12. Kakar,S.N. and Wagner,R.P. (1964) Genetics 49, 213-222
13. Petersen,J.G.L., Kielland-Brandt,M.C., Holmberg,S. and Nilsson-Tillgren, T. (1983) Carlsberg Res. Commun. 48, 21-34
14. Polaina,J. (1984) Carlsberg Res. Commun. 49, 577-584
15. Bussey,H. and Umbarger,H.E. (1969) J. Bacteriol. 98, 623-628
16. Magee,P.T. and Hereford,L.M. (1969) J. Bacteriol. 98, 857-862
17. Bollon,A.P. and Magee,P.T. (1973), J. Bacteriol. 113, 1333-1344
18. Ryan,E.D. and Kohlhaw,G.B. (1974) J. Bacteriol. 120, 631-637
19. Schürch,A., Miozzari,G. and Hütter,R. (1974) J. Bacteriol. 117, 1131-1140
20. Wolfner,M., Yep,D., Messenguy,F. and Fink,G.R. (1975) J. Mol. Biol. 96, 273-290
21. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Nat. Acad. Sci. USA 74, 5463-5467
22. Sanger,F., Coulson,A.R., Barrell,G., Smith,A.J.H. and Roe,B.A. (1980) J. Mol. Biol. 143, 161-178
23. Hong,G.F. (1982) J. Mol. Biol. 158, 539-549
24. Holmberg,S., Kielland-Brandt,M.C., Nilsson-Tillgren,T. and Petersen,J.G. L. (1985) Carlsberg Res. Commun. 50, 163-178

25. Yang,R.C-A., Lis,J. and Wu,R. (1979) in Methods in Enzymology, Wu,R. Ed., Vol 68, pp. 176-182, Academic Press, New York and London
26. Botstein,D. and Davis,R.W. (1982) in The Molecular Biology of the Yeast Saccharomyces. Metabolism and Gene expression. Strathern,J.N., Jones, E.W. and Broach,J.R. Eds., Vol 11B, pp. 607-636, Cold Spring Harbor Laboratory, New York
27. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, New York
28. StJohn,T.P., Scherer,S., McDonell,M.W. and Davis,R.W. (1981) J. Mol. Biol. 152, 317-334
29. Boeke,J.D., LaCroute,F. and Fink,G.R. (1984) Mol. Gen. Genet. 197, 345-346
30. Messing,J. (1979) in Recombination DNA Tech. Bull., NIH publication, No. 79-99, 2, 43-48
31. Alwine,J.C., Kemp,D.J. and Stark,G.R. (1977) Proc. Nat. Acad. Sci. USA 74, 5350-5354
32. Baker,C.C. and Ziff,E.B. (1981) J. Mol. Biol. 149, 189-221
33. Berk,A.J. and Sharp,P.A. (1978) Proc. Nat. Acad. Sci. USA 75, 1274-1278
34. Miozzari,G., Niederberger,P. and Hütter,R. (1978) J. Bacteriol. 134, 48-59
35. Miozzari,G.F., Niederberger, P. and Hütter, R. (1978) Anal. Biochem. 90, 220-233
36. Laemmli,U.K. (1970) Nature 227, 680-685
37. Hay,R., Böhni,P. and Gasser,S. (1984) Biochim. Biophys. Acta 779, 65-87
38. Kyte,J. and Dolittle,R.F. (1982) J. Mol. Biol. 157, 105-132
39. Bennetzen,J.L. and Hall,B.D. (1982) J. Biol. Chem. 257, 3026-3031
40. Kielland-Brandt,M.C., Holmberg,S., Petersen,J.G.L. and Nilsson-Tillgren, T. (1984) Carlsberg Res. Commun. 49, 567-575
41. Falco,S.C., Dumas,K.S. and Livak,K.J. (1985) Nucl. Acids Res. 13, 4011-4027
42. Struhl,K. and Davis,R.W. (1981) J. Mol. Biol. 152, 535-552
43. Zalkin,H. and Yanofsky,C. (1982) J. Biol. Chem. 257, 1491-1500
44. Donahue,T.F., Daves,R.S., Lucchini,G. and Fink,G.R. (1983) Cell 32, 89-98
45. Messenguy,F., Feller,A., Crabeel,M. and Pierand,A. (1983) EMBO J. 2, 1249-1254
46. Aebi,M., Furter,R., Prantl,F., Niederberger,P. and Hütter,R. (1984) Curr. Genet. 8, 165-172
47. Struhl,K. (1982) Nature 300, 284-287
48. Hinnebusch,A.G. and Fink,G.R. (1983) J. Biol. Chem. 258, 5238-5247
49. Hinnebusch,A.G., Lucchini,G. and Fink,G.R. (1985) Proc. Nat. Acad. Sci. USA 82, 498-502
50. Gallwitz,D., Perrin,F. and Seidel,R. (1981) Nucl. Acids Res. 9, 6339-6350
51. Dobson,M.J., Tuite,M.F., Roberts,N.A., Kingsman,A.J. and Kingsman,S.M. (1982) Nucl. Acids Res. 10, 2625-2637
52. Carlson,M., Taussig,R., Kustu,S. and Botstein,D. (1983) Mol. Cell. Biol. 3, 439-447
53. Holland,J.P., Labieniec,L., Swimmer,C. and Holland,M.J. (1983) J. Biol. Chem. 258, 5291-5299
54. Breatnach,R. and Chambon,P. (1981) Ann. Rev. Biochem. 50, 349-383
55. Sentenac,A. and Hall,B.D. (1982) in The Molecular Biology of the Yeast Saccharomyces. Metabolism and Gene Expression, Strathern,J.N., Jones, E.W. and Broach,J.R. Eds., Vol 11B, pp. 561-606, Cold Spring Harbor Laboratory, New York
56. Jones,E.W. and Fink,G.R. (1982) in The Molecular Biology of the Yeast Saccharomyces. Metabolism and Gene Expression, Strathern,J.N., Jones,

E.W. and Broach,J.R. Eds., Vol 11B, pp. 181-299, Cold Spring Harbor Laboratory, New York
57. Martinez-Arias,A., Yost,H.J. and Casadaban,M.J. (1984) Nature 307, 740-742
58. Hsu,Y.-P. and Schimmel,P. (1984) J. Biol. Chem. 259, 3714-3719
59. Beltzer,J.P., Chang,L.-F.L., Hinkkanen,A.E. and Kohlhaw,G.B. (1986) J. Biol. Chem. 261, 5160-5167
60. Niederberger,P., Miozzari,G. and Hütter,R. (1981) Mol. Cell. Biol. 1, 584-593
61. Bennetzen,J.L., Hall,B.D. (1982) J. Biol. Chem. 257, 3018-3025
62. Henikoff,S., Kelly,J.D. and Cohen,E.H. (1983) Cell 33, 607-614
63. Zaret,K.S. and Sherman,F. (1982) Cell 28, 563-573
64. Suissa,M. and Schatz,G. (1982) J. Biol. Chem. 257, 13048-13055
65. van Loon,A.P.G.M., Brändli,A.W., Schatz,G. (1986) Cell 44, 801-812