Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 8732–8736, November 1989
Biochemistry

Cloning and nucleotide sequence of the gene for protein X from *Saccharomyces cerevisiae*

(pyruvate dehydrogenase complex/polymerase chain reaction/evolutionary conservation)

ROBERT H. BEHAL, KAREN S. BROWNING, T. BRUCE HALL, AND LESTER J. REED

Clayton Foundation Biochemical Institute and Department of Chemistry, The University of Texas at Austin, Austin, TX 78712

Contributed by Lester J. Reed, August 31, 1989

ABSTRACT **The gene encoding the protein X component of the pyruvate dehydrogenase complex from *Saccharomyces cerevisiae* has been cloned and sequenced. A 487-base fragment of yeast genomic DNA encoding the amino-terminal region of protein X was amplified by the polymerase chain reaction using synthetic oligonucleotide primers based on amino-terminal and internal amino acid sequences. This DNA fragment was used as a probe to select two genomic DNA restriction fragments, which were cloned and sequenced. A 2.1-kilobase insert contains the complete sequence of the protein X gene. This insert has an open reading frame of 1230 nucleotides encoding a presequence of 30 amino acid residues and a mature protein of 380 amino acid residues (M_r, 42,052). Hybridization analysis showed that there is a single copy of the protein X gene and that the size of the mRNA is ≈1.5 kilobases. Comparison of the deduced amino acid sequences of yeast protein X and dihydrolipoamide acetyltransferase indicates that the two proteins evolved from a common ancestor. The amino-terminal part of protein X (residues 1–195) resembles the acetyltransferase, but the remainder is quite different. There is strong homology between protein X and the acetyltransferase in the amino-terminal region (residues 1–84) that corresponds to the putative lipoyl domain. Protein X lacks the highly conserved sequence His-Xaa-Xaa-Xaa-Asp-Gly near the carboxyl terminus, which is thought to be part of the active site of all dihydrolipoamide acyltransferases.**

Eukaryotic pyruvate dehydrogenase complexes are organized about a core consisting of the oligomeric dihydrolipoamide acetyltransferase (E_2; acetyl-CoA:dihydrolipoamide *S*-acetyltransferase, EC 2.3.1.12), around which are arranged multiple copies of pyruvate dehydrogenase (lipoamide) [E_1; pyruvate:lipoamide 2-oxidoreductase (decarboxylating and acceptor-acetylating), EC 1.2.4.1] and dihydrolipoamide dehydrogenase (E_3; dihydrolipoamide:NAD^+ oxidoreductase, EC 1.8.1.4.) bound by noncovalent bonds (1). The E_2 core has the appearance of a pentagonal dodecahedron in the electron microscope and apparently consists of 60 subunits arranged with octahedral (5 3 2) symmetry (2). The E_2 subunit has a multidomain structure (2–4). The amino-terminal segment contains one (yeast) or two (mammalian) lipoyl-bearing domains (5, 6), followed by a domain that is involved in binding E_1 and/or E_3, and then the catalytic inner-core domain. The domains are linked to each other by protease-sensitive segments that are rich in the conservatively substituted residues alanine, proline, serine, and threonine and in charged amino acid residues. These interdomain linker segments (hinge regions) are thought to provide flexibility to the lipoyl domains, facilitating active-site coupling within these multienzyme complexes.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

The mammalian and yeast pyruvate dehydrogenase complexes also contain small amounts of a protein of unknown function, designated protein X or component X, of apparent M_r ≈50,000 (7, 8). In the bovine kidney pyruvate dehydrogenase complex, about six molecules of protein X are associated with the E_2 core. Bovine protein X resembles bovine E_2 in that both proteins contain a lipoyl moiety that undergoes reductive acetylation and acetyl transfer reactions (8–11). There is also some sequence similarity in the amino-terminal segment of protein X and E_2 (12). However, immunological, peptide mapping, and limited proteolysis studies indicate that protein X is a distinct polypeptide (7, 8, 13). It is not clear whether protein X is bound to the E_2 core or whether it is an integral part of the E_2 core. The latter alternative implies a possible structural role for protein X in the assembly of the E_2 core. Evidence has been presented that in the pyruvate dehydrogenase complex from bovine kidney, protein X contributes to the binding and function of E_3, perhaps by facilitating the transfer of reducing equivalents to E_3 (14, 15). It should be noted that protein X has been detected only in eukaryotic pyruvate dehydrogenase complexes and not in other α-keto acid dehydrogenase complexes. To gain further insight into the nature and function of protein X, we have cloned and sequenced the gene encoding protein X from *Saccharomyces cerevisiae*.* In this report, we present these results and compare the deduced amino acid sequences of yeast E_2 and protein X.

MATERIALS AND METHODS

Materials. Restriction endonucleases and DNA-modifying enzymes were purchased from Bethesda Research Laboratories, New England Biolabs, Promega Biotech, and Boehringer-Mannheim. The GeneAmp kit was obtained from Perkin–Elmer/Cetus. A Sequenase DNA sequencing kit was obtained from United States Biochemical. Radiolabeled nucleotides were obtained from New England Nuclear. Immobilon poly(vinylidene difluoride) membrane was obtained from Millipore. Yeast genomic DNA was obtained from Clontech Laboratories. The GenClean kit was obtained from Bio 101 (La Jolla, CA).

Reductive Acetylation of the Lipoyl Moieties. The incubation mixture contained 5 mg of highly purified yeast pyruvate dehydrogenase complex (16), 0.2 mM thiamine diphosphate, 0.3 mM [2-^{14}C]pyruvate, 1 mM $MgCl_2$, 0.6 mM *N*-ethylmaleimide, 1 mM EDTA, and 50 mM potassium phosphate buffer (pH 7.0) in a total vol of 1 ml. The [2-^{14}C]pyruvate was added last. After incubation for 10 min at 4°C, the reaction was stopped by addition of 0.5 ml of 40% trichloroacetic acid. The precipitate was collected by centrifugation, washed with

Abbreviations: E_1, pyruvate dehydrogenase; E_2, dihydrolipoamide acetyltransferase; E_3, dihydrolipoamide dehydrogenase; PCR, polymerase chain reaction.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M28222).

ether, and dissolved in 1 ml of 1X Laemmli sample buffer (17) without heating. The complex was resolved on a SDS/polyacrylamide (12.5% acrylamide) minislab gel. The gel was stained with Coomassie brilliant blue, fixed, and dried. Autoradiography was performed with Kodak XAR-5 film.

Amino-Terminal and Peptide Sequence Analysis. Highly purified pyruvate dehydrogenase complex was subjected to SDS/polyacrylamide gel electrophoresis in 12.5% slab gels, with running buffer containing 45 mM Tris borate (pH 8.0), 2 mM EDTA, 0.1% SDS. The resolved protein bands were electroblotted onto a poly(vinylidene difluoride) membrane (18). Protein bands were visualized by staining with Coomassie brilliant blue, and the protein band corresponding to protein X was cut out. The membrane was washed in deionized water, air-dried, and subjected to automated sequence analysis with an Applied Biosystems model 470A gas-phase sequenator equipped with a model 120A on-line phenylthiohydantoin amino acid analyzer.

To obtain internal amino acid sequence, a gel slice containing ≈50 μg of protein X was applied to a second SDS/polyacrylamide slab gel in the presence of 0.1 μg of *Staphylococcus aureus* V8 protease for 45 min as described by Cleveland *et al.* (19). After electrophoresis, the peptides were electroblotted and sequenced as described above.

Preparation of Oligonucleotides. Oligonucleotide primers for polymerase chain reaction (PCR) and for DNA sequencing were synthesized on an Applied Biosystems model 381A DNA synthesizer and were used without further purification.

PCRs. PCR amplification was performed in a Perkin–Elmer/Cetus DNA thermal cycler for 30 cycles (1 min at 94°C, 2 min at 55°C, 3 min at 74°C; the extension step in the last cycle was increased to 10 min) with the GeneAmp kit according to the manufacturer's instructions.

Selected Genomic DNA Library Construction. Samples (5 μg) of yeast genomic DNA were digested with selected restriction enzymes overnight at 37°C. The digests were resolved by electrophoresis in a 0.8% agarose gel in TAE buffer (40 mM Tris acetate, pH 8.0/1 mM EDTA) and blotted onto nitrocellulose in 20× SSC (1× SSC is 0.15 M NaCl/15 mM sodium citrate, pH 7.0). Probes were prepared from either DNA fragments obtained by PCR amplification or from plasmid inserts and were labeled with [α-^{32}P]dCTP by random priming. Blots were prehybridized for 60 min at 65°C in 5× SSC/3× Denhardt's solution (1× Denhardt's solution is 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone)/0.3% SDS/50 μg of denatured salmon testes DNA per ml. The probe was denatured by heating to 100°C for 5 min, cooled in an ice bath, and added directly to the prehybridization mixture. Blots were allowed to hybridize overnight at room temperature and were washed for 1 hr in 2× SSC/0.2% SDS at 65°C with one change of wash buffer. Autoradiography was performed for 1–3 days at −70°C with Kodak XAR-5 film and an intensifying screen. Restriction fragments of 1–4 kilobases (kb) that hybridized to the probe were identified. To obtain restriction fragments for cloning, 30 μg of yeast genomic DNA was digested with the appropriate restriction enzyme(s), and the restriction fragments were resolved by electrophoresis on a 14-cm-long 0.8% agarose gel. DNA was visualized by sparing use of UV light, the band of interest was excised with a razor blade, and the DNA was purified with GeneClean. The plasmid vector Bluescript (Stratagene) was digested with the appropriate restriction enzyme(s), treated with calf intestinal alkaline phosphatase, if necessary, and gel purified. Restriction fragments were ligated into the prepared Bluescript vector and the resulting plasmid was used to transform *Escherichia coli* strain XL1-B (Stratagene). Positive colonies were identified, under the conditions described above, by colony hybridization to the appropriate probes.

DNA Sequencing. DNA fragments obtained by PCR amplification were purified essentially as described (20). Double-stranded DNA from PCR amplification or from a plasmid was sequenced with Sequenase according to the manufacturer's instructions, with the following modifications. DNA templates from PCR or plasmid DNA were mixed with sequencing primer in an equimolar ratio, heated to 100°C for 5 min, frozen immediately in a dry ice/ethanol bath, and annealed at room temperature for 10 min prior to sequencing.

RNA Blot Analysis. A 30-μg sample of total RNA from *S. cerevisiae* strain 20B-12 was denatured with formaldehyde, fractionated in a 0.8% agarose gel, and blotted onto nitrocellulose as described (21). This blot was analyzed with a probe prepared by random-primer labeling of an 891-base DNA fragment of the coding region of clone XX1, using hybridization conditions as described (21).

Chromosome Analysis. A dried gel containing yeast chromosomes resolved by pulsed-field gel electrophoresis was obtained from Clontech Laboratories. The gel was hybridized with the 891-base DNA probe according to the manufacturer's instructions.

Computer Analysis. Beckman Microgenie programs, versions 5 and 6, were used to analyze the DNA sequence data. Synthetic oligonucleotide primers were designed by using the yeast codon preference data in the oligonucleotide design program BIGprobe (obtained from the Genetics Software Center, University of Arizona).

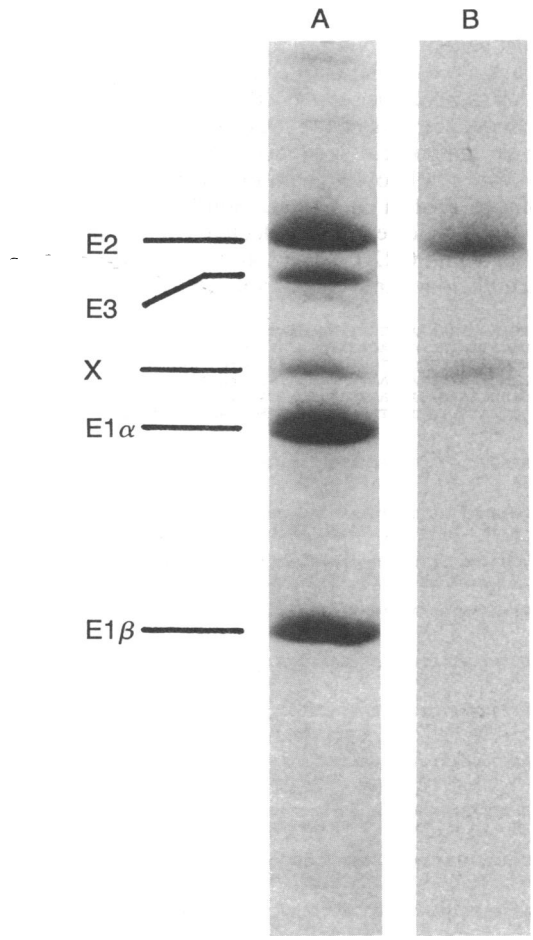

FIG. 1. Reductive acetylation of the lipoyl moieties in E_2 and protein X. Reductive acetylation in the presence of [2-^{14}C]pyruvate was performed as described. Lane A, pattern obtained by SDS/polyacrylamide gel electrophoresis of the labeled pyruvate dehydrogenase complex after staining with Coomassie brilliant blue; lane B, autoradiogram obtained with the dried gel.

RESULTS AND DISCUSSION

Reductive Acetylation of the Lipoyl Moieties in E_2 and Protein X. In the presence of E_1, Mg^{2+}, thiamine diphosphate and [2-^{14}C]pyruvate, the lipoyl moieties of the E_2 and protein X components of the bovine pyruvate dehydrogenase complex are acetylated (8, 10, 11). Comparable results were obtained with the yeast pyruvate dehydrogenase complex (Fig. 1). Two radioactive bands corresponding to E_2 and protein X were detected on an autoradiogram after resolution of the acetylated complex by SDS/polyacrylamide minislab gel electrophoresis.

Amino Acid Sequence Analysis. The amino-terminal sequence of protein X isolated from yeast pyruvate dehydrogenase complex was determined to be Ala-Val-Lys-Thr-Phe-Ser-Met-Pro-Ala-Met-Ser-Pro-Thr-Met-Glu-Lys-Gly-Gly-Ile-Val-Ser-Xaa-Lys-(Met or Tyr)-Xaa-Val-Gly-Glu-Pro-(Phe or Asp), where Xaa indicates residues that could not be identified. An internal sequence, obtained from a peptide released by *S. aureus* V8 protease, was Ile-Xaa-Pro-Ser-Gly-Ser-Asn-Gly-Arg-Leu-Leu-Lys-Gly-Asp-Val.

PCR Amplification of a Genomic DNA Fragment Encoding a Segment of Protein X. The degenerate oligonucleotide primer NTX-ol (ATGCC$^{T}_{A}$GC$^{T}_{C}$ATGTC$^{T}_{C}$CC$^{T}_{A}$AC$^{T}_{C}$ATG-GA) corresponds to amino acids 7–15 of the amino-terminal sequence. Primer XV-1-rol-1 (CTACCGTTAGAACCAGAT-GG) is an antisense nondegenerate oligonucleotide corresponding to amino acids 9–3 of the internal peptide sequence. Using yeast genomic DNA as a template, a DNA fragment of ≈500 base pairs, designated PCRX1, was generated by PCR. This DNA fragment was sequenced directly by using the PCR primers. Sequence analysis revealed an open reading frame encoding an amino acid sequence that exhibited extensive similarity with the lipoyl domain of yeast E_2 (5).

Isolation of Protein X Clones. Southern blots of yeast genomic DNA digested with selected restriction enzymes were probed with PCRX1, and a 2.4-kb *Hin*dIII fragment was identified and cloned into Bluescript. This insert, designated H1, was partially sequenced and was found to contain the experimentally determined amino-terminal and internal amino acid sequences of protein X. The insert encoded the putative presequence and the first 262 amino acid residues of the mature protein. Sequence analysis revealed the presence of an *Xho* I restriction site immediately preceding the presequence coding region. Further Southern blot analysis of genomic DNA digested with combinations of *Xho* I and other restriction endonucleases yielded a 2.1-kb *Xho* I/*Xba* I restriction fragment, designated XX1. The XX1 fragment was cloned and sequenced and was found to contain the complete protein X gene.

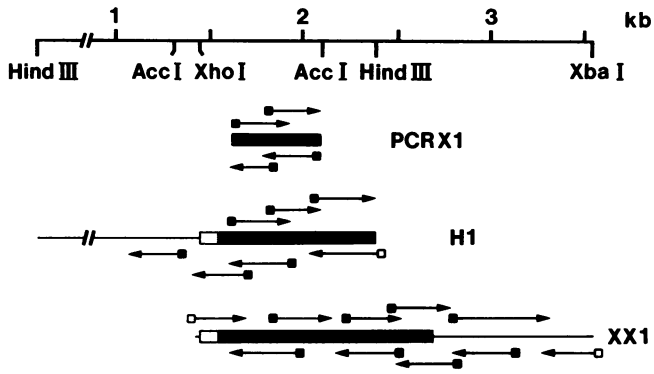

FIG. 2. Restriction endonuclease map and sequencing strategy for protein X gene. Solid bars, coding regions for the mature protein; open bar, coding region for the putative presequence. Arrows denote the direction and extent of nucleotide sequencing. Arrows beginning with open boxes represent sequences determined with vector sequencing primers. Arrows beginning with solid boxes represent sequences determined by use of synthetic oligonucleotide primers based on previously obtained sequences.

Sequence Analysis of the DNA Encoding Protein X. The restriction map and sequencing strategy for the protein X gene are shown in Fig. 2. The PCR amplified DNA fragment PCRX1 contains 487 nucleotides encoding amino acid resi-

```
AAAAGACTCTCAATCTAAAAATATGTTTGAATAATTTATCATGCCCTCACAAGTACACAC   60
AAACACAGACACATAATATACATACATATATATATATCACCGTTATTATGCGTGCACATG  120
ACAATGCCCTTGTATGTTTCGTATACTGTAGCAAGTAGTCATCATTTTGTTCCCNGTTCG  180
GAAAATGAGAAAAAAAGTAAAATCAATAAATGAAGAGTAAAAAACAATTTATGAAAGGGTG  240
AGCGACCAGCAACGAGAGAGACAAATCAAATTAGCGCTTTCCAGTGAGAATATAAGAGAG  300
CATTGAAAGAGCTAGGTTATTGTTAAATCATCTCGAGGAAAAAATTCAAAAATGCTAAGT  360
                                                  M  L  S   -28

GCAATTTCCAAAGTCTCCACTTTAAAATCATGTACAAGATATTTAACCAAATGCAACTAT  420
 A  I  S  K  V  S  T  L  K  S  C  T  R  Y  L  T  K  C  N  Y   -8

CATGCATCAGCTAAATTACTTGCTGTAAAGACATTTTCAATGCCTGCAATGTCTCCTACT  480
 H  A  S  A  K  L  L  A  V  K  T  F  S  M  P  A  M  S  P  T   13

ATGGAGAAAGGGGGGATTGTGTCTTGGAAATATAAAGTTGGCGAACCATTCAGCGCGGGC  540
 M  E  K  G  G  I  V  S  W  K  Y  K  V  G  E  P  F  S  A  G   33

GATGTGATATTAGAAGTGGAAACAGATAAATCTCAAATTGATGTGGAAGCACTGGACGAT  600
 D  V  I  L  E  V  E  T  D  K  S  Q  I  D  V  E  A  L  D  D   53

GGTAAACTAGCTAAGATCCTGAAAGATGAAGGCTCTAAAGATGTTGATGTTGGTGAACCT  660
 G  K  L  A  K  I  L  K  D  E  G  S  K  D  V  D  V  G  E  P   73

ATTGCTTATATTGCTGATGTTGATGATGATTTAGCTACTATAAAGTTACCCCAAGAGGCC  720
 I  A  Y  I  A  D  V  D  D  D  L  A  T  I  K  L  P  Q  E  A   93

AACACCGCAAATGCGAAATCTATTGAAATTAAGAAGCCATCCGCAGATAGTACTGAAGCA  780
 N  T  A  N  A  K  S  I  E  I  K  K  P  S  A  D  S  T  E  A  113

ACACAACAACATTTAAAAAAAGCCACAGTTACACCAATAAAAACCGTTGACGGCAGCCAA  840
 T  Q  Q  H  L  K  K  A  T  V  T  P  I  K  T  V  D  G  S  Q  133

GCCAATCTTGAACAGACGCTATTACCATCCGTGTCATTACTACTGGCTGAGAACAATATA  900
 A  N  L  E  Q  T  L  L  P  S  V  S  L  L  L  A  E  N  N  I  153

TCCAAACAAAAGGCTTTGAAGGAAATTGCGCCATCTGGTTCCAACGGTAGACTATTAAAG  960
 S  K  Q  K  A  L  K  E  I  A  P  S  G  S  N  G  R  L  L  K  173

GGTGATGTGCTAGCATACCTAGGGAAAATACCACAAGATTCGGTTAACAAGGTAACAGAA 1020
 G  D  V  L  A  Y  L  G  K  I  P  Q  D  S  V  N  K  V  T  E  193

TTTATCAAGAAGAACGAACGTCTCGATTTATCGAACATTAAACCTATACAGCTCAAACCA 1080
 F  I  K  K  N  E  R  L  D  L  S  N  I  K  P  I  Q  L  K  P  213

AAAATAGCCGAGCAAGCTCAAACAAAAGCTGCCGACAAGCCAAAGATTACTCCTGTAGAA 1140
 K  I  A  E  Q  A  Q  T  K  A  A  D  K  P  K  I  T  P  V  E  233

TTTGAAGAGCAATTAGTGTTCCATGCTCCCGCCTCTATTCCGTTTGACAAACTGAGTGAA 1200
 F  E  E  Q  L  V  F  H  A  P  A  S  I  P  F  D  K  L  S  E  253

TCATTGAACTCTTTCATGAAAGAAGCTTACCAGTTCTCACACGGAACACCACTAATGGAC 1260
 S  L  N  S  F  M  K  E  A  Y  Q  F  S  H  G  T  P  L  M  D  273

ACAAATTCGAAATACTTTGACCCTATTTTCGAGGACCTTGTCACCTTGAGCCCAAGAGAG 1320
 T  N  S  K  Y  F  D  P  I  F  E  D  L  V  T  L  S  P  R  E  293

CCAAGATTTAAATTTTCCTATGACTTGATGCAAATTCCCAAAGCTAATAACATGCAAGAC 1380
 P  R  F  K  F  S  Y  D  L  M  Q  I  P  K  A  N  N  M  Q  D  313

ACGTACGGTCAAGAAGACATATTTGACCTCTTAACAGGTTCAGACGCGACTGCCTCATCA 1440
 T  Y  G  Q  E  D  I  F  D  L  L  T  G  S  D  A  T  A  S  S  333

GTAAGACCCGTTGAAAAGAACTTACCTGAAAAAAACGAATATATACTAGCGTTGAATGTT 1500
 V  R  P  V  E  K  N  L  P  E  K  N  E  Y  I  L  A  L  N  V  353

AGCGTCAACAACAAGAAGTTTAATGACGCGGAGGCCAAGGCAAAAAGATTCCTTGATTAC 1560
 S  V  N  N  K  K  F  N  D  A  E  A  K  A  K  R  F  L  D  Y  373

GTAAGGGAGTTAGAATCATTTTGAATAAAAAACACGCTTTTTCAGTTCGAGTTTATCATT 1620
 V  R  E  L  E  S  F

ATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTAGTGATTTTCCTAACTTT 1680
ATTTAGTCAAAAAATTAGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGG 1740
GCGGGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCAT 1800
CCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAAAGAATCCCAG 1860
CACCAAAAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAAC 1920
TACAGAGAACAGGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAA 1980
CCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTT 2040
CTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAAGGTTG 2100
AAACCAGTTCCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTA 2160
TTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAGT 2220
CTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAACA 2280
AAATGGTTAGAGTTGC
```

FIG. 3. Nucleotide sequence of the yeast protein X gene and the deduced amino acid sequence (one-letter amino acid symbols). The experimentally determined amino acid sequences are underlined. Two putative polyadenylylation signals are overlined.

dues 7–169 of the mature protein X. Genomic DNA insert XX1 contains the nucleotide sequence encoding the precursor of protein X as well as some 5′ and 3′ flanking sequence. Genomic DNA insert H1 contains an 897-base overlap with XX1 and encodes the 5′ flanking region and the first 292 amino acid residues of the protein X precursor.

The nucleotide sequence and the deduced amino acid sequence of protein X are shown in Fig. 3. The open reading frame of 1230 nucleotides encodes a putative presequence of 30 amino acid residues and a mature protein of 380 amino acid residues, with a calculated M_r of 42,052. The presequence is similar to the presequences of other mitochondrial matrix proteins in that it contains no acidic amino acid residues and is rich in basic and hydroxylated amino acid residues (22). The deduced amino acid sequence of the mature protein contains the experimentally determined amino acid sequences of protein X (underlined).

RNA Blot Analysis. RNA blot analysis of yeast total RNA showed that the size of the mRNA is ≈1.5 kb (Fig. 4).

Chromosome Analysis. The gene encoding protein X is located on chromosome XV (data not shown).

Comparison of Amino Acid Sequences of Yeast Protein X and E_2. Comparison of the deduced amino acid sequences of yeast protein X and E_2 (Fig. 5) reveals that the amino-terminal part of protein X (residues 1–195) resembles E_2, but the carboxyl-terminal part (residues 196–380) is quite different. The two proteins exhibit 50% sequence identity in the amino-terminal segment that corresponds to the putative lipoyl-bearing domain of E_2 (residues 1–84). This extensive homology indicates that protein X and E_2 evolved from a common ancestor.

Protein X lacks the carboxyl-terminal segment of E_2 (residues 375–454) that contains the highly conserved sequence

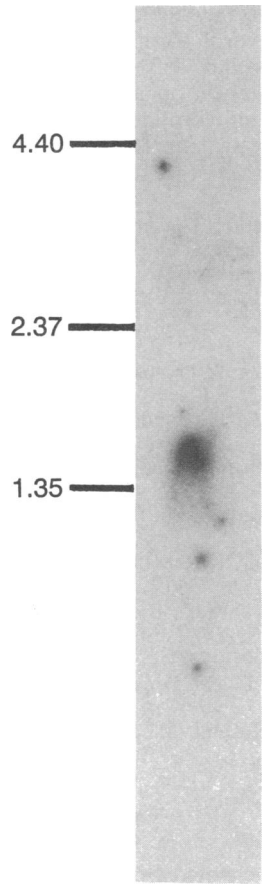

FIG. 4. RNA blot analysis of yeast total RNA. Molecular size markers, in kb, are indicated.

```
X                A V K T F S M P A M S P T M E K  16
E2 A S Y P E H T I I G M P A L S P T M T Q  20

   G G I V S W K Y K V G E P F S A G D V I  36
   G N L A A W T K K E G D Q L S P G E V I  40

   L E V E T D K S Q I D V E A L D D G K L  56
   A E I E T D K A Q M D F E F Q E D G Y L  60

   A K I L K D E G S K D V D V G E P I A -  75
   A K I L V P E G T K D I P V N K P I A V  80

   Y I A D V D D D L A T I K L P Q E A N T  95
   Y V E D K A D V P A F K D F K L E D S G 100

   A N A K S I E I K K P S A D S T E A T Q 115
   S D S K T S T K A Q P A E P Q A E K K Q 120

   Q H L K K A T V T P I K T V D G S Q A N 135
   E A P A E E T K T S A P E A K K S D V A 140

   L E Q T L L P S V S L L L A E N N I S K 155
   A P Q - G R I F A S P L A K T I A L E K 159

   Q K A L K E I A P S G S N G R L L K G D 175
   G I S L K D V H G T G P R G R I T K A D 179

   V L A Y L G K I P Q D S V N K V T E F I 195
   I E S Y L E K S S K Q S - - S Q T S G A 197

   K K N E R L D L S N I K P I Q L K P K I 215
   A A A T P A A A T S S T T A G S A P S P 217
```

```
X  A E Q A Q T K A A D K P K I T - P V E F 234
E2 S S T A - - S Y E D V P I S T M R S I I 235

   E E Q L V F H A P A S I P F D K L S E S 254
   G E R L - L Q S T Q G I P S Y I V S - S 253

   L N S F M K E A Y Q F S H G T P L M D T 274
   K I S I S K - - - L L K L R Q S L N A T 270

   - N S K Y F D P I F E D L V T L S P R E 293
   A N D K Y K L S I N D L L V K A I T V A 290

   P R F K F S Y D L M Q I P K A N N M Q D 313
   A K R V P D A N A Y W L P N E N V I R K 310

   T Y G Q E D I F D L L T G S D A T A S S 333
   F K N V D V S V A V A T P T G L L T P I 330

   V R P V E - K N L - P E K N E Y I L A L 351
   V K N C E A K G L S Q I S N E - I K E L 349

   N V S V N N K K F N D A E A K A K R F L 371
   V K R A R I N K L A P E E F Q G G T I C 369

X  D Y V R E L E S F
E2 I S N M G M N N A V N M F T S I I N P P 389

   Q S T I L A I A T V E R V A V E D A A A 409
   E N G F S F D N Q V T I T G T F D H R T 429
   I D G A K G A E F M K E L K T V I E N P 449
   L E M L L
```

FIG. 5. Comparison of amino acid sequences (single-letter code) of yeast protein X and E_2. The deduced sequences of protein X and E_2 are aligned for maximum similarity. Identical residues are indicated by asterisks. The lipoyllysine residue in the lipoyl domain and residues in the putative active site of E_2 are indicated by solid circles.

His-Xaa-Xaa-Xaa-Asp-Gly, which is thought to be part of the putative catalytic site of all dihydrolipoamide acyltransferases (23). This finding suggests that yeast protein X, in contrast to E_2, is not able to catalyze an acetyl transfer between the protein-bound *S*-acetyldihydrolipoyl moiety and coenzyme A. The availability of the structural gene for yeast protein X should facilitate elucidation of its function.

We thank the Protein Sequencing Center, University of Texas at Austin, for amino acid sequence analyses and Dr. Russell F. Doolittle for helpful comments on the manuscript. This work was supported in part by Grant GM06590 from the U.S. Public Health Service.

1. Reed, L. J. (1974) *Acc. Chem. Res.* **7,** 40–46.
2. Oliver, R. M. & Reed, L. J. (1982) in *Electron Microscopy of Proteins,* ed. Harris, J. R. (Academic, London), Vol. 2, pp. 1–48.
3. Miles, J. S. & Guest, J. R. (1988) *Biochem. Soc. Symp.* **54,** 45–65.
4. Perham, R. N., Packman, L. C. & Radford, S. E. (1988) *Biochem. Soc. Symp.* **54,** 67–81.
5. Niu, X.-D., Browning, K. S., Behal, R. H. & Reed, L. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7546–7550.
6. Coppel, R. L., McNeilage, L. J., Surh, C. D., Van de Water, J., Spithill, T. W., Whittingham, S. & Gershwin, M. E. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7317–7321.
7. De Marcucci, O. & Lindsay, J. L. (1985) *Eur. J. Biochem.* **149,** 641–648.
8. Jilka, J. M., Rahmatullah, M., Kazemi, M. & Roche, T. E. (1986) *J. Biol. Chem.* **261,** 1858–1867.
9. De Marcucci, O. G. L., Hodgson, J. A. & Lindsay, J. G. (1986) *Eur. J. Biochem.* **158,** 587–594.
10. Hodgson, J. A., De Marcucci, O. G. & Lindsay, J. G. (1986) *Eur. J. Biochem.* **158,** 595–600.
11. Rahmatullah, M. & Roche, T. E. (1987) *J. Biol. Chem.* **262,** 10265–10271.
12. Rahmatullah, M., Gopalakrishnan, S., Andrews, P. C., Chang, C. L., Radke, G. A. & Roche, T. E. (1989) *J. Biol. Chem.* **264,** 2221–2227.
13. Rahmatullah, M., Gopalakrishnan, S., Radke, G. A. & Roche, T. E. (1989) *J. Biol. Chem.* **264,** 1245–1251.
14. Powers-Greenwood, S. L., Rahmatullah, M., Radke, G. A. & Roche, T. E. (1989) *J. Biol. Chem.* **264,** 3655–3657.
15. Gopalakrishnan, S., Rahmatullah, M., Radke, G. A., Powers-Greenwood, S. L. & Roche, T. E. (1989) *Biochem. Biophys. Res. Commun.* **160,** 715–721.
16. Uhlinger, D. J., Yang, C.-Y. & Reed, L. J. (1986) *Biochemistry* **25,** 5673–5677.
17. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
18. Matsudaira, P. (1987) *J. Biol. Chem.* **262,** 10035–10038.
19. Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. (1977) *J. Biol. Chem.* **252,** 1102–1106.
20. Gulcher, J. R. & Stefansson, K. (1988) *Nucleic Acids Res.* **16,** 10931.
21. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
22. von Heijne, G. (1986) *EMBO J.* **5,** 1335–1342.
23. Guest, J. R. (1987) *FEMS Microbiol. Lett.* **44,** 417–422.