THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 256, No. 3, Issue of February 10, pp. 1385-1395, 1981
Printed in U.S.A.

The Primary Structures of Two Yeast Enolase Genes

HOMOLOGY BETWEEN THE 5′ NONCODING FLANKING REGIONS OF YEAST ENOLASE AND GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE GENES*

(Received for publication, June 30, 1980, and in revised form, September 4, 1980)

Michael J. Holland‡, Janice P. Holland, Gregory P. Thill§, and Kimberly A. Jackson

From the Department of Biochemistry, The University of Connecticut Health Center, Farmington, Connecticut 06032

Segments of yeast genomic DNA containing two enolase structural genes have been isolated by subculture cloning procedures using a cDNA hybridization probe synthesized from purified yeast enolase mRNA. Based on restriction endonuclease and transcriptional maps of these two segments of yeast DNA, each hybrid plasmid contains a region of extensive nucleotide sequence homology which forms hybrids with the cDNA probe. The DNA sequences which flank this homologous region in the two hybrid plasmids are nonhomologous indicating that these sequences are nontandemly repeated in the yeast genome. The complete nucleotide sequence of the coding as well as the flanking noncoding regions of these genes has been determined. The amino acid sequence predicted from one reading frame of both structural genes is extremely similar to that determined for yeast enolase (Chin, C. C. Q., Brewer, J. M., Eckard, E., and Wold, F. (1981) *J. Biol. Chem.* 256, 1370-1376), confirming that these isolated structural genes encode yeast enolase. The nucleotide sequences of the coding regions of the genes are approximately 95% homologous, and neither gene contains an intervening sequence. Codon utilization in the enolase genes follows the same biased pattern previously described for two yeast glyceraldehyde-3-phosphate dehydrogenase structural genes (Holland, J. P., and Holland, M. J. (1980) *J. Biol. Chem.* 255, 2596-2605). DNA blotting analysis confirmed that the isolated segments of yeast DNA are colinear with yeast genomic DNA and that there are two nontandemly repeated enolase genes per haploid yeast genome. The noncoding portions of the two enolase genes adjacent to the initiation and termination codons are approximately 70% homologous and contain sequences thought to be involved in the synthesis and processing of messenger RNA. Finally there are regions of extensive homology between the two enolase structural genes and two yeast glyceraldehyde-3-phosphate dehydrogenase structural genes within the 5′ noncoding portions of these glycolytic genes.

Enolase prepared from a wide variety of eucaryotic cells appears to exist as multiple cellular forms (1). These results have led to the hypothesis that the structural gene for enolase is repeated and that some of these multiple forms of the native dimer are isozymes. Multiple yeast enolases have been resolved by anion exchange chromatography (2-4), gel electrophoresis (5), isoelectric focusing (6), and countercurrent distribution techniques (7, 8). Westhead and McLain (2) originally showed that the multiple forms of yeast enolase resolved by *O*-(triethylaminoethyl)-Sephadex chromatography arise as a consequence of aging cell extracts, and they suggested that these forms of the enzyme may arise as a consequence of *in vitro* deamidation of the enzyme. Porcelli *et al.* (4) have recently shown that the major form of yeast enolase resolved after *O*-(triethylaminoethyl)-Sephadex chromatography is the precursor to the two minor forms of the enzyme. These latter investigators suggest that the minor forms arise by a process which involves dissociation of the native dimer. Based on these data it is clear that some portion of the heterogeneity observed in preparations of yeast enolase is the result of *in vitro* modifications. It is not clear, however, if all multiple forms of yeast enolase correspond to artifacts of isolation.

Enolase messenger RNA is one of the most abundant mRNAs in vegetative yeast cells (9). Previous reports from this laboratory have shown that glyceraldehyde-3-phosphate dehydrogenase mRNA, an abundant yeast mRNA, can be obtained in sufficient purity to permit the cloning of its respective structural genes (10). We have used a similar approach to isolate and identify enolase structural genes utilizing a cDNA probe synthesized from partially purified enolase mRNA. We report the isolation and characterization of two nontandemly repeated yeast enolase structural genes. The primary structures of the polypeptides encoded by these two enolase genes differ in 20 out of 436 amino acids. The primary structure of one of these polypeptides is in close agreement to that determined by Chin *et al.* (17) for the major "A" form of yeast enolase resolved after *O*-(triethylaminoethyl)-Sephadex chromatography. Evidence is presented that both enolase structural genes are expressed in vegetative yeast cells suggesting that isozymes of yeast enolase do exist in the cell.

The primary structures of the two enolase genes are 95% homologous within their respective coding regions, and the primary structures of the 5′ and 3′ noncoding regions adjacent to the translational initiation and termination codons are approximately 70% homologous. Comparison of the 5′ and 3′ noncoding regions of the two enolase genes with two nontandemly repeated yeast glyceraldehyde-3-phosphate dehydrogenase structural genes (11, 12) shows little if any homology among the four genes within the 3′ noncoding regions; however, the 5′ noncoding regions of all four glycolytic genes are extremely homologous within the first 100 nucleotides adjacent to the translational initiation codons. These latter results suggest that the 5′ noncoding regions of enolase and glyceraldehyde-3-phosphate dehydrogenase structural genes may have evolved from a common nucleotide sequence.

* This research was supported by a grant from The National Institute of General Medical Sciences and a grant from the March of Dimes Birth Defects Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ American Heart Association Established Investigator.

§ Present address, Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205.

EXPERIMENTAL PROCEDURES

The miniprint supplement contains all experimental details.[1]

RESULTS

Subculture Cloning of peno8 and peno46—Segments of yeast DNA containing sequences complementary to yeast enolase mRNA were isolated by subculture cloning procedures as described in the miniprint supplement. The results of the screening experiments are illustrated in Table I. The subculture cloning procedure used to identify transformants containing yeast enolase structural genes was the same as that previously described for the isolation of a yeast glyceraldehyde-3-phosphate dehydrogenase structural gene (10). A series of cultures was started with a defined number of *Escherichia coli* transformants containing hybrid plasmids composed of the bacterial vector pSF2124 and randomly sheared segments of yeast genomic DNA. The yeast DNA was joined into the *Eco* RI restriction endonuclease cleavage site by the A/T joining procedure as previously described (10). Total hybrid plasmid DNA was isolated from each culture, depurinated, and hybridized against a cDNA probe synthesized from partially purified yeast enolase mRNA. Hybridization reactions were carried out in solution with a vast excess of depurinated hybrid plasmid DNA. Hybridization was carried out to saturation of the probe, and the percentage of the labeled cDNA which formed hybrids was determined by digestion of the reaction mixture with the single strand specific nuclease, S1. Since the enolase mRNA was estimated to be approximately 20% pure, cultures which contained an enolase structural gene were expected to protect 20% more of the probe sequences than cultures which lacked an enolase gene. The hybridization detected in cultures which lacked an enolase structural gene is due to yeast DNA sequences which are complementary to contaminants in the enolase cDNA. In the case of *peno8*, approximately 22% of the cDNA synthesized from partially purified enolase mRNA formed hybrids with the isolated hybrid plasmid DNA whereas *peno46* plasmid DNA hybridized with 16% of the cDNA sequences. Although the extent of hybridization of each hybrid plasmid DNA with the cDNA differed, a mixture of *peno8* and *peno46* plasmid DNA hybridized to the same extent as *peno8* DNA alone confirming that the two plasmid contain homologous sequences. We reproducibly observe reduced hybridization of the cDNA with the *peno46* plasmid relative to the *peno8* plasmid suggesting that *peno8* contains sequences which are complementary to a larger percentage of the cDNA. The significance of these hybridization results will be analyzed under "Discussion."

Restriction Endonuclease Mapping of peno8 and peno46—In order to establish the extent of nucleotide sequence homology between the two isolated plasmids, a restriction endonuclease cleavage map of *peno8* and *peno46* was generated as described in the miniprint supplement. As illustrated in Fig. 1, each hybrid plasmid contains a region of extensive nucleotide sequence homology (indicated by the dark region of the map in Fig. 1). The nucleotide sequences which flank this region of homology are not homologous indicating that these two hybrid plasmids contain different portions of yeast genomic DNA.

[1] "Experimental Procedures" are presented in a miniprint supplement immediately following this paper. The restriction mapping analysis of *peno8* and *peno46*, Tables Is and IIs, and additional references are also described in the miniprint supplement. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9650 Rockville Pike, Bethesda, Maryland 20014. Request Document 80M-1329, cite author(s), and include a check or money order for $2.40 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

TABLE I

Subculture cloning of peno8 and peno46

Total plasmid DNA was isolated from each culture (inoculated with the number of transformants indicated), depurinated, and hybridized against cDNA synthesized from partially purified enolase mRNA. An average of 6 to 12 cultures were tested at each step of the subculture cloning procedure. The percentage hybridization values shown are the average for cultures which were considered either positive or negative. Purified *peno8* and *peno46* plasmid DNAs were depurinated and hybridized against cDNA synthesized from partially purified enolase mRNA. The percentage hybridization values indicated are for *peno8* and *peno46*, respectively, as well as a mixture of the two hybrid plasmid DNAs.

Transformants/culture	% Hybridization	
	Negative	Positive
peno8		
400	20	41
100	13	38
12	7	27
1	4	22
peno46		
600	14	40
41	9	26
6	13	25
1	15	27

Plasmid DNA	% Hybridization
peno8	22
peno46	16
peno8 plus *peno46*	21

In order to establish that the homologous segments of the plasmids contain sequences which hybridize with the cDNA, a series of DNA blotting experiments was carried out. Initially, yeast genomic DNA was digested with the restriction endonucleases *Xba* I and *Xho* I, electrophoresed on a 0.8% agarose slab gel, transferred to nitrocellulose filter paper, and hybridized with purified enolase cDNA as described in the miniprint supplement. Two *Xba* I fragments (12 kb[2] and 6.4 kb) and two *Xho* I fragments (12 kb and 10 kb) formed hybrids with the cDNA, respectively. Since the isolated plasmid DNAs are not cleaved by *Xba* I or *Xho* I within the homologous regions of the plasmids, we conclude that the commercial baker's yeast (strain F1) contains only two sequences which form hybrids with the cDNA. In order to confirm that these sequences are contained within *peno8* and *peno46*, yeast genomic DNA and the two isolated plasmids were cleaved with the restriction endonucleases, *Eco* RI and *Hind* III, electrophoresed on a 0.8% agarose slab gel, transferred to nitrocellulose, and hybridized with the cDNA. As illustrated in Fig. 2, a 3.2-kb and a 0.27-kb *Hind* III fragment from yeast genomic DNA formed hybrids with the cDNA. The 3.2-kb fragment corresponds to a 3.2-kb fragment generated from *peno8*, while the 0.27-kb fragment corresponds to a fragment generated from *peno46*. In the case of yeast DNA digested with *Eco* RI, a 2.0-kb fragment formed a strong hybrid while a 1.6-kb fragment formed a weak hybrid with the probe. These *Eco* RI fragments correspond to two *Eco* RI fragments generated from *peno46*. Based on these data, we conclude that sequences present within the right end of the homologous region shown in Fig. 1 are capable of hybridizing with the cDNA probe. Since the cDNA used in this experiment is not a full length copy of the enolase mRNA (average size 600 bases), we conclude that the sequences which form hybrids with the probe are complementary to only the sequences at the 3′ end of the mRNA. This conclusion is confirmed by the nucleotide sequence described below. The fact that the isolated sequences in *peno8* and *peno46* which form hybrids with enolase cDNA correspond with sequences in yeast genomic DNA

[2] The abbreviation used is: kb, kilobase pair.

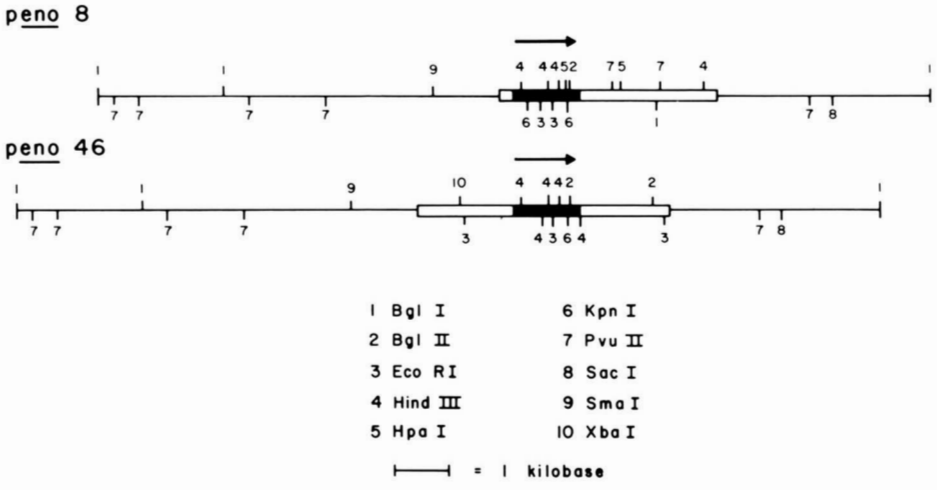

FIG. 1. **Restriction endonuclease cleavage maps of p*eno*8 and p*eno*46.** The *single horizontal lines* indicate pSF2124 vector DNA sequences and the *double horizontal lines* indicate the yeast DNA portions of the hybrid plasmids. The *shaded regions* indicate the location of the enolase structural genes in the plasmids, and the *arrow* indicates the direction of transcription of the enolase structural genes. These cleavage maps were derived from the data reported in the miniprint supplement.

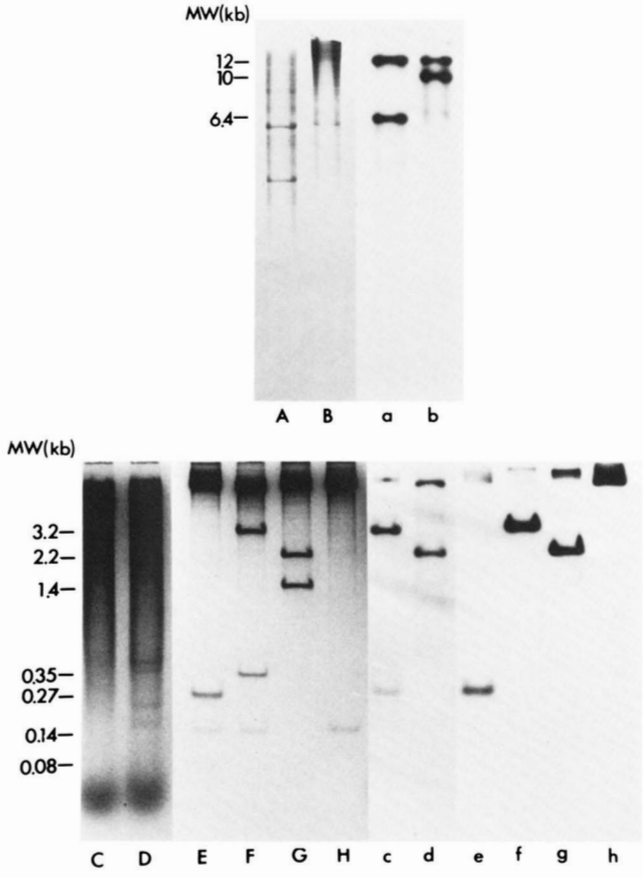

FIG. 2. **Hybridization of DNA filter blots of restriction endonuclease cleavage fragments of p*eno*8, p*eno*46, and yeast genomic DNA with an enolase cDNA probe.** Lanes *A* and *B* are agarose gels of yeast genomic DNA (strain F_1) digested with *Xba*I and *Xho*I, respectively, visualized after ethidium bromide staining. Lanes *a* and *b* are autoradiograms of hybrids formed between ^{32}P-labeled enolase cDNA and the *Xba*I and *Xho*I digested yeast DNA, respectively, after transfer to a nitrocellulose filter. Lanes *C* and *D* are agarose gels of yeast genomic DNA digested with *Hin*dIII and *Eco*R1, respectively, and lanes *c* and *d* are autoradiograms of hybrids formed between enolase cDNA and the DNA fragments shown in lanes *C* and *D*, respectively. Lane *E*, an agarose gel of p*eno*46 digested with *Hin*dIII; lane *e*, an autoradiogram of hybrids formed between ^{32}P-labeled enolase cDNA and *Hin*dIII-digested p*eno*46; lane *F*, an agarose gel of p*eno*8 digested with *Hin*dIII; lane *f*, an autoradiogram of hybrids formed between ^{32}P-labeled enolase cDNA and *Hin*dIII-digested p*eno*8; lane *G*, an agarose gel of p*eno*46 digested with *Eco*R1; lane *g*, an autoradiogram of hybrids formed between ^{32}P-labeled enolase cDNA and *Eco*R1-digested p*eno*46; lane *H*, an agarose gel of p*eno*8 digested with *Eco*R1; lane *h*, an autoradiogram of hybrids formed between ^{32}P-labeled enolase cDNA and *Eco*R1-digested p*eno*8. Restriction endonuclease cleavage fragments were electrophoresed on agarose gels and transferred to a nitrocellulose filter before hybridization. Enolase cDNA was purified prior to hybridization as described in the miniprint supplement.

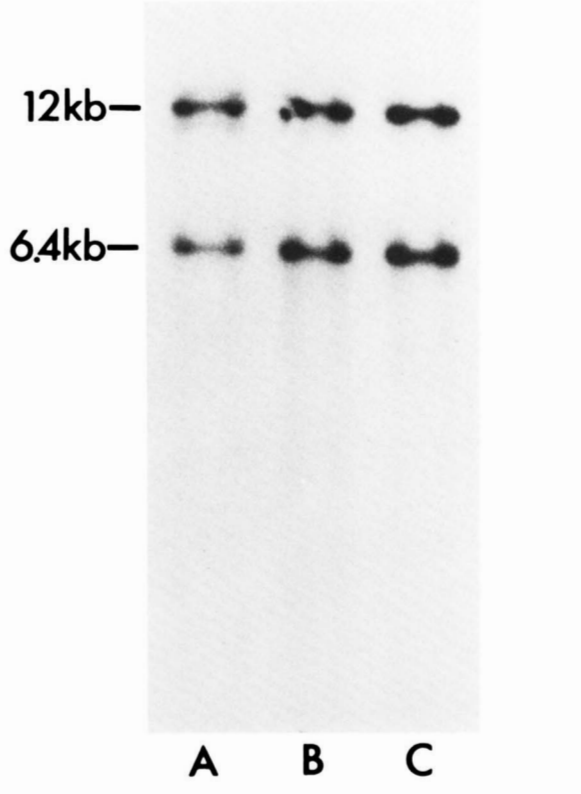

FIG. 3. **Hybridization of yeast genomic DNA digested with *Xba*I with an enolase cDNA probe.** Genomic yeast DNA isolated from commercially grown strain F1 (lane *A*), strain F1 grown in YPD media (lane *B*) and strain D273-10B grown in YPD media (lane *C*) digested with *Xba*I, electrophoresed on a 1% agarose slab gel, transferred to a nitrocellulose filter, and hybridized with a ^{32}P-labeled enolase cDNA. The enolase cDNA was purified prior to hybridization as described in the miniprint supplement.

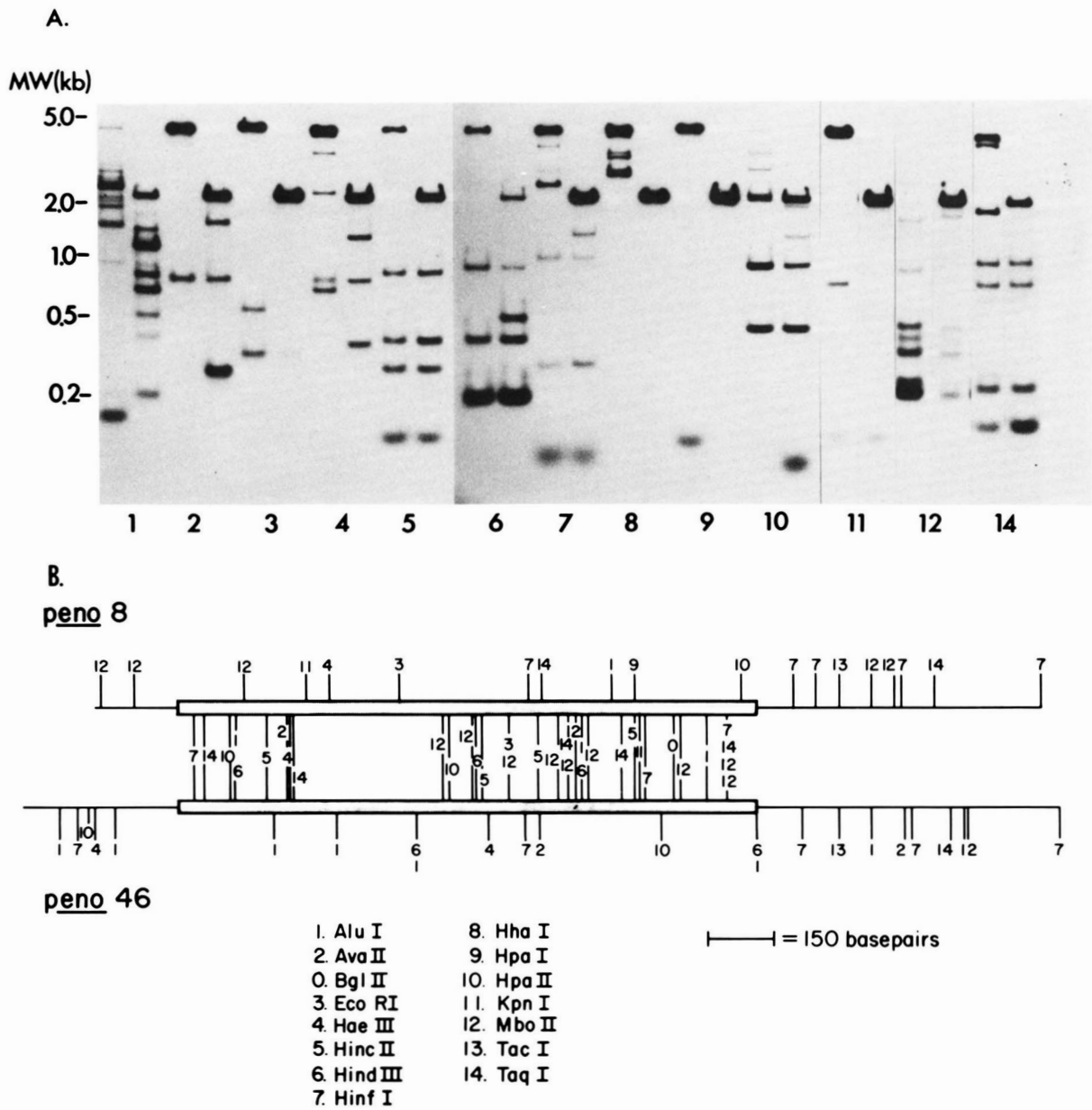

FIG. 4. **Restriction endonuclease cleavage maps of the enolase structural gene portions of p*eno*8 and p*eno*46.** *A*, an autoradiogram of a 1.5% agarose slab gel of partial restriction endonuclease digests of 4.4-kb and 2.0-kb fragments from p*eno*8 and p*eno*46, respectively. Each fragment was ^{32}P-labeled at the 5′ terminus of a common *Bgl*II cleavage site within each respective enolase structural gene. The 4.4-kb fragment from p*eno*8 extends from the *Bgl*II cleavage site to a *Pvu*II cleavage site within the vector DNA to the left of the enolase structural gene (Fig. 1). The 2.0-kb fragment from p*eno*46 extends from the *Bgl*II cleavage site to the *Xba*I cleavage site within the hybrid plasmid. Partial digests of these fragments with the restriction endonucleases indicated were electrophoresed in parallel lanes of the slab gel. The fragments derived from p*eno*8 and p*eno*46 are in the left and right parallel lanes, respectively. *B*, the cleavage map predicted from the partial digests shown in *A* as well as the data described in the text. The numbers below the parallel lanes in *A* correspond to the restriction endonuclease key in *B*.

supports the conclusion that the yeast DNA sequences in the hybrid plasmids are colinear with genomic DNA sequences.

The DNA blotting analysis shown in Fig. 2 was carried out with DNA isolated from commercially grown baker's yeast (strain F1). Similar DNA blots were carried out with DNA isolated from the haploid *Saccharomyces cerevisiae* strain D273-10B and strain F1 grown under the same conditions as the haploid strain. As illustrated in Fig. 3, the DNA blots with DNA isolated from haploid and commercial baker's yeast after digestion with the restriction endonuclease *Xba* I demonstrate a 12-kb and a 6.4-kb fragment which form hybrids with enolase cDNA. These data confirm that there are two nontandemly repeated enolase structural genes per haploid yeast genome. Furthermore, the number and relative location of the two genes in the genome of strain F1 is not influenced by the growth conditions for propagating the cells.

Fig. 4 shows a more detailed restriction endonuclease cleavage map of the homologous regions of p*eno*8 and p*eno*46. The restriction endonuclease cleavage map was generated as described in the miniprint supplement utilizing the partial cleavage method described by Smith and Birnstiel (13). The molecular weights of each of the restriction endonuclease cleavage fragments were determined from limit digests with each of the restriction endonucleases tested. Fig. 4*A* shows a series of partial digests of two DNA fragments generated by digestion of p*eno*8 with *Bgl* II and *Pvu* II (4.4 kb) and by digestion of p*eno*46 with *Bgl* II and *Xba* I (2.0 kb). Both fragments were labeled at the common *Bgl* II cleavage site with [α-^{32}P]ATP

and polynucleotide kinase as described in the miniprint supplement. As illustrated by the partial cleavages and the cleavage map shown in Fig. 4, *A* and *B*, the DNA sequences surrounding the common *Bgl* II cleavage site are extremely homologous in the two hybrid plasmids. The region of nucleotide sequence homology between the two plasmids is the approximate size needed to encode yeast enolase. This observation is confirmed by the nucleotide sequence described below. As initially pointed out from the data shown in Fig. 1, the nucleotide sequences which flank the homologous portions of the plasmids are nonhomologous indicating that the genes are nonallelic and are nontandemly repeated in the yeast genome. Based on the restriction endonuclease map (Fig. 4*B*) approximately 62% of the cleavage sites are in common between the two genes. These data, therefore, suggest that the polypeptides encoded by these two genes have diverged in primary structure.

The Nucleotide Sequence of the Coding Portions of peno8 and peno46—The complete nucleotide sequence of the coding regions of p*eno*8 and p*eno*46 is shown in Fig. 5. The nucleotide sequence was determined by the method of Maxam and Gilbert (14) as described in the miniprint supplement using the strategy outlined in Fig. 6. The continuous nucleotide sequence shown in Fig. 5 corresponds to the enolase gene in p*eno*46. The nucleotide sequence of the coding region of p*eno*8 is identical with that from p*eno*46 except at the positions indicated by the codons above the continuous sequence. At these positions the entire codon in p*eno*8 is indicated when that codon differs from the respective codon in p*eno*46. The amino acid sequence predicted from one reading frame of p*eno*46 is indicated below the nucleotide sequence in Fig. 5. In the case of a codon change in p*eno*8 resulting in a change in the predicted amino acid sequence relative to p*eno*46, the new amino acid is indicated above the p*eno*8 codon. The polypeptides encoded by the two structural genes differ in 20 out of 436 residues. The polypeptide encoded by p*eno*46 agrees in primary structure extremely well with the yeast enolase primary structure determined by Chin *et al.* (17). Based on these data, it is likely that the enolase protein which was sequenced by Chin *et al.* is a dimer of the polypeptide encoded by p*eno*46. The relevance of these two gene products to the multiple forms of yeast enolase will be discussed below. These data confirm the identification of two nontandemly repeated yeast enolase structural genes contained within the hybrid plasmids described here. The primary structures of the two enolase genes are 95% homologous within the coding regions. Based on the agreement between the primary structure determined for yeast enolase and those predicted from the gene sequences, we can conclude that neither gene contains intervening sequences. There are a total of 33 silent third position changes of a total of 67 nucleotide differences between the two coding regions.

The codon utilization pattern for the two enolase structural genes is shown in Table II. The observed codon bias is extremely similar to that observed for two yeast glyceraldehyde-3-phosphate dehydrogenase structural genes (12). The major exception to the bias observed in glyceraldehyde-3-phosphate dehydrogenase is the occurrence of the UUA codon for leucine twice in p*eno*8 and three times in p*eno*46. The CGU codon for arginine is present at position 14 in both enolase gene sequences and was not observed in the glyceraldehyde-3-phosphate dehydrogenase structural genes. Cysteine, glutamine, glutamate, histidine, and glycine are encoded by a single codon. Arginine, asparagine, phenylalanine, proline, and tyrosine are also encoded by a single codon with only seven exceptions. Lysine and leucine are encoded by AAG and UUG codons approximately 90% of the time, and finally

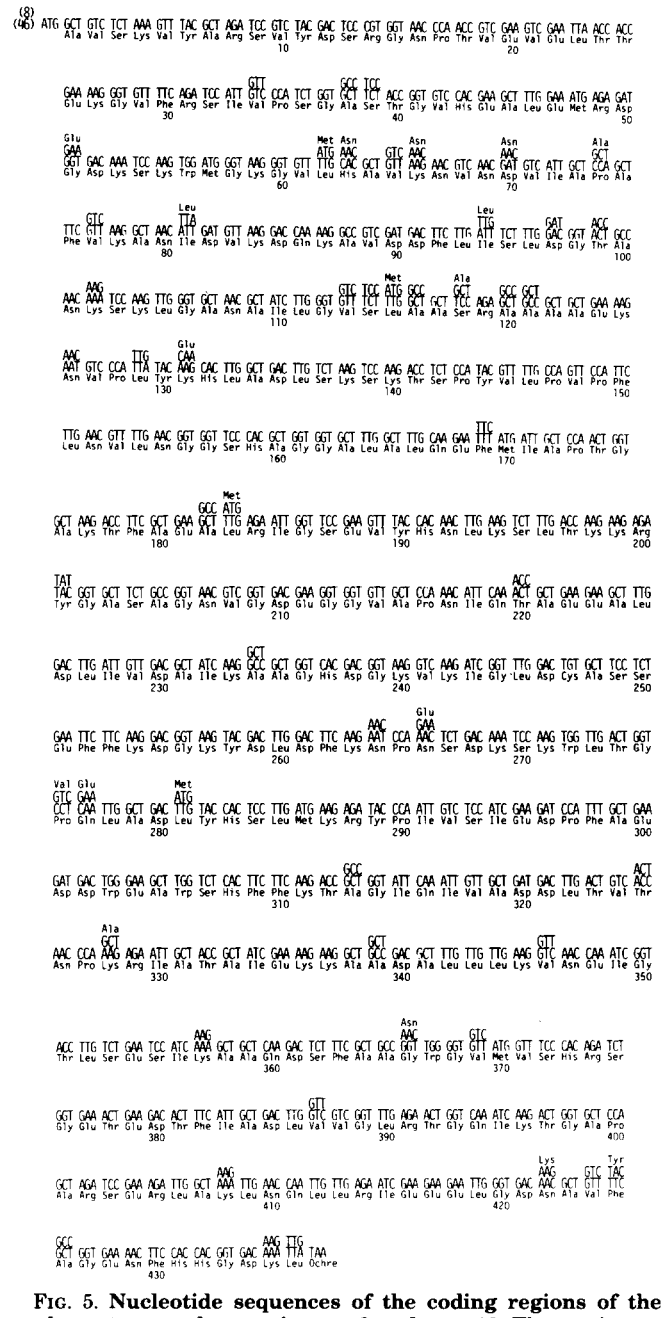

FIG. 5. **Nucleotide sequences of the coding regions of the enolase structural genes in *peno*8 and *peno*46.** The continuous nucleotide sequence is for the enolase structural gene in p*eno*46 extending from the ATG initiation codon to the TAA termination codon. The amino acid sequence predicted from one reading frame of the nucleotide sequence is shown below the continuous nucleotide sequence. The nucleotide sequence of the enolase structural gene within p*eno*8 is identical with the p*eno*46 sequence with the exception of the codons illustrated above the continuous nucleotide sequence. In cases where a codon in p*eno*8 encodes a different amino acid than the codon p*eno*46, the new amino acid is illustrated above the p*eno*8 codon.

alanine, aspartate, isoleucine, serine, threonine, and valine are encoded by two codons which contain either U or C in third position. As was the case with the glyceraldehyde-3-phosphate dehydrogenase genes, we observe only 10 to 20% change between U and C in the third positions of the codons for the six amino acids which are encoded by two codons with either C or U in third position. This variation is significantly less than predicted for random change at these third positions,

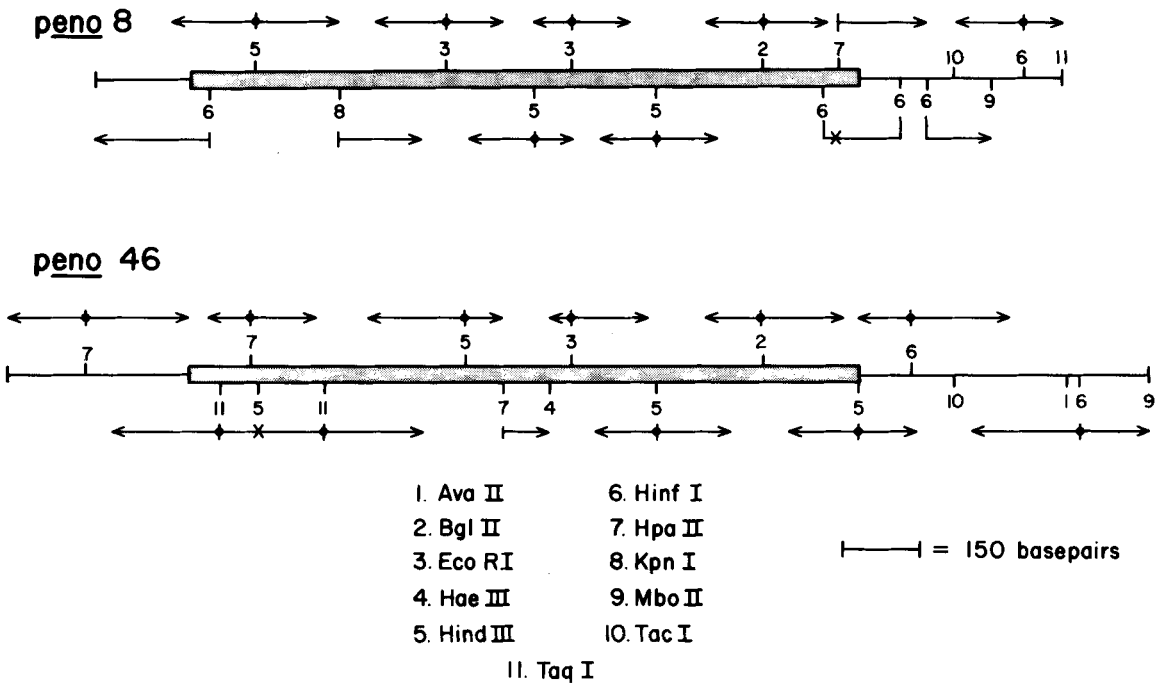

FIG. 6. **Strategy for determining the nucleotide sequences of the enolase structural genes in *peno8* and *peno46*.** The *arrows* indicate the direction and amount of sequence determined from each end-labeled restriction endonuclease fragment.

TABLE II

Codon utilization in peno8 and peno46

The codon utilization within the coding portions of *peno8* and *peno46* was tabulated from the primary structure shown in Fig. 5. The codon usage pattern for *peno8* is shown in parentheses.

ALA GCA (0) 0	ARG CGA (0) 0	ASN AAT (0) 2	ASP GAT (7) 7
T (48) 48	T (1) 1	C (21) 17	C (23) 24
C (10) 7	C (0) 0		
G (0) 0	G (0) 0		
	AGA (13) 13		
	G (0) 0		
CYS TGT (1) 1	GLN CAA (9) 9	GLU GAA (28) 25	GLY GGA (0) 0
C (0) 0	G (0) 0	G (0)	T (35) 37
			C (0) 0
			C (0) 0
HIS CAT (0) 0	ILEU ATA (0) 0	LEU TTA (2) 3	LYS AAA (3) 7
C (10) 10	T (11) 13	G (36) 37	G (32) 30
	C (9) 10	CTA (0) 0	
		(0) 0	
		T (0) 0	
		C (0) 0	
		G (0) 0	
MET ATG (9) 5	PHE TTT (1) 2	PRO CCA (13) 14	SER TCA (0) 0
	C (14) 14	T (0) 1	T (13) 15
		C (0) 0	C (17) 16
		G (0) 0	G (0) 0
			AGT (0) 0
			C (0) 0
THR ACA (0) 0	TRP TGG (5) 5	TYR TAT (1) 0	VAL GTA (0) 0
T (8) 9		C (9) 9	T (15) 17
C (12) 11			C (20) 17
G (0) 0			G (0) 0

suggesting that divergence of the duplicated genes for enolase relative to each other may be restricted.

Nucleotide Sequence of the 5′ and 3′ Noncoding Flanking Regions of peno8 and peno46—The nucleotide sequences of the 5′ and 3′ noncoding regions of *peno8* and *peno46* are shown in Fig. 7, *A* and *B*, respectively. The nucleotide sequences of the two structural genes have been aligned to maximize homology between the two genes. The alignments necessitated leaving gaps in one of the two sequences. Within the 5′ noncoding sequences of the genes the A+T composition is approximately 75% for the first 150 nucleotides adjacent to the ATG initiation codon. Approximately 77% of the first 75 nucleotides adjacent to the initiation codon in *peno8* are homologous to *peno46*. In order to maximize homology between *peno8* and *peno46* within this region, it is necessary to allow a 36-nucleotide deletion in the *peno8* sequence. The 25 nucleotides upstream from this deletion in *peno8* are greater than 85% homologous to sequences in *peno46*. The structural gene in *peno46* contains a -TATAAA- sequence 140 nucleotides from the initiation codon. The position of these sequences is similar to the -TATAAA-sequences in two yeast glyceraldehyde-3-phosphate dehydrogenase structural genes (12) and the cytochrome C-1 gene from yeast (15). *peno8* does not contain -TATAAA- sequences within 175 nucleotides of the initiation codon. As previously described for the yeast glyceraldehyde-3-phosphate dehydrogenase structural genes (12), both enolase genes contain short nucleotide sequences in the 5′ noncoding regions which are repeated in noninverted fashion in the 3′ noncoding portions of the respective genes. These repetitious sequences are indicated by the boxed regions in Fig. 7, *A* and *B*.

The nucleotide sequences adjacent to the termination codons in *peno8* and *peno46* are also extremely homologous. Both genes are approximately 65% homologous for 130 nucleotides beyond the termination codon and are 62% homologous within the 80 nucleotides surrounding a -AATAA- sequence in both genes. The -AATAA- sequence is homologous to similar sequences located in the 3′ noncoding regions of many eucaryotic genes and is usually located approximately 15 to 20 nucleotides from the site of poly A addition to the mRNA. Based on DNA blotting experiments carried out with enolase cDNA probe and restriction endonuclease fragments generated from both plasmid DNAs which contain the 3′ noncoding regions, we conclude that enolase mRNA sequences extend to the -AATAA- sequences (data not shown). The A+T composition of the first 300 nucleotides beyond the termination codon in both genes is 68%. The nucleotide sequences which are repeated in the 5′ noncoding regions of the genes are indicated by the boxed regions in Fig. 7*B*.

Comparison of the 5′ Noncoding Regions of the Two Enolase Genes with Two Yeast Glyceraldehyde-3-Phosphate Dehydrogenase Structural Genes—As described above for

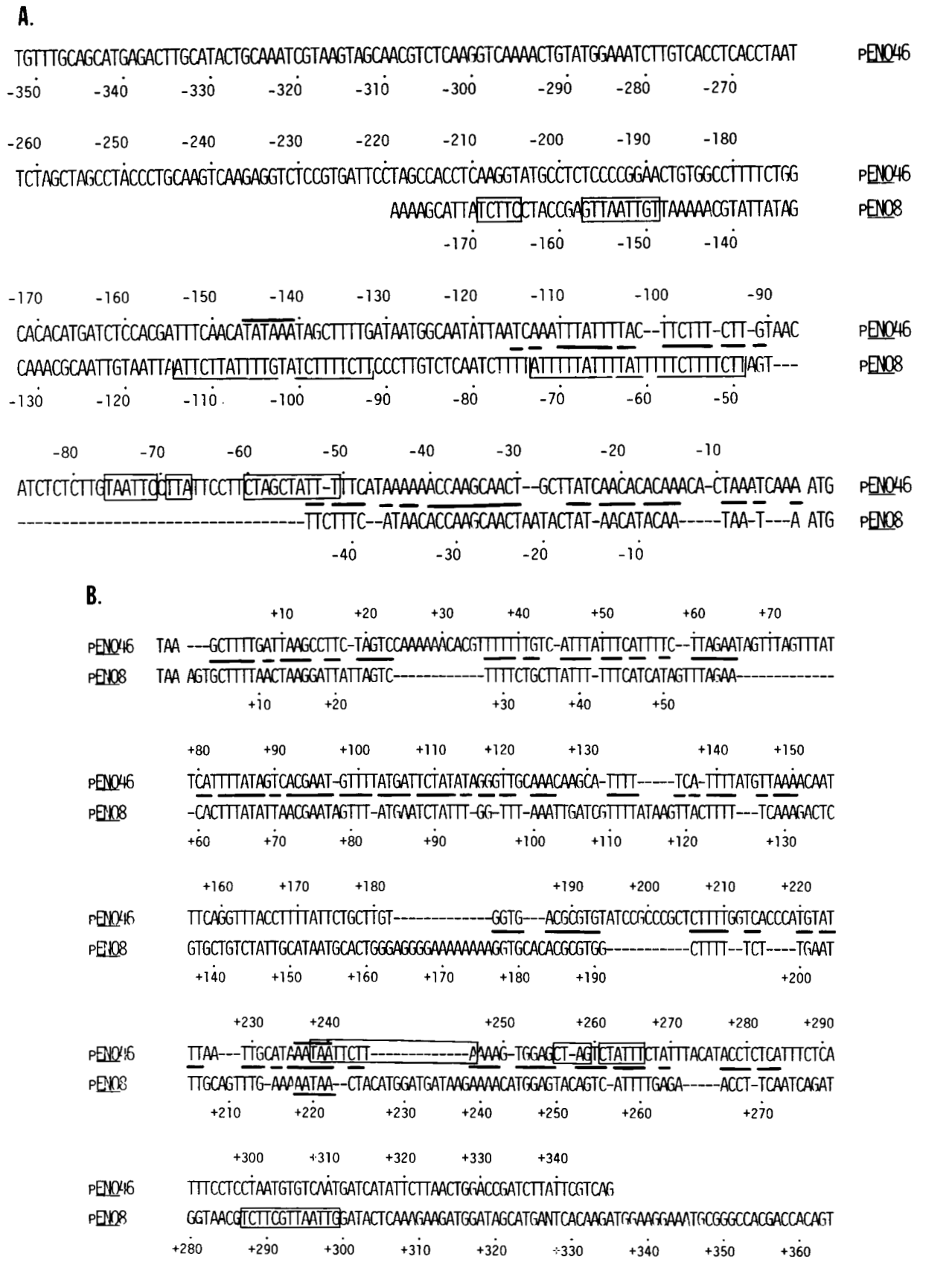

FIG. 7. **Nucleotide sequences of the 5′ and 3′ noncoding flanking regions of the enolase structural genes; comparison of the corresponding regions in *peno*8 and *peno*46.** *A*, shows the 5′ noncoding flanking regions of p*eno*8 and p*eno*46; *B*, shows the 3′ noncoding flanking regions of p*eno*8 and p*eno*46. The *solid bars* between the p*eno*8 and p*eno*46 sequences indicate homologous regions of the sequences. The *boxed nucleotides* are repeated in the 5′ and 3′ noncoding regions of p*eno*8 and p*eno*46, respectively. Two homologous nucleotide sequences in the 5′ noncoding region of p*eno*8 are indicated by *brackets*.

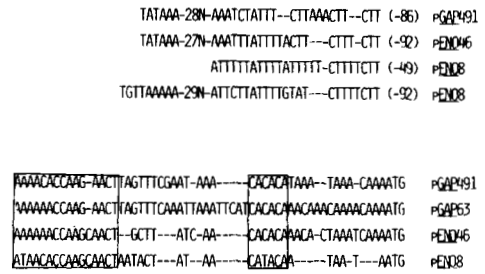

FIG. 8. **Comparison of the 5′ noncoding flanking regions of two glyceraldehyde-3-phosphate dehydrogenase structural genes and two enolase structural genes.** The 5′ noncoding flanking sequences of p*gap*491, p*gap*63, p*eno*8, and p*eno*46 are aligned to maximize homology among the sequences. Two regions of extensive sequence homology among all four genes are indicated by the *boxed sequences*. The *numbers within parentheses* refer to the number of nucleotides upstream from the ATG initiation codon and the number of nucleotides downstream from the TATAAA sequence, respectively.

the enolase genes and for the yeast glyceraldehyde-3-phosphate dehydrogenase structural genes (12), the nontandemly repeated genes share extensive nucleotide sequence homology within the first 100 nucleotides upstream from the translational initiation codon. We have analyzed the 5′ noncoding regions of the enolase and glyceraldehyde-3-phosphate dehydrogenase genes to determine if these glycolytic genes share homology. As illustrated in Fig. 8 there is extensive nucleotide sequence homology among the enolase and glyceraldehyde-3-phosphate dehydrogenase genes within the 5′ noncoding regions. Three regions of extensive nucleotide sequence homology are indicated in Fig. 8. The sequence -CACACA- is homologous among all four genes and is also found at approximately the same position relative to the translational initiation codon in the yeast cytochrome C-1 structural gene (15). Beyond this sequence is a 16-nucleotide segment of DNA which is highly conserved in the four glycolytic genes. We do not find this sequence within the 5′ noncoding region of either of the yeast cytochrome *c* genes (16). A third region of homology is present 30 nucleotides downstream from the -TATAAA- sequence in the p*gap*491 glyceraldehyde-3-phosphate dehydrogenase gene and the p*eno*46 enolase gene. This nucleotide sequence is homologous to two sequences in p*eno*8 as illustrated in Fig. 8. This latter nucleotide sequence is located approximately 90 nucleotides upstream from the initiation codon in p*gap*491 and p*eno*46 and is repeated 50 and 90 nucleotides upstream from the initiation codon in p*eno*8. This sequence is not present in the second glyceraldehyde-3-phosphate dehydrogenase gene (p*gap*63). Based on the extensive homology between the 5′ noncoding regions of these two sets of glycolytic genes, it appears that the 5′ noncoding regions may have evolved from a common precursor and that the sequences have undergone divergent evolution.

DISCUSSION

Two segments of yeast genomic DNA have been isolated by subculture cloning techniques which form hybrids with cDNA synthesized from partially purified yeast enolase messenger RNA. Based on the primary structures of these isolated hybrid plasmids each segment of DNA encodes a polypeptide with an amino acid sequence which is in close agreement with that determined by Chin *et al.* (17) for a chromatographically homogeneous form of yeast enolase. The primary structure of the polypeptide encoded by one of the isolated hybrid plasmids (p*eno*46) is identical with that determined by Chin *et al.* (17) in all but 9 residues. Careful examination of all of the nucleic acid and protein sequence data (17) suggests that four or five of these differences in amino acid sequence may in fact correspond to real differences in the primary structures of the enolase polypeptides in the yeast strains from which the protein and the genes were isolated. The primary structures of the enolase polypeptides encoded by the two isolated structural genes differ in 20 of 436 residues. Interestingly, none of the discrepancies between the amino acid sequence of yeast enolase and that predicted from the structural gene contained in p*eno*46 can be reconciled with the polypeptide sequence predicted from the gene contained in p*eno*8. Based on this observation we conclude that the enolase dimer which was sequenced by Chin *et al.* (17) is not a mixture of polypeptides derived from both enolase genes but rather that the protein is probably a dimer of the polypeptide encoded by the structural gene contained in p*eno*46. The presence of two enolase structural genes per haploid yeast genome raises the possibility that true isozymes of enolase exist in the cell. As yet there is no direct evidence demonstrating the presence of the polypeptide encoded by the gene in p*eno*8 in yeast cells, however.

Based on the primary structures of the enolase structural genes, both hybrid plasmids contain all of the sequences which would be complementary to cDNA synthesized from enolase mRNA. For this reason it is necessary to consider the fact that p*eno*46 and p*eno*8 do not hybridize to the same extent with the cDNA and that a mixture of the two hybrid plasmid DNAs hybridizes to the same extent as p*eno*8 alone. We believe that the molecular basis for these observations is related to the 3′ noncoding regions of the two structural genes. Based on DNA blotting experiments with restriction endonuclease cleavage fragments generated from the 3′ noncoding regions of the two genes, ^{32}P-labeled enolase cDNA hybridizes with both genes approximately 200 to 300 nucleotides beyond the respective termination codons.[3] The 3′ noncoding regions of the two genes are partially homologous; however, they would not be expected to cross-hybridize under stringent hybridization conditions. Since the cDNA is initiated within the 3′ noncoding regions of the mRNA and the average size of the cDNA is approximately 600 nucleotides, the 3′ noncoding portions of each mRNA are highly represented in the cDNA population. We conclude, therefore, on the basis of the hybridization data shown in Fig. 1 that the mRNA sequences derived from p*eno*46 are not as abundant in the yeast cell as are those derived from the gene contained in the p*eno*8 plasmid. These data suggest that there are differences in the rates of transcription of the genes and/or the half-lives of the two mRNAs. Since the mRNA encoded by the structural gene contained in p*eno*8 appears to be the most abundant enolase mRNA in the cell and the form of enolase which was sequenced by Chin *et al.* (17) appears to be a dimer of the polypeptide encoded by the gene contained in p*eno*46 it is likely that both enolase genes are expressed during vegetative yeast cell growth.

The coding regions of the two enolase structural genes are 95% homologous as are the coding regions of two yeast glyceraldehyde-3-phosphate dehydrogenase structural genes (12) indicating that these two duplicated glycolytic genes have diverged to the same extent. Divergence of the amino acid sequences of the polypeptides encoded by the enolase and glyceraldehyde-3-phosphate dehydrogenase genes is also 5%. The codon usage pattern in the enolase genes is extremely similar to the biased usage pattern previously reported for the two yeast glyceraldehyde-3-phosphate dehydrogenase genes (12). Comparison of the codon usage patterns of all four glycolytic genes reveals that cysteine, glutamine, glutamate, histidine, and glycine are encoded by a single codon while alanine, aspartate, isoleucine, serine, threonine, and valine are

[3] M. J. Holland and J. P. Holland, unpublished observations.

exclusively encoded by two codons which contain either C or U in third position. Arginine, asparagine, phenylalanine, proline, and tyrosine are encoded by a single codon with only 10 exceptions out of 258 codons. Ninety per cent of the lysines are encoded by AAG, and 95% of the leucines are encoded by UUG. The most abundant codons utilized in the two yeast cytochrome *c* genes are the same as those observed in the glycolytic genes; however, the degree of the bias in the cytochrome *c* genes is significantly lower (16). The molecular basis for the observed codon bias remains unclear; however, it may be relevant to the efficiency of translation of the glycolytic mRNAs. This hypothesis is consistent with the fact that enolase and glyceraldehyde-3-phosphate dehydrogenase are among the most abundant enzymes in yeast cells (2, 9, 18).

Based on the divergence of the nucleotide sequences of the repeated enolase and glyceraldehyde-3-phosphate dehydrogenase genes it is possible to estimate the age of the gene duplication. Utilizing the estimated rate of divergence of glyceraldehyde-3-phosphate dehydrogenase, for example (1% amino acid sequence divergence per 20 million years (19)) one would predict a duplication event which is approximately 100 million years old. This estimate is only valid, however, if the repeated yeast genes evolve independently. In the case of tandemly repeated genes there is considerable evidence that gene repeats do not evolve independently. Recent studies on the yeast ribosomal cistrons suggest that conservation of the primary structures of each cistron within the tandemly repeated array is achieved by unequal crossing over between tandem repeats during miosis (20). These crossing over events result in conservation of the primary structure of each cistron, and as a consequence the entire gene repetition evolves in concert. Concerted evolution of the tandemly repeated genes which code for the α chains of hemoglobin has also been proposed as an explanation for the fact that the α genes in a single primate species diverge in primary structure at one-tenth the rate observed among different primate species (21). Since the yeast enolase and glyceraldehyde-3-phosphate dehydrogenase genes are not tandemly repeated one might argue that the genes do not recombine during evolution. Alternatively, these genes may recombine by mechanisms which are different from those which occur for tandemly repeated genes. Multiple forms of enolase and glyceraldehyde-3-phosphate dehydrogenase have been resolved from a wide variety of eucaryotic cells (1, 22) suggesting that multiple genes for these glycolytic enzymes are present in these cells. In the case of glyceraldehyde-3-phosphate dehydrogenase the primary structures of the enzymes from yeast, lobster, and pig are 60% homologous (23). If duplicated glyceraldehyde-3-phosphate dehydrogenase genes existed in a common precursor to these eucaryotes, the estimated age of the duplication event based on the primary structures of the two yeast genes is a significant underestimation. Given these considerations, concerted evolution of the yeast genes cannot be ruled out. Concerted evolution of the yeast enolase and glyceraldehyde-3-phosphate dehydrogenase gene repeats would provide an explanation for the fact that divergence at the third position of the codons for the six amino acids which are encoded by two codons is less than 20% of that predicted for random drift at these positions. Such a mechanism would also explain the fact that unusual codons such as the CGU arginine codon at residue 14 in the enolase genes are conserved at the same position in both genes and that the 5′ and 3′ noncoding flanking regions of the gene repeats are 70% homologous.

The 5′ and 3′ noncoding flanking regions of the two enolase structural genes are approximately 70% homologous for 100 nucleotides upstream from the translational initiation codon and 280 nucleotides beyond the translational termination codon. Similar homology was observed in the 5′ and 3′ noncoding flanking regions of two yeast glyceraldehyde-3-phosphate dehydrogenase genes (12) and the genes which encode the β chains of mouse hemoglobin (24). Maximizing homology between the enolase genes within the noncoding regions necessitates leaving gaps in one of the gene sequences. It is assumed that these gaps arise as a consequence of deletions and/or insertions within one of the gene sequences. Similar gaps are generated when the yeast glyceraldehyde-3-phosphate dehydrogenase (12) and mouse β globin gene sequences are aligned (24). The gene contained in p*eno*46 contains a TATAAA sequence 140 nucleotides upstream from the initiation codon while this hexanucleotide is not present in the 5′ noncoding region of the gene contained in p*eno*8. The TATAAA sequence was found in the 5′ noncoding regions of both of the yeast glyceraldehyde-3-phosphate dehydrogenase genes (12) and in one of the yeast cytochrome *c* genes (15). Both enolase genes contain an AATAA sequence 220 to 240 nucleotides beyond the respective termination codons. Finally, short nucleotide sequence repetitions are observed in the 5′ and 3′ noncoding regions of the two enolase genes. Similar noninverted sequence repetitions were observed in the two yeast glyceraldehyde-3-phosphate dehydrogenase genes (12); however, these latter sequences are not homologous to those found in the enolase genes.

The 5′ and 3′ noncoding flanking regions of the enolase and glyceraldehyde-3-phosphate dehydrogenase (12) structural genes were compared to determine if homologous sequences were shared by these two glycolytic gene repeats. Little if any significant homology was found within the 3′ noncoding regions of the genes; however, the 5′ noncoding regions of all four genes are extremely similar for approximately 100 nucleotides upstream from the respective translational initiation codons. As illustrated in Fig. 8, all four genes contain three regions of extensive nucleotide sequence homology. There is a hexanucleotide adjacent to the initiation codon which is homologous in all four genes and is also present at a similar position in the yeast cytochrome *c*-1 gene (15). Beyond the hexanucleotide is a 16-nucleotide long region which is shared by the four glycolytic genes but is not found in the yeast cytochrome *c* genes (16). A third region of extensive homology is located approximately 90 nucleotides upstream from the initiation codons in the two enolase genes and the glyceraldehyde-3-phosphate dehydrogenase gene contained in the p*gap*491 hybrid plasmid (11). This latter sequence is located 30 nucleotides downstream from the TATAAA sequence in the genes contained in the p*eno*46 and p*gap*491 hybrid plasmids. This sequence is repeated twice in the enolase gene contained in p*eno*8, 49 and 92 nucleotides upstream from the translational initiation codon. The nucleotide sequences of the enolase and glyceraldehyde-3-phosphate dehydrogenase 5′ noncoding regions are also homologous within the regions between the sequences described above suggesting that the entire 5′ noncoding region of all four genes evolved from a common nucleotide sequence. Since enolase and glyceraldehyde-3-phosphate dehydrogenase are nonhomologous proteins, this common nucleotide sequence must have been fused to the coding regions of these two glycolytic genes during evolution. The similarities in nucleotide sequences within the 5′ noncoding regions of enolase and glyceraldehyde-3-phosphate dehydrogenase genes are in marked contrast to the fact that the 5′ noncoding regions of the coordinately expressed mouse α and β globin structural genes are nonhomologous (24, 25). While the structures of the mouse globin genes suggest that conservation of nucleotide sequence is not required for coordinate expression it is tempting to speculate that the homologous 5′ noncoding regions of the yeast enolase

and glyceraldehyde-3-phosphate dehydrogenase genes are involved in coordinate expression of these glycolytic genes. This possibility can now be tested by *in vitro* alteration of these sequences and subsequent analysis of the expression of these altered genes *in vivo*.

Acknowledgments—We would like to express our appreciation to Drs. Chin, Brewer, and Wold for making available to us their unpublished amino acid sequence of yeast enolase and for helpful discussions. We are indebted to Ms. Laura Labieniec for excellent technical assistance.

REFERENCES

1. Wold, F. (1971) in *The Enzymes* (Boyer, P. D., ed) 3rd Ed. pp. 499–538, Academic Press, New York
2. Westhead, E. W., and McLain, G. (1964) *J. Biol. Chem.* **239,** 2464–2468
3. Oh, S.-K., Travis, J., and Brewer, J. M. (1973) *Biochim. Biophys. Acta* **310,** 421–429
4. Porcelli, L. J., Jr., Small, E. D., and Brewer, J. M. (1978) *Biochem. Biophys. Res. Commun.* **82,** 316–321
5. Pfleiderer, G., Neufahrt-Kreiling, A., Kaplan, R. W., and Fortnagel, P. (1966) *Biochem. Z.* **346,** 269–281
6. Susor, W. A., Kochman, M., and Rutter, W. J. (1969) *Science* **165,** 1260–1262
7. Shanbhag, V., Blomquist, G., Johansson, G., and Hartman, A. (1972) *FEBS Lett.* **22,** 105–108
8. Blomquist, G. (1976) *Biochim. Biophys. Acta* **420,** 81–86
9. Holland, M. J., and Holland, J. P. (1978) *Biochemistry* **17,** 4900–4907
10. Holland, M. J., and Holland, J. P. (1979) *J. Biol. Chem.* **254,** 5466–5474
11. Holland, J. P., and Holland, M. J. (1979) *J. Biol. Chem.* **254,** 9839–9845
12. Holland, J. P., and Holland, M. J. (1980) *J. Biol. Chem.* **255,** 2596–2605
13. Smith, H. O., and Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3,** 2387–2398
14. Maxam, A. M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 560–564
15. Smith, M., Leung, D. W., Gillam, S., Astell, C. R., Montgomery, D. L., and Hall, B. D. (1979) *Cell* **16,** 753–761
16. Montgomery, D. L., Leung, D. W., Smith, M., Shalit, P., Faye, G., and Hall, B. D. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 541–545
17. Chin, C. C. Q., Brewer, J. M., Eckard, E., and Wold, F. (1981) *J. Biol. Chem.* **256,** 1370–1376
18. Krebs, E. G. (1953) *J. Biol. Chem.* **200,** 471–478
19. Wilson, A. C., Carlson, S. S., and White, J. (1977) *Annu. Rev. Biochem.* **46,** 573–639
20. Petes, T. D. (1980) *Cell* **19,** 765–774
21. Zimmer, E. A., Martin, S. L., Beverly, S. M., Kan, Y. W., and Wilson, A. C. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 2158–2162
22. Lebherz, H. G., and Rutter, W. J. (1967) *Science* **157,** 1198–1200
23. Olsen, K. W., Moras, D., Rossmann, M. G., and Harris, J. I. (1975) *J. Biol. Chem.* **250,** 9313–9321
24. Konkel, D. A., Maizel, J. V., Jr., and Leder, P. (1979) *Cell* **18,** 865–873
25. Nishioka, Y., and Leder, P. (1979) *Cell* **18,** 875–882

Additional references are found on p. 1395.

SUPPLEMENTARY MATERIAL

TO

The Primary Structures of Two Yeast Enolase Genes:
Homology between the 5' Noncoding Flanking Regions
of Yeast Enolase and Glyceraldehyde-3-Phosphate Dehydrogenase Genes

BY

Michael J. Holland, Janice P. Holland, Gregory P. Thill and Kimberly A. Jackson

EXPERIMENTAL PROCEDURES

Materials. [γ-^{32}P]ATP (2000-3000 Ci/mmol) was purchased from New England Nuclear. Restriction endonucleases *Alu* I, *Ava* II, *Bam*HI, *Hin*d III, *Hin*f I, *Hpa* II, *Sal* I, *Xho* I, *Xba* I, *Bgl* II, *Sma* I, *Tac* I, *Taq* I and bacterial alkaline phosphatase were obtained from Bethesda Research Laboratories; *Bgl* I, *Eco* RI, *Hae* III, *Hha* I, *Mbo* II, *Pvu* II, *Sac* I were from New England Biolabs. Polynucleotide kinase was purchased from PL Biochemicals. Bacterial strains and growth conditions were as described in Holland and Holland (1). Plasmid DNA was prepared according to the procedure of Clewell and Helinski (2).

Subculture Cloning of *peno*8 and *peno*46. Subculture cloning of *E. coli* transformants which contain two nontandemly repeated enolase structural genes was carried out by the method described for the isolation of a glyceraldehyde 3-phosphate dehydrogenase structural gene (1,3). Total hybrid plasmid DNA composed of the *E. coli* plasmid vector, pSF2124 (4), and randomly sheared yeast genomic DNA, was isolated from a series of subcultures which were innoculated with a known number of transformants. Plasmid DNA was depurinated and hybridized in solution with a [^{3}H]cDNA synthesizes from partially purified enolase mRNA (5) as described previously (1).

Restriction Endonuclease Mapping of *peno*8 and *peno*46. The restriction endonuclease cleavage maps of *peno*8 and *peno*46 were established by digesting 0.5-1 µg of hybrid plasmid DNA with the following enzymes for 1 hr at 37°C in 30 µl reaction volumes under the suppliers' prescribed conditions: *Bgl* I, *Bgl* II, *Eco* RI, *Hin*d III, *Hpa* I, *Kpn* I, *Pvu* II, *Sac* I, *Sma* I and *Xba* I.

Bgl II, *Hin*d III, *Hpa* I, *Kpn* I and *Xba* I do not cleave the vector *Sac* I and *Sma* I cleave pSF2124 once, *Bgl* I twice and *Pvu* II five times. The *Eco* RI cleavage site in pSF2124 was destroyed in the construction of the hybrid plasmids (1). The yeast segments of *peno*8 and *peno*46 are not cleaved by *Sal* I or *Sma* I. *peno*8 is cleaved once by *Bgl* I and *Bgl* II, twice by *Eco* RI, *Hpa* I, *Kpn* I and *Pvu* II and four times by *Hin*d III. Orientation of these sites relative to the vector was achieved by double digestion with either *Sac* I or *Sma* I. More precise measurements were made by multiple digestions with enzymes cleaving within the yeast portion of *peno*8. Table IS summarizes these data. The yeast portion of *peno*46 is not cleaved by *Bgl* I, or *Pvu* II, is cleaved once by *Kpn* I and *Xba* I twice by *Bgl* II, three times by *Eco* RI and five times by *Hin*d III. The orientation of these cleavage sites was established as described above and these data are summarized in Table IIS.

A more detailed restriction endonuclease cleavage map involving enzymes with four, five and six nucleotide recognition sites was established for *peno*8 and *peno*46 employing the partial digestion procedure of Smith and Birnstiel (6). The cleavage sites were oriented relative to the ^{32}P-end labeled *Bgl* II sites of both enolase genes. A series of partial digestions of the *peno*8 4.4 kb *Pvu* II/*Bgl* II fragment which contained the major portion of the enolase structural gene and the *peno*46 2.0 kb *Xba* I/*Bgl* II fragment containing the same segment of the structural gene were carried out in parallel (3000 cpm end labeled DNA per digestion) with *Alu* I, *Ava* II, *Eco* RI, *Hae* III, *Hin*f II, *Hin*d III, *Hin*f I, *Xba* I, *Hpa* I, *Hpa* II, *Kpn* I, *Mbo* II, and *Taq* I. The cleavage patterns were displayed on a 1.5% agarose slab gel (30 x 13.5 x 0.3 cm) and visualized by autoradiography with Kodak X-OMAT R film and a Dupont Lightning Plus intensifying screen for exposure times of 24-48 h.

Restriction maps extending in the opposite direction from the *Bgl* II cleavage sites in the plasmids were generated as follows: The 1.5 kb *Bgl* II/*Bgl* II cleavage fragment from *peno*46, ^{32}P-labeled at its 5' termini, was cleaved with *Hin*d III. The 1.3 kb *Bgl* II/*Hin*d III fragment was then analyzed by partial hydrolysis with restriction endonucleases as described above. The restriction endonuclease cleavage map of the 0.2 kb *Bgl* II/*Hin*d III fragment from *peno*46 was determined from the nucleotide sequence of the fragment. A 0.75 kb *Bgl* II/*Pvu* II fragment, labeled at the *Bgl* II cleavage site, was used to generate the cleavage map of *peno*8 by the partial cleavage method described above.

Southern Blotting Analysis of *peno*8 and *peno*46. 0.5 µg each of *peno*8 and *peno*46 and 2.5 µg of yeast DNA strain F1, grown commercially and prepared by the method of Holland et al (7), were limit digested with the restriction enzymes *Xba* I, *Xho* I, *Eco* RI and *Hin*d III and the cleavage patterns displayed on a 0.8% agarose slab gel electrophoresed at 60 V for 5 hr. 2.5 µg of DNA prepared from strains F1 and D273-10B grown in YPD media (2% peptone, 1% yeast extract and 2% glucose) and strain F1 grown commercially by The Red Star Corp. were limit digested with *Xba* I and electrophoresed in parallel as described above. After electrophoresis DNA was transferred to a nitrocellulose filter by the modified procedure of the Southern technique (8) outlined by Ketner and Kelley (9). The filter was pretreated for 3 h at 65° in 3 X SSC (1 X SSC = 0.15 M sodium chloride, 0.015 M sodium citrate, pH 7.0) containing Denhardt's solution (10): 0.02% Ficoll, 0.02% polyvinylpyrolidine, and 0.02% bovine serum albumin. The filter was then pretreated for 1 h at 65° with the above solution supplemented with 0.1% sodium dodecyl sulfate and 10 µg/ml depurinated salmon sperm DNA and blotted dry. Hybridization was performed in the latter solution at 65° for 36 h under mineral oil utilizing 2.5 x 10^{5} cpm of a [^{32}P]dCTP labeled enolase cDNA (specific activity = 10^{9} cpm/µg) synthesized as outlined in Holland and Holland (1). Following incubation the filter was washed 3 times with chloroform at room temperature to remove the oil and 6 times with 300 ml 6 X SSC at 65° for a total of 90 min. Autoradiography was performed as previously described.

Complementary DNA synthesized from partially purified enolase mRNA was purified for DNA blotting experiments involving yeast genomic DNA. The 1.6 kb *Eco* RI fragment from *peno*46 was ligated into the *Eco* RI site of pSF2124. The resulting hybrid plasmid was linked to Sepharose 4B after sonication and denaturation by the procedure described by Gilboa et al (13). Enolase cDNA was purified by annealing the cDNA synthesized from partially purified enolase mRNA to this affinity column and subsequent elution with formamide as previously described (1).

DNA Sequence Analysis of *peno*8 and *peno*46. DNA fragments were prepared for sequence analysis by dephosphorylating with bacterial alkaline phosphatase and phosphorylating with γ-^{32}P ATP and polynucleotide kinase according to the procedure of Maxam and Gilbert (11). Unincorporated γ-^{32}P ATP was removed and at the same time end labeled DNA fragments resolved by preparative agarose or acrylamide gel electrophoresis as described previously (12). Secondary cleavages followed by preparative gel electrophoresis were performed and DNA fragments labeled at a single end were isolated. DNA sequence analysis was carried out according to the procedure of Maxam and Gilbert (11) employing 0.4 mm 20% (or 12%) polyacrylamide gel electrophoresis at 1100 V with a 7, 20, and 36 h (14 h) loading schedule. Autoradiography was carried out as previously described.

DNA sequence analysis of the coding regions of the *peno*8 enolase structueal gene was performed according to the strategies diagrammed in Figure 6 and involved endlabeling of the *Bgl* II, *Eco* RI, *Hin*d III and *Kpn* I cleavage sites. DNA sequence analysis of the 5' noncoding region of the *peno*8 enolase gene involved digesting the 1.7 kb *Sma* I/*Kpn* I fragment with *Hin*f I, endlabeling, isolating the 1.1 kb *Hin*f I/*Hin*f I fragment cleaving secondarily with *Hpa* II and sequencing from the *Hin*f I site at the extreme 5' end of the coding region. Analysis of the 3' flanking region of the *peno*8 enolase gene involved *Hin*f I digestion of the 2.7 kb *Hin*d III fragment and endlabeling the resulting fragments. The fragments ~ 0.3l and 0.19 kb in size were identified by secondary digestion with *Taq* I and *Tac* I, respectively, or containing 3' noncoding sequences and were isolated by preparative acrylamide gel electrophoresis. The 0.047 kb *Hin*f I/*Hin*f I fragment located 85 nucleotides from the ochre codon was sequenced from an endlabeled *Hpa* II fragment which included the 3' end of the structural gene.

Analysis of the coding region of the *peno*46 structural gene was performed according to the strategies diagrammed in Figure 6. In order to facilitate the isolation of restriction fragments for DNA sequence analysis, the 2.0 and 1.6 kb *Eco* RI fragments of *peno*46 were each ligated into the *Eco* RI site of p2124. These fragments were then prepared in quantity by *Eco* RI digestion and sucrose density gradient centrifugation. DNA sequence analysis of the enolase coding region represented in the 2.0 kb *Eco* RI fragment was accomplished by endlabeling the *Bgl* II, *Eco* RI and *Hin*d III cleavage sites. The 3' noncoding region of the *peno*46 enolase structural gene was analyzed by endlabeling the two *Hin*f I sites within 450 nucleotides of the ochre codon. DNA sequence analysis of the coding region contained in the 1.6 kb *Eco* RI fragment was achieved by endlabeling *Hin*d III, *Hpa* II and *Taq* I cleavage sites as diagrammed in Figure 6. Fragments generated by *Hpa* II and *Taq* I digestions also yielded DNA sequence information in the 5' flanking region of the *peno*46 enolase structural gene.

REFERENCES

1. Holland, M.J. and Holland, J.P. (1979) J. Biol. Chem. 254, 5466-5474.
2. Clewell, D.B. and Helinski, D.R. (1970) Biochemistry 9, 4428-4440.
3. Holland, J.P. and Holland, M.J. (1980) J. Biol. Chem. 255, 2596-2605.
4. So, M., Gill, R. and Falkow, S. (1975) Molec. Gen. Genet. 142, 239-249.
5. Holland, M.J. and Holland J.P. (1978) Biochemistry 17, 4900-4907.
6. Smith, H.O. and Birnstiel, M.L. (1976) Nuc. Acids. Res. 3, 2387-2398.
7. Holland, M.J., Hager, G.L. and Rutter, W.J. (1977) Biochemistry 16, 8-16.
8. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
9. Ketner, G. and Kelly, T.J., Jr. (1976) Proc. Natl. Acad. Sci. USA 73, 1102-1106.
10. Denhardt, D.T. (1966) Biochem. Biophys. Res. Comm. 23, 641-646.
11. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
12. Holland, J.P. and Holland, M.J. (1979) J. Biol. Chem. 254, 9839-9845.
13. Gilboa, E., Prives, C.L. and Aviv, H. (1975) Biochemistry 14, 4215-4220.

TABLE IS

Molecular Weight of Restriction Endonuclease

Cleavage Fragments of peno8

Restriction Endonuclease	Fragment Molecular Weight (kb)
Bgl I	7.9, 5.1, 2.3
Bgl II	15.3
Eco RI	15.1, 0.25
*Hin*d III	11.8, 2.7. 0.54, 0.24
Hpa I	14.3, 1.0
Kpn I	14.6, 0.75
Pvu II	5.2, 2.8, 2.5, 2.0, 1.45, 0.9, 0.45
Sac I	15.3
Sma I	15.3
Sma I/*Bgl* II	12.8, 2.5
Sma I/*Eco* RI	13.0, 2.0, 0.25
Sma I/*Hin*d III	10.2, 2.7, 1.6, 0.54, 0.24
Sac I/*Hin*d III	9.4, 2.7, 2.4, 0.54, 0.24
*Hin*d III/*Eco* RI	12.1, 2.7, 0.38, 0.16, 0.16, 0.08
Sma I/*Hpa* I	11.9, 2.4, 1.0
Sma I/*Kpn* I	12.9, 1.7, 0.75

TABLE IIS

Molecular Weight of Restriction Endonuclease

Cleavage Fragments of peno46

Restriction Endonuclease	Fragment Molecular Weight (kb)
Bgl I	13.5, 2.3
Bgl II	14.3, 1.5
Eco RI	12.2, 2.0, 1.6
*Hin*d III	14.6, 0.41, 0.40, 0.24, 0.13
Kpn I	15.8
Pvu II	9.4, 2.5, 2.0, 1.45, 0.45
Sac I	15.8
Sma I	15.8
Xba I	15.8
Sma I/*Bgl* II	10.3, 4.0, 1.5
Sac I/*Bgl* II	11.9, 2.4, 1.5
Sma I/*Eco* RI	10.2, 2.0, 2.0, 1.6
Sac I/*Eco* RI	10.0, 2.2, 2.0, 1.6
Bgl II/*Eco* RI	12.2, 1.6, 1.45, 0.35, 0.2
Sma I/*Hin*d III	11.5, 3.1, 0.41, 0.40, 0.24, 0.13
Bgl II/*Hin*d III	13.3, 1.3, 0.41, 0.24, 0.20, 0.20, 0.13
Eco RI/*Hin*d III	12.2, 1.4, 1.0, 0.41, 0.40, 0.16, 0.13, 0.08
Sma I/*Kpn* I	11.9, 3.9
Eco RI/*Kpn* I	12.2, 1.6, 1.5, 0.5
Sma I/*Xba* I	13.8, 2.0