Sequence of a *Drosophila* DNA segment that functions in *Saccharomyces cerevisiae* and its regulation by a yeast promoter

Steven Henikoff* and Clement E. Furlong⁺

*Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104, and

⁺ Departments of Genetics and Medicine, Division of Medical Genetics Center for Inherited Diseases, University of Washington, Seattle, WA 98195, USA

Received 1 November 1982; Revised and Accepted 3 January 1983

ABSTRACT

We have determined the complete DNA sequence of a segment derived from *Drosophila melanogaster* that complements a yeast adenine-8 (ade8) mutation. This sequence, in conjunction with transcriptional analysis, predicts that a 206 amino acid sequence accounts for complementation. We have ligated a 1.5 kb promoter segment of the yeast alcohol dehydrogenase I (ADHI) gene just upstream of the open reading frame. This construction results in ADE8+ function that responds to glucose and ethanol in a manner similar to that for the chromosomal ADHI gene, as determined by a simple quantitative color assay for ADE8+ function.

INTRODUCTION

The isolation of genes from higher eukaryotes has been accomplished primarily by reverse transcription of mRNA sequences coding for proteins that are abundant in certain cell types. In contrast, few higher eukaryotic genes have been cloned that are normally expressed at low levels. This is unfortunate since the latter situation holds for the bulk of the genetic information that encodes proteins. The question arises whether the features found for abundantly expressed genes are typical of higher eukaryotic genes in general, or rather just the subset of genes coding for abundant proteins in differentiated cells. We previously demonstrated that at least one sequence with metabolic function could be isolated from a higher eukaryote using genetic complementation in the yeast *Saccharomyces cerevisiae* (1). A pool of cloned genomic segments was used to transform mutant adenine-8 cells allowing isolation of a *Drosophila* DNA segment that complemented the defect. Here we present the complete DNA sequence of the functional segment. We show that the expression and regulation of this segment can

be modified by ligation of a well-characterized yeast promoter upstream of the predicted protein sequence.

EXPERIMENTAL PROCEDURES

Construction of pYFADE8 plasmids and cloning in S. cerevisiae and E. coli have been described (1). Restriction fragments were 3' end-labeled with ^{32}P using the large E. coli DNA polymerase I fragment (2) or were labeled at the 5' ends using alkaline phosphatase and T4 DNA kinase (3). These were either strand-separated or cut with restriction enzymes and sequenced by the method of Maxam and Gilbert (3). Sequence information from both strands was obtained for 92% of the sequence.

Plasmid DNA was used to transform JK24 (a ade2 ade8 leu2 trp1) yeast, selecting for LEU2 as described by Beggs (4). Transformants that had been streaked onto synthetic complete plates lacking leucine were used to inoculate overnight cultures containing 3 ml of a synthetic complete-minus leucine medium (5). This medium has been modified by reducing the amount of adenine by one-half; the resulting medium was found to give enhanced red color of ade2 controls. Experimental tubes included either 8% glucose or 3% ethanol + 0.1% glucose as carbon source. Fresh overnight cultures in 0.25% glucose were pelleted and resuspended in sterile water. Aliquots of 0.05 ml were used to inoculate experimental tubes. Cultures were incubated at room temperature at a 30° angle on a New Brunswick Gyrotary Shaker Model G2 at 300rpm. Standards were prepared by mixing an equal volume of fully red cells (score of 10) with fully white cells (score of 0) to give a culture with a score of 9. Equal volumes of these cells and fully white cells gave a culture with a score of 8, and so on. Three ml cultures in glass tubes were centrifuged and pellets were compared to those of 3 ml standards by transmitted light. Tubes were vortexed to resuspend the cells and replaced on the shaker after each determination. Pigment began to accumulate when cells approached maximum density, allowing color to be estimated from pellets of uniform size.

RESULTS & DISCUSSIONDNA sequence and transcription of *Drosophila* adenine-8 in yeast

It was previously shown that a single-copy segment derived from subdivision 27C of the *D. melanogaster* genome and carried on a recombinant DNA plasmid (pYFADE8) complemented a yeast adenine-8 mutation (1). The complementation behavior of subclones derived from this plasmid narrowed the functional region to a 1.6 kilobase (kb) segment defined by a BglIII site on one side and a BamHI site on the other side. Figure 1 shows the complete sequence of this functional segment. The BamHI site located at -63 is important in the analysis that follows.

Previous work showed that a poly(A)⁺ RNA transcript extracted from yeast cells carrying a pYFADE8 plasmid begins very near the BamHI site at -63 and ends about 900 bases to the right. No intervening sequence was detected. We have extended this transcriptional analysis in order to determine more precisely the start of transcription in yeast. A 448 bp TaqI DNA fragment spanning the vicinity of transcriptional initiation was used in an S1-nuclease protection experiment. Poly(A)⁺ RNA was extracted from cells carrying PYFADE8(2.0) (1) which includes the entire functional region. Protection by the TaqI fragment from S1 nuclease digestion and electrophoresis on a sequencing gel indicates that transcription starts are mostly clustered between -126 and -113, upstream of the BamHI site at -63 (Figure 2, lane A). PolyA⁺ RNA also was extracted from cells carrying just the 1.2kb BamHI fragment cloned in an orientation that failed to complement ade8 (1). Transcription of this segment would give rise to an S1-protected fragment migrating at the position corresponding to -63 on the gel. No significant protected fragment could be detected (Figure 2, lane B), as if removal of *Drosophila* sequences upstream of the BamHI site at -63 essentially abolished transcription. These results map the *Drosophila* ADE8 "promoter" to the region upstream of -63.

About 50 base pairs upstream of the first transcription initiation site is the sequence TATAA which is similar to the "TATA box" sequence found just upstream of most eukaryotic genes that have been sequenced (6). It appears that in *Drosophila*,

BglII
 5' AGATCTATGTTAAGGCATTGTCCACTCTGCTGTCCCGCGGGAACACGGCATCAAGGCACTAGCCACATCAGGAGGAGGCCTGAGCGAGAATATA
 -300
 CCGCGAGTATTGCGTAAGGATTGGCCGTACGCTGGATGCCAATAAGTTCCAGCTTCCGCCCTTTTTCGCTGGCTGCAGCGGCCGGAATATCAGCT HpaII
 -200
 CCAGTCTAGCTGCAGCGCACATATAACTGCGGCTTGGGTTAGTTCTGGTGGTGGCCCCGAGGAAGTCGAGGACGTTCTTAAGGAGCTGCGTTATCCGCA TaqI
 -100
 AAGAGCAGCTGTGGTCCGCGAGGTGGTGGCACGCAAGGATCCGAAGAAGTCGAGGTGGTGGTCCAGAACTTTGAGGCATCATTTGCCAGGACACAGAAG
 BamHI AvaII
 1
 MetLeuSerGlnArgArgLysArgValAlaValLeuIleSerGlyThrGlySerAsnLeuGlnAlaLeuIleAspAlaThrArgAspSerAlaGlnGlyI
 ATGCTCTCGCAGCGTCGGAACAGAGTTGCTGTCTCATTTCGGGAACGGGCAGCAACTGCAAGCACTAATTGATGCAACGCGGACTCGGCACAGGGTA
 HpaII
 200
 leHisAlaAspValValLeuValIleSerAsnLysProGlyValLeuGlyLeuGlnArgAlaThrGlnAlaGlyIleProSerLeuValIleSerHisLy
 TCCATGCCGATGTTGTGCTGGTAATCAGCAATAAGCCTGGTGTCTGGGTCTGCAGCGAGCTACGCAAGCCGGAATACCTTCGTTGGTCATCTCCACAA
 HpaII
 300
 sAspPheAlaSerArgGluValTyrAspAlaGluLeuThrArgAsnLeuLysAlaAlaArgValAspLeuIleCysLeuAlaGlyPheMetArgValLeu
 GGATTCGCTAGCCGAGAGGTTTATGACGCGGAACTAACGCGAATCTGAAGCGGCTCGTGTGGATTGATCGCTGCCTGCGGCTTCATGCGTTCCTG
 TaqI
 400
 SerAlaProPheValArgGluTrpArgGlyArgLeuValAsnIleHisProSerLeuLeuProLysTyrProGlyLeuHisValGlnLysGlnAlaLeuG
 AGTGTCTCCATTTCGTCGAGAGTGGCGTGGACGGCTCGTCAATATCCATCCATCGTGTCTGCCAAGTATCTCGGCTGCATGTCAGAGCAGGCTTGG
 HpaII
 500
 luAlaGlyGluLysGluSerGlyCysThrValHisPheValAspGluGlyValAspThrGlyAlaIleIleValGlnAlaAlaValProIleLeuProAs
 AGGCGGGTGAAAAGGAGTCCGGCTGCACGGTTCACTTCGTGGACGAGGGTGTGGACACGGGAGCGATCATTTGTCAGGCTGCTGTTCCCATTTTGGCCAGA
 600
 pAspAspAlaAspSerLeuThrGlnArgIleHisLysAlaGluHisTrpAlaPheProArgAlaLeuAlaMetLeuValAsnGlyThrAlaLeuIleSer
 TGACGACGAGGATTTCGCTGACCCACGATCCACAAGGCTGAGCACTGGGCTTTTCAAGGGCATTAGCAATGTGGTCAACGGAACAGCTCTCATTTCT
 HpaII
 700
 ProGluValSerSerGlnEnd
 CCGGAAGTCAGCAGTCAGTAAATGCCAGCTTGTGAAACCTGACAATTTAAATGGAACAAGCAGGGATGAGTTTATCTTACTCGAAGACTTTTGTCAAAT
 TaqI
 800
 TAGTTATAATTATACTTGAACATTCAACGATAATTTGACGTTAATAAGTTGTATAGCAATTTTATAGGTTTAGGCTGCGTTTCTTTTATATAAAAA
 3' poly(A)
 900
 CTACAAAAATATATTAATAAACTTCAAGTTTGATTAAAAAAATCGCAAAAAATTTGTACACATGAACATACTATATAAAAAAATCCAATATTTTTT
 TaqI
 1000
 ATCACTCGATTTCACAGAAAGAACATATCAATGATTTTAATGTTACAATAACAAAAAGGTCTCTTGGATTATTGAACATTTTCGTACATATATGTACAC
 1100
 ATAAACGTACAGCACCTTAATCTACATCTAATATCACATATTTTACAATCACACATTTTAAATACAATAGTTTGCTAATTAGGAGAATAAACTTGTAC
 XbaI BamHI
 TAGCGCTGCCCTGATTCTCTAGATTGCTGCTGGGCTTGATTAAATTCGGGTCTGATTACGTTACGTTGCAATTCGGGATCC

Figure 1 Complete DNA sequence of the *Drosophila melanogaster* ADENINE-8 segment that functions in yeast. The ends of message were determined in the S1 nuclease digestion experiments shown in Figures 2 and 3. Multiple sites are found at both ends as indicated by arrows. Beginning with the first ATG in the transcribed region, a 206 amino acid protein is predicted. Restriction sites used in the sequencing analysis are shown.

however, the 5' end of this gene is found even farther upstream (preliminary observations), as if initiation within this region does not occur for the segment in its natural state.

Transcription appears to terminate within a region that is 75% A + T. Figure 3 shows an S1 nuclease protection experiment that maps poly(A) addition. A cluster of 3' polyadenylated termini are seen between nucleotides 830 and 850 (lane A). These

Figure 2 Mapping of transcription initiation by S1 nuclease protection. The TaqI restriction fragment spanning the sequence from -133 to +315 (Figure 1) was 5' end-labeled, hybridized to poly (A)⁺ RNA from cells carrying pYFADE8(2.0), and S1 nuclease-treated as described (1), where conditions allowed RNA-DNA, but not DNA-DNA hybridization. The resulting hybrid (A) was electrophoresed on a 6% acrylamide sequencing gel (3) and autoradiographed. As a control (B), the TaqI fragment was hybridized to an equivalent amount of poly(A)⁺ RNA from cells carrying pYFADE8(1.2R), which is the 1.2 kb BamHI Drosophila DNA fragment subcloned into YEpl3 in the orientation that gives no detectable growth under adenine selection (1). Arrows indicate the expected migration distances of the untreated Taq I fragment (-133) and of the same fragment reduced in size by 70 bases using BamHI (-63). Three major initiation sites mapped in A are indicated in Figure 1.

occur about 10 to 30 base pairs beyond the sequence AATAAA, which is thought to signal cleavage and polyadenylation in higher eukaryotes (7). Preliminary results suggest that, unlike initiation, termination of transcription occurs in the same region for both yeast and Drosophila.

The first transcribed ATG is followed by the longest open reading frame in the sequence. It predicts a 206 amino acid polypeptide. Such a long open reading frame is expected to occur by chance from the first transcribed ATG with a probability of 0.00005, so that this predicted polypeptide is presumed to account for adenine-8 complementation. Since adenine-8 mutants accumulate glycineamide ribonucleotide (GAR)

Figure 3 Mapping of transcription termination by S1 nuclease protection. The HpaII restriction fragment spanning the sequence from 601 to 1177 (Figure 1) was 5' end-labeled, strand-separated, and the anti-sense strand was hybridized to poly (A)+ RNA from cells carrying either (A) pYFADE8(2.0), or (B) pADHAD8. As a control, this fragment was hybridized to *E. coli* tRNA. (C) Each mixture included 2 ng DNA and 2.5 μ g RNA in a 10 μ l reaction volume containing 2% formamide, 0.4M NaCl, 50mM PIPES pH6.4, 1mM EDTA. After 14 hr at 48 $^{\circ}$, each mixture was treated with 40 units S1 nuclease (Sigma) in a 100 μ l volume for 30 min at 22 $^{\circ}$ (1). Following ethanol precipitation, samples were electrophoresed on a sequencing gel and autoradiographed. Arrows indicate the positions for fragments corresponding to 825 and 850 on the sequence (Figure 1).

and since mutants in all other reactions in the first part of the de novo purine biosynthetic pathway have been identified (8), the 206 amino acid polypeptide is expected to be GAR transformylase. In support of this expectation, P. Russell (personal communication) has found that both Drosophila and yeast ADE8 plasmid derivatives will complement a Schizosaccharomyces pombe ADE5 mutation which blocks GAR transformylase in this organism.

Control of Drosophila adenine-8 by the yeast ADHI promoter.

We had previously shown that the 334bp segment between the BglIII and BamHI sites within which transcription initiates is required for complementation (1). Since translation appears to

initiate downstream from this segment, it seemed possible that another promoter segment could be inserted in place of the 334 bp segment and restore ADE8 function. Figure 4 shows such a substitution. Plasmid pYFADE8(1.6R) is the 1575bp Drosophila functional segment inserted into the yeast cloning vector YEpl3. Plasmid pADHAD8 is a construction in which the 334bp Drosophila "promoter" segment has been replaced by a 1.5kb promoter segment derived from the yeast ADHI gene that has been shown to promote transcription of human leukocyte interferon (9). Sequence analysis of pADHAD8 reveals that the ADHI promoter end point is

Figure 4 Substitution of a yeast ADHI promoter fragment for the Drosophila ADE8 "promoter". BamHI digestion of pYFADE8(1.2R) (1) yielded a mixture of fragments, half of which were broken between the BamHI site at -63 and that of the YEpl3 vector. A pool of ca. 1.5kb ADHI promoter fragments flanked by BamHI linkers (9, G. Ammerer, personal communication) was ligated into this vector and used to transform E. coli strain RRI cells. Plasmid DNA from the resulting pool of transformants was used to transform ade8 leu2 yeast cells (5028-1D α), selecting for ADE8⁺ LEU2⁺. One such transformant was cloned, used to isolate plasmid which was then used to transform E. coli RRI. Restriction and sequence analysis showed that this plasmid, pADHAD8, was the fusion indicated in the Figure. For comparison, pYFADE8(1.6R) is also depicted; the complete sequence of its insert is shown in Figure 1. Arrows above the junction sequence indicate transcription initiation sites for the ADHI promoter (10). B, BamHI; X, XbaI; wavy arrows, transcription; open boxes, inserts; hatched area, open reading frame.

identical to that of a promoter fragment used in the interferon study and found to be relatively active in interferon synthesis.

When pADHAD8 was used to transform an ade8 strain (5028-1D α), growth on plates and in liquid medium lacking adenine was similar to that for pYFADE8(1.6R) and pYFADE8(2.0) tested previously (1). In order to determine whether ADE8 function of pADHAD8 transformants resulted from abundant transcription of the Drosophila segment, we isolated poly(A)⁺ RNA. Figure 3 (lane B) shows that ADE8 3' termini are as readily detected when the gene is controlled by the ADHI promoter as when controlled by the Drosophila "promoter". These results indicate that the ADHI promoter fragment is active in functional transcription of the Drosophila ADE8 sequence.

Alcohol dehydrogenase I is active during yeast fermentation, catalyzing the final step of glycolysis. Recently, Denis and co-workers (11) showed that as glucose is converted to ethanol in the growth medium, ADHI protein synthesis and mRNA levels decline about 5-10 fold. Their results indicate transcriptional regulation of the gene during the conversion from glucose to ethanol. A similar regulation appears to occur for another glycolytic enzyme, enolase (12). We wondered whether the ADE8 coding sequence controlled by the ADHI promoter would be affected as well.

In order to determine whether the promoter fragment used in the construction of pADHAD8 is sensitive to a change from glucose to ethanol, we compared ADE8 function for this plasmid to that of pYFADE8(1.6R) (Figure 4). JK24 (a ade2 ade8 leu2 trp1) cells carrying either pYFADE8(1.6R) or pADHAD8 were grown under leucine selection in liquid cultures that included either 8% glucose or 3% ethanol. As a control, cells carrying a plasmid with the wild-type yeast ADE8 gene, pYeADE8(4.9) (1), were similarly grown.

To measure activity of the enzyme encoded by the Drosophila ADE8 open reading frame, we have adapted a red color screening method that has been a useful tool in other studies (13, 14). Adenine-2 cells accumulate a red pigment when adenine is limiting. Six genes for de novo purine biosynthesis, ade3 through ade8, control enzymes needed for the precursor utilized

by the ade2 gene product. Lesions in any of these six genes can prevent the synthesis of precursor and so prevent the accumulation of red pigment in ade2 cells. Since the pigment is inhibitory, ade3 through ade8 mutations that occur in ade2 cells have a growth advantage and form papillae on the red ade2 colonies. We have found that the Drosophila ADE8 segment on a replicating plasmid in ade2 ade8 yeast gives rise to lighter red colonies than does the natural yeast ADE8 segment similarly cloned and maintained (data not shown). This is the behavior expected for the somewhat reduced ADE8⁺ function of the Drosophila gene that was previously detected by prototrophy (1).

In order to quantify this visual assessment of ADE8⁺ function, we grew liquid cultures in tubes, centrifuged them, and compared the shades of red of the translucent pellets to standards made by mixing fully red ade2 control cells with fully white ade2 ade8 cells similarly treated. We found that ten successive 1:1 dilutions of red with white cells could be distinguished visually, covering about three orders of magnitude. A score could be assigned to each ADE8 transformant based on the color of its pellet when compared to standards.

Figure 5 (left panel) shows that with glucose as the carbon source, pigment accumulates most rapidly for pYeADE8(4.9) as expected from earlier comparisons measuring growth under adenine selection (1). In cells carrying pADHAD8, pigment accumulates at first more rapidly than in cells carrying pYFADE8(1.6R). Later, pigment accumulation appears to level off where the ADE8 gene is controlled by the ADHI promoter. Eventually, the rate of pigment accumulation decreases in these cells when compared to cells in which the gene is controlled by the Drosophila "promoter". This decrease for cells in which ADE8 is controlled by the ADHI promoter is similar to the behavior of the yeast ADHI gene as cells ferment glucose to ethanol (11).

With ethanol as the carbon source (Figure 5, right panel), pigment accumulation is again most rapid for pYADE8(4.9). However, a comparison between pYFADE8(1.6R) and pADHAD8 reveals that pigment accumulates much more slowly in the latter case, where the same gene is controlled by the ADHI promoter. The difference in pigment accumulation can be roughly estimated to

Figure 5 Red color accumulation of cells carrying ADE8 plasmids. Tube cultures were scored by comparison to standards over several days during which cultures were in stationary growth. A color doubling is an estimate of intensity based on standards that differ in amount of pigment by factors of 2. The total scale represents approximately a 3 order of magnitude difference in pigment.

be about an order of magnitude. Assuming that the *Drosophila* "promoter" sequence is not regulated by carbon source, this estimate is in accord with the reduction in ADHI activity and mRNA accumulation that occurs in later stages of fermentation (11). The fact that a gene placed under control of the ADHI promoter responds to a change in carbon source as does the ADHI gene itself indicates that carbon source regulation occurs at the level of transcription within the promoter fragment. Furthermore, this result suggests that carbon source regulation is not impaired by placing the promoter on a multiple-copy plasmid.

The above results indicate that changes in transcription of the *Drosophila* adenine-8 gene can be detected using a simple red color assay. We expect that this assay will prove to be useful in other studies such as the testing of sequences for enhancement or suppression of transcription and the selection

and characterization of mutations in regulatory elements.

Our color screening system resembles that of Guarente and Ptashne (15) who were able to screen for function in yeast using fusions to the *E. coli lacZ* gene. One difference is that in addition to color screening, selection both for the *ADE8*⁺ phenotype (prototrophy) or against it (white papilla overgrowth) allows for detection of very infrequent mutations in either direction. Also, we have noted that *ade2 ade8* cells carrying the *Drosophila ADE8* sequence on a 2 micron plasmid vector show considerable red and pink fine structure variegation of colonies on plates (preliminary results). This variegation probably reflects variable copy number, making the system potentially useful for the rapid screening of mutations in elements that control plasmid segregation, such as centromeres.

ACKNOWLEDGMENTS

We thank Gustav Ammerer for providing the *Bam*HI-linked ADHI promoter fragment and Bruce Veit for help in plasmid construction. This work was supported in part by United States Public Service Grants GM29009 and GM15253.

REFERENCES

1. Henikoff, S., Tatchell, K., Hall, B.D., and Nasmyth, K.A. (1981) *Nature* (London) 289,33-37.
2. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
3. Maxam, A.M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* 74, 560-564.
4. Beggs, J.D. (1978) *Nature* (London) 275, 104-108.
5. Sherman, F., Fink, G.R. and Hicks, J.B. (1981) *Methods in Yeast Genetics*, Cold Spring Harbor, New York, pp.61-62.
6. Proudfoot, N.J. (1979) *Nature* (London) 279, 376.
7. Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) *Nucleic Acids Res.* 8, 127-142.
8. Woods, R.A. and Jackson, I.E. (1973) *Biochem. Biophys. Res. Comm.* 53, 787-793.
9. Hitzeman, R.A., Hagie, F.E., Levine, H.L., Goeddel, D.V., Ammerer, G. and Hall, B.D. (1981) *Nature* (London) 293, 717-722.
10. Bennetzen, J. and Hall, B.D. (1982) *J. Biol. Chem.* 257, 3018-3025.
11. Denis, C., Ferguson, J. and Young, E.T. (1982) *J. Biol. Chem.* In press.
12. McAlister, L. and Holland, M.S. (1982) *J. Biol. Chem.* 257, 7181-7188.

13. Roman, H. (1956) Compt. Rend. Lab. Carlsberg Ser. Physiol 29, 279-314.
14. Parry, J.M., Parry, E.M. and Barrett, J.C. (1981) Nature 294, 263-265.
15. Guarente, L. and Ptashne, M. (1981) Proc. Natl. Acad. Sci. USA 78, 2199-2203.