Volume 14 Number 20 1986 Nucleic Acids Research

Nucleotide sequence of the *Saccharomyces cerevisiae* MET25 gene

Pierre Kerjan, Hélène Cherest and Yolande Surdin-Kerjan

Laboratoire d'Enzymologie du CNRS, 91190 Gif-sur-Yvette, France

Received 15 July 1986; Revised and Accepted 25 September 1986

ABSTRACT

To elucidate further the molecular basis of the specific regulatory mechanism modulating the expression of the genes implicated in methionine metabolism, we have cloned and characterized two genes, MET3 and MET25, and shown that the regulation of their expression is transcriptional (1, 2).

The sequence of the cloned yeast MET25 gene which encodes the O-acetyl homoserine - O-acetyl serine (OAH-OAS) sulfhydrylase is reported here along with its 5' and 3' flanking regions. The amino acid composition predicted from the DNA sequence is in good agreement with that determined by hydrolysis of the purified enzyme (3). In the 5' flanking region the signal for general amino acid control was not found, corroborating our previous finding that the synthesis of OAH-OAS sulfhydrylase is not submitted to general control. The transcription start points have been determined. The 5' and 3' flanking regions of the MET25 gene suggest initiation and termination signals similar to those associated with other yeast genes.

INTRODUCTION

The MET25 gene of Saccharomyces cerevisiae encodes the enzyme O-acetyl homoserine - O-acetyl serine (OAH-OAS) sulfhydrylase (4, 5). We had shown that the synthesis of this enzyme is submitted to the same regulatory system as other enzymes implicated in methionine metabolism (6, 7). As a first step towards elucidating the molecular basis for repression of enzyme synthesis in this pathway, we have already cloned two genes : MET3 encoding ATP sulfurylase (1) and MET25 (2). We have shown that the regulation of MET3 and MET25 expression was transcriptional and, in the case of MET25 we have identified a 120 bp region necessary for regulation to take place. Here we report the complete nucleotide sequence of the MET25 gene with its 5' and 3' flanking regions.

MATERIALS AND METHODS

Strain and culture conditions

The yeast strain used in this work was the wild type strain FL100

(MATa) obtained from Dr. F. Lacroute. Cultures used for the preparation of RNAs were grown to $2x10^7$ cells per ml in YNB medium (0.7% Yeast Nitrogen base without amino acids, 2% glucose).

DNA sequencing

DNA sequence analysis was performed as described by Maxam and Gilbert (8). End labelling of 5' termini was performed with (γ^{32}P-ATP) using polynucleotide kinase as described in Maniatis et al. (9).

Preparation of RNAs from S. cerevisiae and S1 mapping of the transcripts

The RNAs were extracted from cultures of strain FL100, and poly(A)$^+$RNAs were isolated on oligo-dT cellulose columns as described previously (1).

Mapping experiments with nuclease S1 were performed according to standard protocols (9, 10). The 200 pb BamHI-HpaII fragment from pM25-14 (2 and fig. 3) which includes the 5' end of the MET25 gene was 5' end-labelled using polynucleotide kinase and the two DNA strands were separated on a 5% polyacrylamide sequencing gel. The coding strand was hybridized 4 hours at 44°C to 10 µg of poly(A)$^+$ RNA in 30 µl of hybridization buffer (40 mM PIPES buffer pH 6.4, 1 mM EDTA, 40 mM NaCl, 80% formamide). The hybrid was then diluted 10 fold in S1 buffer (50 mM sodium acetate pH 4.6, 140 mM NaCl, 2.25 mM $ZnSO_4$ and 10 µg/ml of sonicated and denatured thymus DNA) containing 150, 500 and 1500 units of nuclease S1. The S1 digestion was carried out for 30 minutes at 37°C. After recovery by ethanol precipitation, the RNA-DNA hybrids were analysed on a 5% polyacrylamide sequencing gel along with the chemical degradation products of the ^{32}P end-labelled DNA probe used.

Computer analysis

Analysis of the MET25 sequence data was performed on a GOUPIL 3 microcomputer with programs designed for us by Daniel Kerjan, Computer Department of the Lycée Agricole of Le Rheu (France).

RESULTS

Nucleotide sequence analysis of the MET25 gene

We had cloned and studied previously the MET25 gene encoding O-acetyl serine - O-acetyl homoserine (OAS-OAH) sulfhydrylase of S. cerevisiae (2). We had shown that this gene was beared on a 2.4 kb DNA fragment. A restriction map of this region together with the sequencing strategy is shown in Figure 1. We used 5' end labelling and secondary digestion to generate DNA fragments for analysis by the Maxam and Gilbert method. We determined 100% of the sequence on the two strands. The complete sequence of the

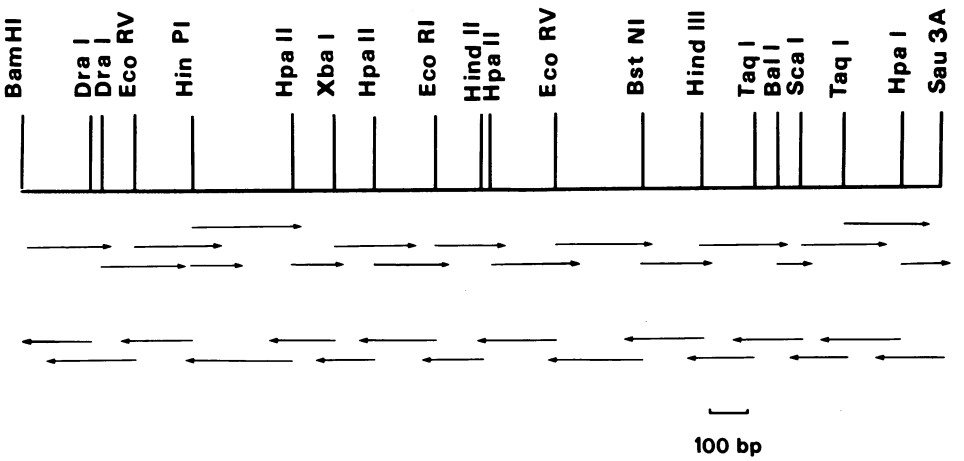
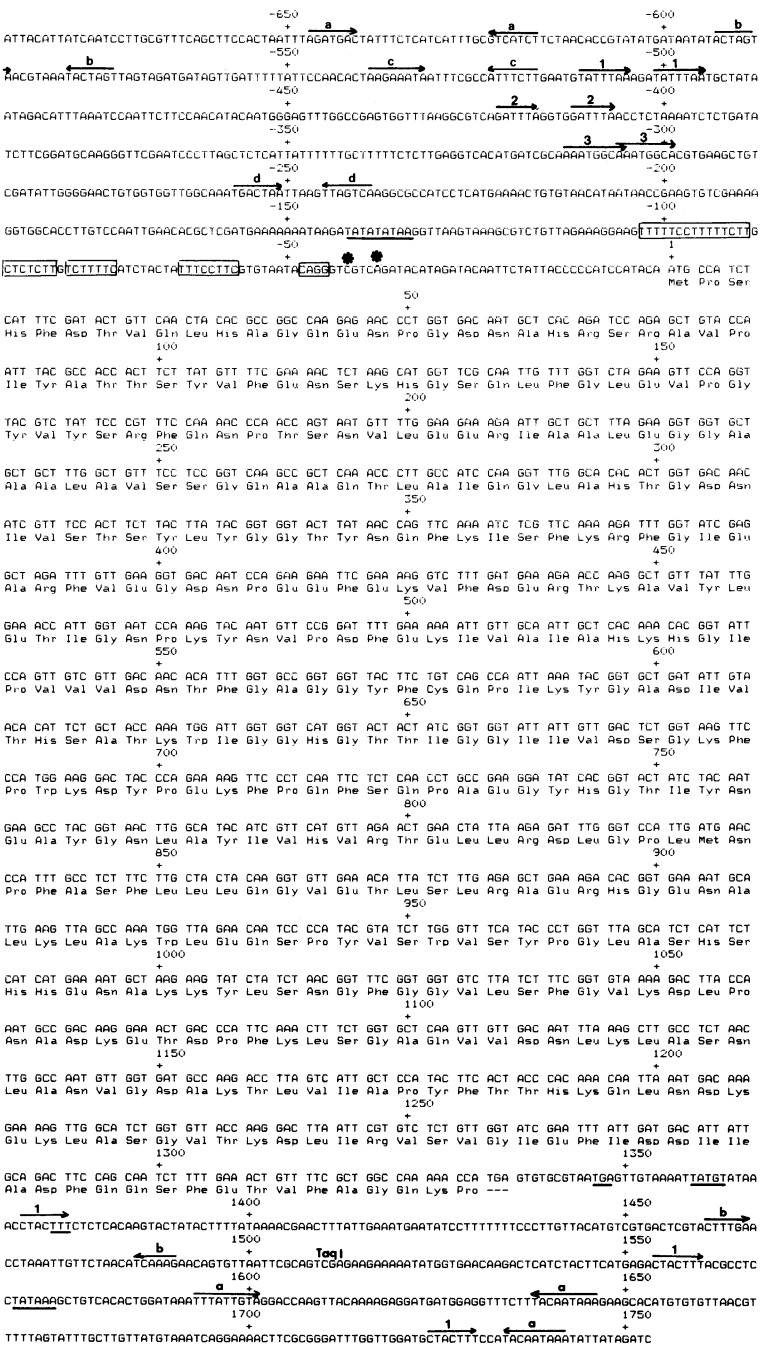


Figure 1: Sequencing strategy for the MET25 gene. The length of each arrow indicates the distance read from the 5' labelled end.

MET25 region is shown in figure 2. Analysis of the nucleotide sequence revealed only one open reading frame (ORF) in the region previously shown to be required to complement met25 mutations. This open reading frame was found between positions +1 and +1335 beginning with a methionine codon and ending with a TGA codon. Short ORFs beginning with ATG codons located before position +1 can be found but none would yield a polypeptide chain longer than 23 amino acid residues. All other ORFs beginning with ATG codons located after position +1 would not yield a polypeptide chain compatible with the published molecular weight value for the subunit of OAH OAS sulfhydrylase (3). The polypeptide derived from the DNA sequence would contain 444 amino acids residues with a molecular weight of 48,678. These values compare reasonably well with the results obtained by Yamagata (3). This author, who found a molecular weight of 54,000 $\pm$ 3000 daltons for the subunit of the yeast OAH-OAS sulfhydrylase, depending on the method used, also determined the amino acid composition of the subunit. It is reported in table 1 for comparison with that of the polypeptide derived from the MET25 gene sequence. It can be seen that a good homology is found. An interesting feature, already pointed out by Yamagata is the presence of only one cysteine residue per subunit of OAH-OAS sulfhydrylase.

As stated above, a small open reading frame of 23 amino acid residues was found in the 5' region of the MET25 gene. It begins at position -322 and ends at position -251. It is not known whether this region is transcribed.



Also this reading frame is partly in the area previously shown to be required for regulation of transcription of the MET25 gene (2).

Codon usage in MET25

Translation of the 1335 bp open reading frame corresponding to the MET25 gene is shown in figure 2. Codon usage and their frequency of appearance are shown in table 2. Bennetzen and Hall (11) have defined a codon bias index which quantifies the extent of bias towards a set of 22 preferred codons identified in two very highly expressed genes in yeast (encoding Alcohol dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase). Using these authors formulas we found a bias index of 0.54 for MET25 which indicates a moderately expressed gene. This is in good agreement with the experimentally determined percentage of MET25 specific mRNA found to be 0.05% of total mRNAs (2).

Mapping of the MET25 transcription start points

The S1 nuclease mapping technique was used to align the 5' end of MET25 mRNA with the cloned DNA. We used the labelled fragment shown in Figure 3, and the Berk and Sharp S1 mapping method in conditions described in Material and Methods. Results (figure 3) show that two major RNA protected fragments were obtained beginning at positions -42 and -38 which are located by asterisks in figure 2. In addition, four persistent minor bands were found at positions -39, -43, -46 and -47. Our results thus suggest a low heterogeneity in the 5' termini of MET25 transcripts with three very close regions of transcriptions start, one located at positions -47 and -46, one at positions -39 and -38 and one at positions -43 and -42. These positions define potential *in vivo* initiation sites and have been corrected for the fact that Maxam-Gilbert DNA sequence reaction products migrate 1.5 bp faster than the corresponding S1 fragment (12). Zalkin and Yanovsky (13) have pointed out the presence of a sequence complementary to the 3' terminus of yeast 18S RNA near the transcriptional initiation site of the yeast TRP5 gene. This sequence (5'-TAATGATC-3') was found with few

Figure 2: Nucleotide and deduced amino acid sequence of the MET25 gene and its 5' and 3' flanking regions. The non coding strand is shown. The nucleotide sequence is numbered from the first nucleotide of the presumed initiation codon. At the 5' end, the TATA box is underlined, asterisks indicate transcription initiation starts and arrows indicate direct (numbered) or inverted (lettered) repeats. The CT rich regions possibly implicated in promoter activity are boxed. At the 3' end, the sequences possibly involved in transcription termination are underlined. The arrows indicate sequence repeats as in the 5' region.

Table 1
Comparison of the amino acid composition of yeast OAH-OAS sulfhydrylase predicted from DNA sequence with that obtained by amino acid analysis.

Amino Acid residue	Number of residues per subunit of OAH-OAS sulfhydrylase	
	DNA sequence	Composition of enzyme (Yamagata, 3)
Ala	39	41
Arg	12	9
Asp + Asn	43	47
Cys	1	1
Glu + Gln	45	48
Gly	42	45
His	16	15
Ile	24	22
Leu	35	37
Lys	27	27
Met	2	5
Phe	26	27
Pro	22	28
Ser	29	33
Thr	24	30
Trp	4	4
Tyr	20	14
Val	33	35
Total number	444	468

modifications near the transcriptional initiation site of several other yeast genes. Zalkin and Yanovski suggested that this site could function as a binding site for ribosomes. A sequence 5'-TAATACAGGGTCG*-3' (beginning at position -53) is found in the 5' flanking region of the MET25 gene. Nevertheless, it seems to be too distant from the main transcription starting point (marked here by an asterix) and moreover, it would not be transcribed into mRNA.

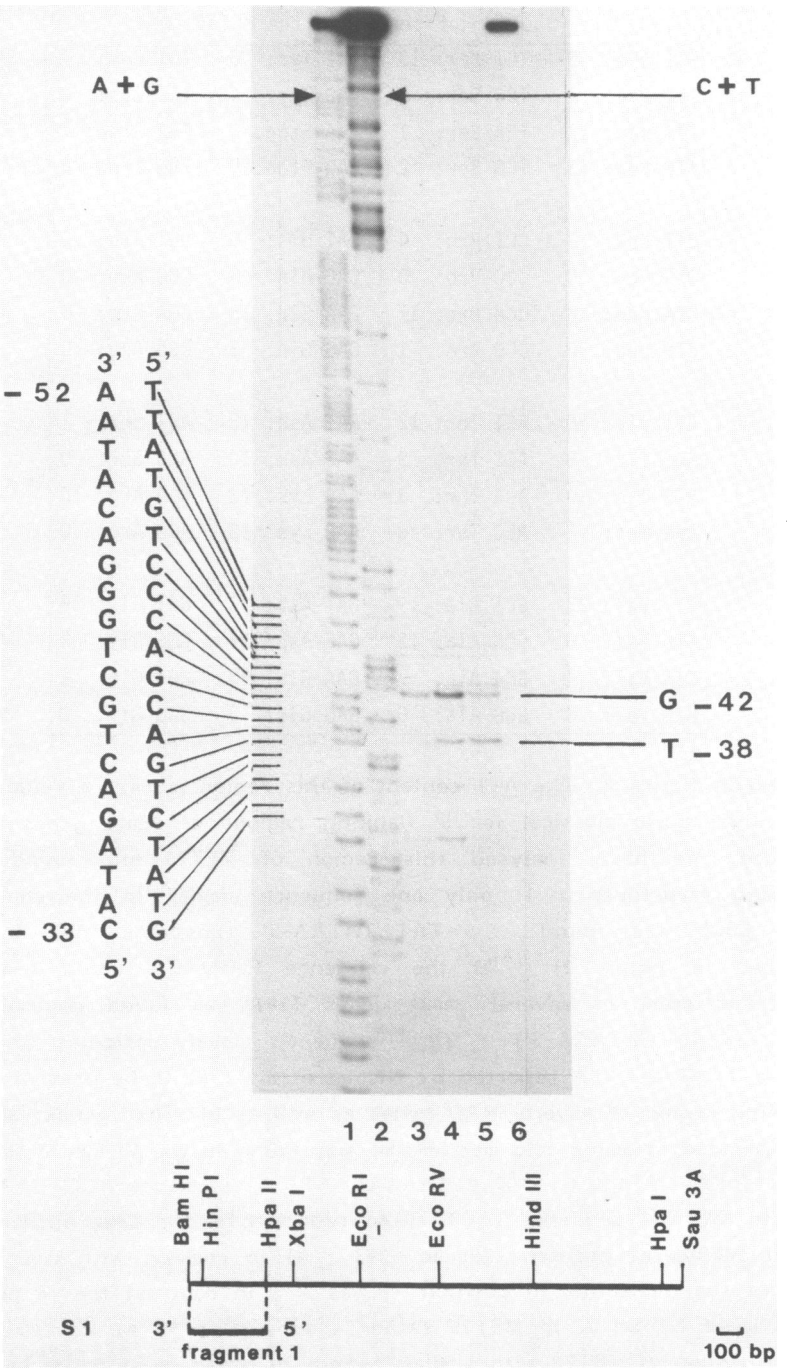
Analysis of the sequences flanking the 5' end of MET25.

The nucleotide sequence of 687 pb 5' to the MET25 gene was determined

Table 2
Distribution of codons in the MET25 gene

TTT Phe: 9	TCT Ser: 19	TAT Tyr: 6	TGT Cys: 1
TTC Phe: 17	TCC Ser: 6	TAC Tyr: 14	TGC Cys: 0
TTA Leu: 13	TCA Ser: 1	TAA Stop: 0	TGA Stop: 1
TTG Leu: 13	TCG Ser: 2	TAG Stop: 0	TGG Trp: 4
CTT Leu: 3	CCT Pro: 4	CAT His: 8	CGT Arg: 2
CTC Leu: 0	CCC Pro: 0	CAC His: 8	CGC Arg: 0
CTA Leu: 6	CCA Pro: 17	CAA Gln: 15	CGA Arg: 0
CTG Leu: 0	CCG Pro: 1	CAG Gln: 3	CGG Arg: 0
ATT Ile: 16	ACT Thr: 12	AAT Asn: 12	AGT Ser: 1
ATC Ile: 8	ACC Thr: 9	AAC Asn: 10	AGC Ser: 0
ATA Ile: 0	ACA Thr: 3	AAA Lys: 12	AGA Arg: 10
ATG Met: 2	ACG Thr: 0	AAG Lys: 15	AGG Arg: 0
GTT Val: 23	GCT Ala: 19	GAT Asp: 7	GGT Gly: 39
GTC Val: 6	GCC Ala: 13	GAC Asp: 14	GGC Gly: 2
GTA Val: 4	GCA Ala: 7	GAA Glu: 25	GGA Gly: 1
GTG Val: 0	GCG Ala: 0	GAG Glu: 2	GGG Gly: 0

as shown in figure 2. The A+T content of this region is 64%, a value which seems to be quite classical for 5' flanking regions of yeast genes so far sequenced. We have analysed this region of MET25 and found many interesting structures : 1) only one sequence similar to the consensus "TATA box" was found : 5'-TATATATAA-3' (positions -142 to -134 underlined in figure 2) ; 2) the sequence CAPyACA found near the translational start of several yeast genes (14) was found beginning at position -6 (5'-CATACA-3') ; 3) a sequence closely matching the CS1 sequence TGAAAAAAAA reported by Nicolet et al. (15) to be present in the 5' flanking region of several RAD genes as well as of other genes, some of them implicated in amino acid metabolism, was found in the MET25 5' flanking region at position -156 to -147 (TGAAAAAAAT). A sequence ACGTGAA very similar to the CS2 sequence (ACTTGAA) reported by the same authors was found in MET25 at position -299 to -293 ; 4) in common with most yeast sequences the nucleotide in position -3 was A ; in contrast the +6 position was occupied by an A in MET25 rather than a T ; 5) we compared the promoter region of MET25 with similar regions of other yeast genes compiled


A + G
C + T
3' 5'
- 52
A T
A T
T A
A T
C G
A T
G C
G C
G C
T A
C G
G C
T A
C G
A T
G C
A T
T A
A T
- 33
C G
5' 3'
G - 42
T - 38
1 2 3 4 5 6
Bam HI
Hin PI
Hpa II
Xba I
Eco RI
Eco RV
Hind III
Hpa I
Sau 3A
S 1
3'
5'
fragment 1
100 bp

by Oberto and Davison (16). We found the CT block followed 8 bp later by a sequence CAGG instead of CAAG situated about 30 bp before the beginning of translation. These structures are boxed in figure 2 ; 6) The 5' flanking region of MET25 contains many direct and inverted repeats. Some of them are indicated by numbered or lettered arrows in figure 2.

Analysis of the sequence flanking the 3' end of MET25

We have analysed the nucleotidic sequence of the 415 bp region downstream of the TGA translational termination codon. This region shows an A+T content of 66% with several tracts of four or more As or Ts scattered throughout. Zaret and Sherman (17) have suggested that the sequence :

TAA

TAG ...1-140...(T rich)...TAG...TA(T)GT...(AT rich)...TTT

TGA

in the 3' non coding region of S. cerevisiae genes may play a role in efficient transcription termination. The 3' region of MET25 shows the related sequence (underlined in figure 2) :

1333 TGA ..(10 bp)..TGA..(10pb)..TATGT..(10bp)..TTT 1376.

We have mapped previously (2) the end of transcription near the TaqI site in position 1509, which would be about 100 pb downstream of the underlined signals. The sequence AATAAA has been implicated as a polyadenylation signal in higher eukaryotes (18). Although such sequences have been found in many yeast genes their role in yeast polyadenylation has not been established. Two AATAAA sequences are found in the 3' region of MET25, at position 1641 and at position 1737 but these sites are situated well after the transcriptional stop. The related sequence TATAAA is found 3 times at positions 1363, 1398 and 1569. The two first ones being situated before the transcriptional termination region, could be polyadenylation signals. In the 3' flanking region of MET25 we have found also many direct and inverted repeats. Some of them are shown by arrows in figure 2.

Figure 3: 5' termini of the MET25 transcript. S1 mapping of MET25 transcription start was performed as indicated in Material and Methods, using the coding strand of the BamHI-HpaII DNA fragment, shown on the lower panel, as a probe. Lanes 1 and 2 show the sequence of the coding strand of this fragment. Lanes 3, 4 and 5 show the mRNA protected fragment submitted to 500, 2000, 5000 u/ml of nuclease S1 respectively. Lane 6 shows the control without poly(A^+) RNA added. The DNA sequence is shown vertically on the left. All positions are numbered as in figure 2. The positions of the protected fragments indicated on the right have been corrected for the 1.5 bp discrepancy in the migration rate with the sequence sample. All indicated positions are on the coding strand used in this experiment.

General amino acid control

Unpublished results from this laboratory had shown that the synthesis of OAH-OAS sulfhydrylase (encoded by the MET25 gene) was not submitted to the general amino acid control. In the 5' regions of the genes the expression of which is under general amino acid control (19, 20, 21), repeats of the sequence 5'-TGACTC-3' have been found in every case (22). This hexanucleotide is not present in the 600 bp preceding the MET25 coding region. This is further supporting that the expression of gene met25 is not submitted to the general amino acid control.

DISCUSSION

We have determined the nucleotide sequence of the OAH-OAS sulfhydrylase coding gene of S. cerevisiae together with its 5' and 3' flanking regions. The experiments reported here localize the 5' end of MET25 transcript to three close regions of 2 bp beginning 49 to 38 bp upstream of ATG. This variation in the pattern of S1 protection fragments must be interpreted circumspectly as specifying the transcription start point to a particular position is beyond the capabilities of the S1 mapping method. Our data show one TATA box situated about 100 pb upstream of the MET25 transcription initiation. Only one open reading frame was found. The amino acid composition derived from the DNA sequence is very similar to that found by protein analysis (3). All the yeast genes so far studied and which are under the general amino acid control exhibit the short sequence TGACTC in their 5' upstream region (22). Such a signal is absent in the 5' flanking region of MET25 which is consistent with our previous finding that the synthesis of OAH-OAS sulfhydrylase is not derepressed in conditions where the general control is acting.

Specific repression of OAH-OAS sulfhydrylase synthesis is promoted by the presence of exogenous methionine (6, 7), and we have already shown that this regulation needs the presence of specific sequences in the 5' upstream region of MET25 (2). We expect these "regulatory sequences" to be identical in different genes submitted to the same regulation, as the MET3 gene. Consequently, the sequencing of this gene is now under progress.

ACKNOWLEDGEMENT

We are indebted to Mrs Denise Henry for skillful technical assistance.

This work was partly supported by the A.T.P. "Organisation and expression du génome dans les cellules eucaryotes".

REFERENCES

1. Cherest, H., Nguyen Ngoc, T. and Surdin-Kerjan, Y. (1985) Gene 34, 269-281.
2. Sangsoda, S., Cherest, H. and Surdin-Kerjan (1985) Mol. Gen. Genet., 200, 407-414.
3. Yamagata, S. (1976) J. Biochem. (Tokyo), 80, 787-797.
4. Cherest, H., Eichler, F. and de Robichon-Szulmajster, H. (1969) J. Bacteriol., 97, 328-336.
5. Yamagata, S., Takeshima, K. and Naiki, N. (1974) J. Biochem. (Tokyo), 75, 1221-1229.
6. Cherest, H., Surdin-Kerjan, Y., Antoniewski, J. and de Robichon-Szulmajster, H. (1973) J. Bacteriol., 97, 328-336.
7. Cherest, H., Surdin-Kerjan, Y., Antoniewski, J. and de Robichon-Szulmajster (1973) J. Bacteriol., 115, 1084-1093.
8. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
9. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: a laboratory manual., Cold Spring Harbor, p. 122.
10. Berk, A. and Sharp, P. (1977) Cell, 12, 721-732.
11. Bennetzen, J.L. and Hall, B.D. (1982) J. Biol. Chem., 257, 3026-3031.
12. Sollner-Webb, B. and Reeder, R.H. (1979) Cell, 18, 485-499.
13. Zalkin, H. and Yanofsky, C. (1982) J. Biol. Chem. 257, 1491-1500.
14. Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J. and Kingsman, S.M. (91982) Nucl. Ac. Res., 10, 2025-2037.
15. Nicolet, C.M., Chenevert, J.M. and Friedberg, E.C. (1985) Gene, 36, 225-234.
16. Oberto, J. and Davison, J. (1985) Gene, 40, 57-65.
17. Zaret, K.S. and Sherman, F. (1982) Yeast Cell, 28, 563-573.
18. Fitzgerald, M. and Schenk, T. (1981) Cell, 24, 251-260.
19. Schurch, A., Miozzari, G. and Hutter, R. (1974) J. Bacteriol., 117, 1131-1140.
20. Delforge, J., Messenguy, F. and Wiame, J. (1975) Eur. J. Biochem., 57, 231-239.
21. Wolfner, M., Yep, D., Messenguy, F. and Fink, G.R. (1975) J. Mol. Biol., 96, 273-290.
22. Jones, E.W. and Fink, G.R. (1982) The molecular biology of the yeast Saccharomyces - Metabolism and gene expression, 181-299.