Volume 13 Number 17 1985 Nucleic Acids Research

Cloning and characterization of the gene for the yeast cytoplasmic threonyl-tRNA synthetase

Louise K.Pape and Alexander Tzagoloff

Department of Biological Sciences, Columbia University, New York, NY 10027, USA

Received 13 May 1985; Revised and Accepted 24 July 1985

ABSTRACT

A fragment of DNA from the yeast nuclear gene *MST1* that codes for the mitochondrial $tRNA_1^{Thr}$ synthetase was used as a probe to screen for other yeast threonyl-tRNA synthetase genes. At low stringency, the *MST1* probe hybridizes strongly to a 6.6 kb *Eco*RI fragment of yeast genomic DNA with the homologous gene and in addition hybridizes more weakly to a smaller 3.6 kb *Eco*RI fragment with a second threonyl-tRNA synthetase gene (*THS1*). To clone *THS1*, a library was constructed by ligation to pUC18 of size selected (3-4.5 kb) *Eco*RI fragments of genomic DNA. Several clones containing the 3.6 kb *Eco*RI fragment were isolated. A 2,202 nucleotide long open reading frame corresponding to *THS1* has been identified in the cloned fragment of DNA. The predicted protein encoded by *THS1* is 38% identical to the *E. coli* threonyl-tRNA synthetase over the latter's length (642 amino acids) and is 42% identical to the predicted *MST1* product over its 462 residues. *In situ* disruption of the chromosomal copy of *THS1* is lethal to the cell, indicating that this gene codes for the cytoplasmic threonyl-tRNA synthetase.

INTRODUCTION

At present there are only a few reported instances of yeast mutants with defective cytoplasmic tRNA synthetases (1-5). As a result only a small number of the genes have been cloned by transformation, and there is relatively little known about the extent to which the eucaryotic synthetases have diverged from the better studied procaryotic enzymes. In the present study we demonstrate that it may be feasible to isolate the genes for the yeast cytoplasmic aminoacyl-tRNA synthetases by using yeast mitochondrial synthetase probes.

We have recently reported the cloning of the yeast nuclear gene *MST1* (6). This gene was shown to code for a mitochondrial aminoacyl-tRNA synthetase that charges the $tRNA_1^{Thr}$ with the anomalous 3'-GAU-5' anticodon (6). Mutations in this synthetase gene were found to abolish charging of $tRNA_1^{Thr}$ but had no effect on charging of $tRNA_2^{Thr}$ whose anticodon recognizes the conventional ACN family of codons (6). Due to a high degree of homology between the MST1 gene and the gene coding for the yeast cytoplasmic threonyl-tRNA synthetase (here referred to as *THS1*), we were able to use an *MST1* probe to select bacterial clones harboring recombinant plasmids with the *THS1* gene. Further characterization of this gene has revealed that of the two known yeast threonyl-tRNA synthetases, the cytoplasmic species is more closely related to the procaryotic enzyme, both in size and primary sequence.

MATERIALS AND METHODS

Strains and media: The strains of *Saccharomyces cerevisiae* and the composition of media have been described previously (6).

Preparative and analytical procedures: Procedures for the preparation of plasmid DNA, yeast genomic DNA and their subsequent analyses are identical to those described previously (6,7).

Construction and screening of the yeast genomic library: Chromosomal DNA purified from *S. cerevisiae* D273-10B was digested to completion with *Eco*RI. The digestion products were separated on a preparative 1% agarose gel (8), and fragments ranging from 3 to 4.5 kb were recovered by electroelution. The DNA was extracted twice with water-saturated phenol, four times with ether, and was precipitated with alcohol from 2 M ammonium acetate. Following ligation to the *Eco*RI site of pUC18 (9), the mixture was used to transform *E. coli* RR1. Transformants were selected on 1.75% Antibiotic Medium #3 (Difco) containing 40 μg/ml ampicillin. Approximately 3,000 ampicillin resistant colonies were screened with a nick-translated fragment of DNA (*Bgl*II-*Hind*III) from the *MST1* gene coding for residues 45-163 of mitochondrial $tRNA_1^{Thr}$ synthetase (6). The transformant colonies were transferred to two separate nitrocellulose circles prewetted on an ampicillin plate (10). These were incubated face up at 37° C on fresh ampicillin medium. The colonies on the nitrocellulose filters were lysed and washed as described by Grunstein and Hogness (10). Hybridizations were done at 37° C for 36 hours in 43% formamide, 6x SSC, 0.1% sodium dodecyl sulfate (SDS), 5X Denhardt's, 50 mM $NaPO_4$ pH 6.5, 100 μg/ml salmon sperm DNA (11) and in duplicate at 30% formamide. Approximately 10^7 cpm was added to 6 ml of the hybridization solution. The filters were washed two times at room temperature with 2x SSC, 0.1% SDS, and an additional two times at 50° C for 15 min in 2x SSC, 0.1% SDS (11). Colonies that gave positive signals under both hybridization conditions were purified and screened a second time as above except that Whatman 541 paper (12) was used instead of nitrocellulose.

Southern hybridizations: Yeast genomic DNA was separated electrophoreticallly on 1% agarose after digestion with restriction endonucleases and processed by the method of Southern (13). High stringency hybridizations were done either at 65° in 6x SSC, 0.01 M EDTA, 0.5% SDS, 5x Denhardt's, 100 μg/ml salmon sperm DNA or at 45° in the same buffer with 50% formamide. For low stringency hybridizations, the temperature was reduced to 37° with either 43% or 30% formamide added to the standard buffer.

DNA sequence analysis: The sequence of the *THS1* gene was derived from the insert of pUC18/THS1. The 1.5 kb *Eco*RI-*Hind*III, 1.2 kb *Bgl*II-*Sal*I, 750 bp *Bam*HI-*Hind*III, and 650 bp *Sal*I-*Bam*HI fragments were isolated from preparative agarose gels. These were subjected to secondary cleavages with other restriction endonucleases. The digestion products were 5'-end labeled with γ-[^{32}-P]-ATP (5,000 Ci/mmol, ICN) in the presence of T4 polynucleotide kinase. Following denaturation, single strands were separated on 4% polyacrylamide and sequenced by the method of Maxam and Gilbert (14). Most of the sequence was confirmed on both strands.

Extraction of cytoplasmic tRNAs and aminoacyl-tRNA synthetases: Cytoplasmic tRNAs were isolated from D273-10B ρ^0 as described (15). *S. cerevisiae* αW303-1B and an αW303-1B transformant containing YEp351/THS1 were grown to late logarithmic phase in minimal glucose media. The cells were harvested by centrifugation, washed in MTE (0.25 M mannitol, 20 mM Tris-Cl pH 7.5, 1 mM EDTA) and resuspended in the same buffer at a concentration of 8 g. wet weight cells/30 ml. After cell disrup-

tion in a Braun glass bead homogenizer, the cell debris and mitochondria were pelleted at 15,000 rpm for 30 min at 4° C. The supernatant was applied to a DEAE column (Whatman DE52) equilibrated in 0.4 M NaCl, 20 mM Tris-Cl pH 7.5, after which the column was washed in 1 column volume of the same buffer. The flow-through and wash were collected, and dialysed 2 hr at 4° C vs. 10 mM Tris-Cl pH 7.5, 0.1 mM EDTA, 1 mM DTT (Dithiothreitol) and used as the source of yeast aminoacyl-tRNA synthetases. Protein concentration was determined by the Lowry method (16). In vitro aminoacylation reactions were perfomed as described by Macino and Tzagoloff (15), except the reaction mix was 50 µl and contained 0.2 mM DTT and 50 µg of cytoplasmic tRNAs.

RESULTS AND DISCUSSION

Hybridization analysis of yeast genomic DNA with an *MST1* probe under low stringency conditions

One of the yeast mitochondrial threonyl-tRNA synthetase genes (*MST1*) has been cloned and shown to code for an enzyme capable of acylating $tRNA_1^{Thr}$ but not $tRNA_2^{Thr}$ (6). In order to determine whether an *MST1* probe could be used to screen for one or both of the other threonyl-tRNA synthetase genes, chromosomal DNA of wild-type yeast was examined by Southern hybridization analysis with a probe containing the entire coding region of *MST1*. At high stringency the probe detects only a 6.6 kb *Eco*RI fragment of genomic DNA which was previously found to encompass the entire *MST1* gene. As

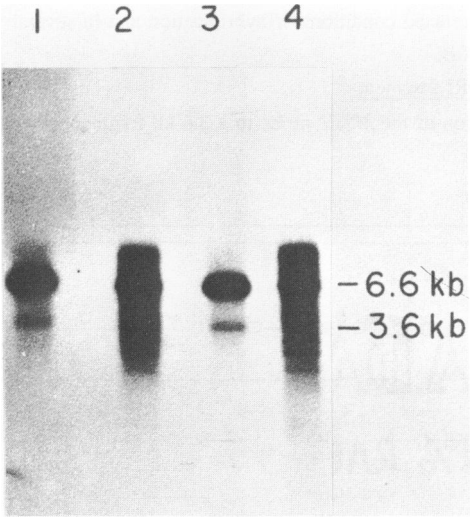


Fig. 1. Low stringency hybridization of wild-type yeast genomic DNA with an *MST1* probe. DNA purified from the wild-type strain 41-8D was digested to completion with *Eco*RI and separated on a 1% agarose gel. The DNA was blotted to nitrocellulose and hybridized with pUC18/MST1, a clone of the 1.8 kb *Ava*I fragment containing the *MST1* gene in pUC18 (6). The hybridizations were done in 6x SSC, 0.01 M EDTA, 0.5% SDS, 5x Denhardt's, 100 µg/ml salmon sperm DNA at the following temperatures and concentrations of formamide. Lane 1: 30% formamide, 42°, lane 2: 30% formamide, 30°, lane 3: 50% formamide, 30°, lane 4: no formamide, 4x SSC, 42°. The positions and sizes of the *Eco*RI fragments containing the *MST1* gene and the homologous *THS1* gene are marked in the margin.

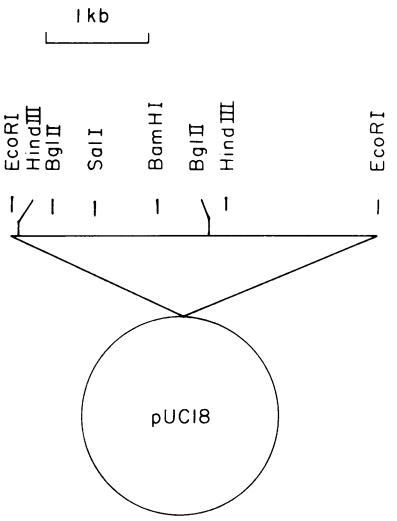


Fig. 2. Partial restiction map of pUC18/THS1. The positions of the *Hind*III, *Bgl*II, and *Sal*I sites in the 3.6 kb *Eco*RI fragment are shown in the upper part of the figure. The fragment was ligated to the *Eco*RI site of the multiple cloning sequence (solid arc) of the pUC18 vector.

shown in Fig. 1, a second *Eco*RI fragment of 3.6 kb is detected by the same probe when hybridizations are done at lower stringency. The signal of this smaller fragment is weaker but is reproducibly observed under the more relaxed conditions of hybridization and in several wild-type strains including 41-8D, W303 and D273-10B.

Cloning of the 3.6 kb *Eco*RI fragment

The cross-hybridization of the *MST1* probe to a 3.6 kb fragment of yeast nuclear DNA suggested

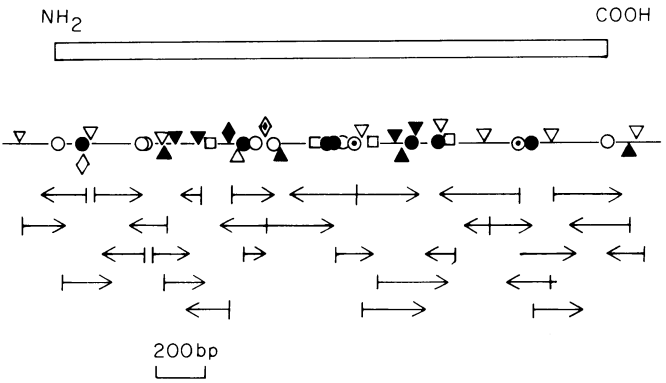


Fig. 3. Restriction sites used to sequence part of the 3.6 kb *Eco*RI insert in pUC18/THS1. The region sequenced starts 260 bp upstream of the unique *Bam*HI site in the insert and extends to the extreme right *Eco*RI site (see Fig. 2). The restriction sites used for 5'-end labeling were *Taq*I (△), *Hinf*I (▲), *Bam*HI (◇), *Dde*I (○), *Rsa*I (●), *Bst*N1 (□), *Hae*III (⊙), *Hind*III (◈), and *Bgl*II (◆). The arrows indicate the directions and approximate lengths of sequences obtained. The location of the open reading frame is shown by the open bar in the upper part of the figure.

that it might contain part or the whole gene for one of the other threonyl-tRNA synthetases. The gene will henceforth be referred to as *THS1* (threonyl-tRNA synthetase). The following strategy was adopted to clone the gene. Chromosomal DNA purified from the wild-type strain D273-10B was digested with *Eco*RI, and the products were separated on a preparative 1% agarose gel. DNA fragments in the range of 3 to 4.5 kb were eluted from the gel and ligated to the unique *Eco*RI site of pUC18 (9). Approximately 3,000 ampicillin resistant transformants harboring the products of the ligation were screened at low stringency with a *Bgl*II-*Hind*III fragment coding for residues 45-163 of $tRNA_1^{Thr}$ synthetase. This probe hybridizes to the 3.6 kb *Eco*RI fragment under low stringency (data not shown). Five clones each having plasmids with a 3.6 kb insert were obtained from the screen. A partial restriction map of the plasmid with the 3.6 kb *Eco*RI fragment (pUC18/THS1) is presented in Fig. 2.

Southern analyses of pUC18/THS1 confirmed the presence of a homologous gene in the 3.6 kb insert and further indicated that its 5' end was proximal to the *Sal1* site (Fig. 2).

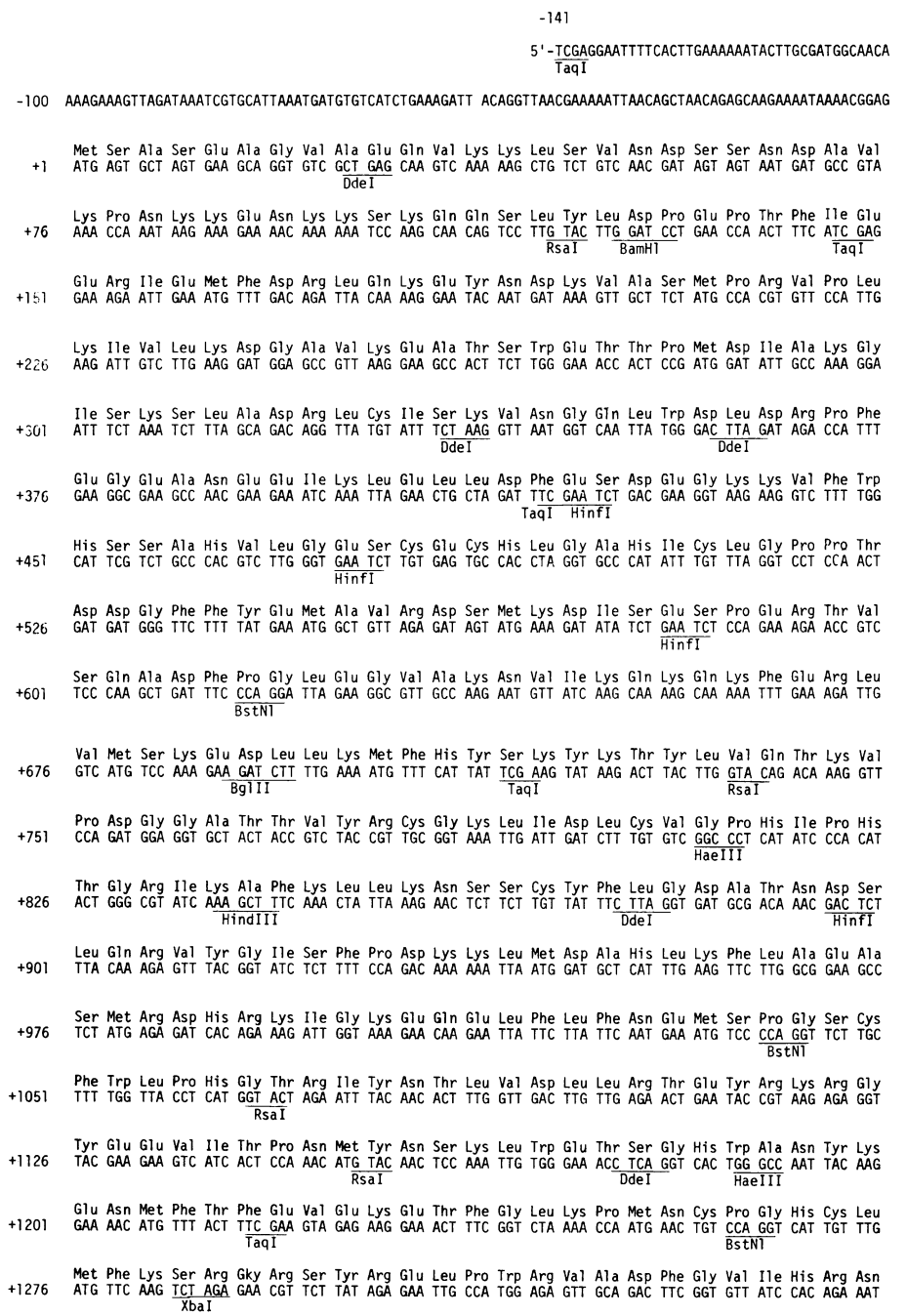
Sequence of part of the 3.6 kb insert in pUC18/THS1

The insert of pUC18/THS1 was sequenced by the strategy outlined in Fig. 3. The sequence was obtained from the 650 bp *Sal*I-*Bam*HI, 750 bp *Bam*HI-*Hind*III and 1,500 bp *Hind*III-*Eco*RI fragments. The *Bam*HI and *Hind*III sites were crossed from preparative *Sal*I-*Bgl*II and *Bgl*II-*Eco*RI fragments. Most of the sequence was confirmed from the complementary strands.

The nucleotide sequence revealed a long continuous reading frame starting with the ATG codon at +1 and ending with an ochre termination codon at +2203 (Fig. 4). The predicted protein encoded in this gene is 734 amino acid residues in length with a molecular weight of 84,380. It has a hydrophilic composition with a net charge of -5. There is a moderate bias in codon usage. Noticeably lacking are the CGC, CGA, and CGG codons for arginine (Table I).

In situ disruption of *THS1*

In order to determine whether *THS1* codes for a cytoplasmic or mitochondrial threonyl-tRNA synthetase, the chromosomal copy of the gene was disrupted by gene substitution (17). The *THS1* gene was isolated on a 3.4 kb *Eco*RI-*Bgl*II fragment and ligated at compatible sites in the multiple cloning region of pUC8 (18). This construct was used to delete a 1.16 kb segment of the coding region between the *BamH1* and *Xba*I sites (Fig. 5). The deleted region of the gene was substituted with a 1.7 kb fragment of yeast nuclear DNA with the *HIS3* gene (19). The isogenic haploid and diploid respiratory competent strains W303 α, a, and a/α with a *his3* marker were each transformed with the linear *Eco*RI-*Sal*I fragment containing the disrupted *THS1* gene. Transformants selected for loss of the histidine auxotrophy were collected and further checked for the presence of the disrupted copy of *THS1* in chromosomal DNA. Southern analysis of numerous haploid transformants indicated that none had acquired the disrupted *THS1* gene. Several diploid transformants complemented for the histidine auxotrophy, however, had one altered copy of *THS1* in addition to one wild-type copy. The results shown in Fig. 5 confirm the presence of the disrupted copy of *THS1* in one such diploid clone. Genomic DNAs from the parental W303 diploid and from the histidine independent transformant were digested with *Eco*RI alone and with a combination of *Eco*RI and *Pst*I. Following separation on agarose and transfer to nitrocellulose, the Southern blot was hybridized with a nick-translated 1.3 kb *Eco*RI-*Bam* fragment containing 5' flanking and 127 bp from the coding region of the gene. As seen in lanes 1 and 2, the *THS1* gene is present on a 3.6 kb *Eco*RI fragment in both the haploid and diploid W303 strains. Although this band is also present in the diploid transformant, there is a second hybridizing band


-141
5'-TCGAGGAATTTTCACTTGAAAAAATACTTGCGATGGCAACA
TaqI
-100 AAAGAAAGTTAGATAAATCGTGCATTAAATGATGTGTCATCTGAAAGATT ACAGGTTAACGAAAAATTAACAGCTAACAGAGCAAGAAAATAAAACGGAG
Met Ser Ala Ser Glu Ala Gly Val Ala Glu Gln Val Lys Lys Leu Ser Val Asn Asp Ser Ser Asn Asp Ala Val
+1 ATG AGT GCT AGT GAA GCA GGT GTC GCT GAG CAA GTC AAA AAG CTG TCT GTC AAC GAT AGT AGT AAT GAT GCC GTA
DdeI
Lys Pro Asn Lys Lys Glu Asn Lys Lys Ser Lys Gln Gln Ser Leu Tyr Leu Asp Pro Glu Pro Thr Phe Ile Glu
+76 AAA CCA AAT AAG AAA GAA AAC AAA AAA TCC AAG CAA CAG TCC TTG TAC TTG GAT CCT GAA CCA ACT TTC ATC GAG
RsaI BamHI TaqI
Glu Arg Ile Glu Met Phe Asp Arg Leu Gln Lys Glu Tyr Asn Asp Lys Val Ala Ser Met Pro Arg Val Pro Leu
+151 GAA AGA ATT GAA ATG TTT GAC AGA TTA CAA AAG GAA TAC AAT GAT AAA GTT GCT TCT ATG CCA CGT GTT CCA TTG
Lys Ile Val Leu Lys Asp Gly Ala Val Lys Glu Ala Thr Ser Trp Glu Thr Thr Pro Met Asp Ile Ala Lys Gly
+226 AAG ATT GTC TTG AAG GAT GGA GCC GTT AAG GAA GCC ACT TCT TGG GAA ACC ACT CCG ATG GAT ATT GCC AAA GGA
Ile Ser Lys Ser Leu Ala Asp Arg Leu Cys Ile Ser Lys Val Asn Gly Gln Leu Trp Asp Leu Asp Arg Pro Phe
+301 ATT TCT AAA TCT TTA GCA GAC AGG TTA TGT ATT TCT AAG GTT AAT GGT CAA TTA TGG GAC TTA GAT AGA CCA TTT
DdeI DdeI
Glu Gly Glu Ala Asn Glu Glu Ile Lys Leu Glu Leu Leu Asp Phe Glu Ser Asp Glu Gly Lys Lys Val Phe Trp
+376 GAA GGC GAA GCC AAC GAA GAA ATC AAA TTA GAA CTG CTA GAT TTC GAA TCT GAC GAA GGT AAG AAG GTC TTT TGG
TaqI HinfI
His Ser Ser Ala His Val Leu Gly Glu Ser Cys Glu Cys His Leu Gly Ala His Ile Cys Leu Gly Pro Pro Thr
+451 CAT TCG TCT GCC CAC GTC TTG GGT GAA TCT TGT GAG TGC CAC CTA GGT GCC CAT ATT TGT TTA GGT CCT CCA ACT
HinfI
Asp Asp Gly Phe Phe Tyr Glu Met Ala Val Arg Asp Ser Met Lys Asp Ile Ser Glu Ser Pro Glu Arg Thr Val
+526 GAT GAT GGG TTC TTT TAT GAA ATG GCT GTT AGA GAT AGT ATG AAA GAT ATA TCT GAA TCT CCA GAA AGA ACC GTC
HinfI
Ser Gln Ala Asp Phe Pro Gly Leu Glu Gly Val Ala Lys Asn Val Ile Lys Gln Lys Gln Lys Phe Glu Arg Leu
+601 TCC CAA GCT GAT TTC CCA GGA TTA GAA GGC GTT GCC AAG AAT GTT ATC AAG CAA AAG CAA AAA TTT GAA AGA TTG
BstNI
Val Met Ser Lys Glu Asp Leu Leu Lys Met Phe His Tyr Ser Lys Tyr Lys Thr Tyr Leu Val Gln Thr Lys Val
+676 GTC ATG TCC AAA GAA GAT CTT TTG AAA ATG TTT CAT TAT TCG AAG TAT AAG ACT TAC TTG GTA CAG ACA AAG GTT
BglII TaqI RsaI
Pro Asp Gly Gly Ala Thr Thr Val Tyr Arg Cys Gly Lys Leu Ile Asp Leu Cys Val Gly Pro His Ile Pro His
+751 CCA GAT GGA GGT GCT ACT ACC GTC TAC CGT TGC GGT AAA TTG ATT GAT CTT TGT GTC GGC CCT CAT ATC CCA CAT
HaeIII
Thr Gly Arg Ile Lys Ala Phe Lys Leu Leu Lys Asn Ser Ser Cys Tyr Phe Leu Gly Asp Ala Thr Asn Asp Ser
+826 ACT GGG CGT ATC AAA GCT TTC AAA CTA TTA AAG AAC TCT TCT TGT TAT TTC TTA GGT GAT GCG ACA AAC GAC TCT
HindIII DdeI HinfI
Leu Gln Arg Val Tyr Gly Ile Ser Phe Pro Asp Lys Lys Leu Met Asp Ala His Leu Lys Phe Leu Ala Glu Ala
+901 TTA CAA AGA GTT TAC GGT ATC TCT TTT CCA GAC AAA AAA TTA ATG GAT GCT CAT TTG AAG TTC TTG GCG GAA GCC
Ser Met Arg Asp His Arg Lys Ile Gly Lys Glu Gln Glu Leu Phe Leu Phe Asn Glu Met Ser Pro Gly Ser Cys
+976 TCT ATG AGA GAT CAC AGA AAG ATT GGT AAA GAA CAA GAA TTA TTC TTA TTC AAT GAA ATG TCC CCA GGT TCT TGC
BstNI
Phe Trp Leu Pro His Gly Thr Arg Ile Tyr Asn Thr Leu Val Asp Leu Leu Arg Thr Glu Tyr Arg Lys Arg Gly
+1051 TTT TGG TTA CCT CAT GGT ACT AGA ATT TAC AAC ACT TTG GTT GAC TTG TTG AGA ACT GAA TAC CGT AAG AGA GGT
RsaI
Tyr Glu Glu Val Ile Thr Pro Asn Met Tyr Asn Ser Lys Leu Trp Glu Thr Ser Gly His Trp Ala Asn Tyr Lys
+1126 TAC GAA GAA GTC ATC ACT CCA AAC ATG TAC AAC TCC AAA TTG TGG GAA ACC TCA GGT CAC TGG GCC AAT TAC AAG
RsaI DdeI HaeIII
Glu Asn Met Phe Thr Phe Glu Val Glu Lys Glu Thr Phe Gly Leu Lys Pro Met Asn Cys Pro Gly His Cys Leu
+1201 GAA AAC ATG TTT ACT TTC GAA GTA GAG AAG GAA ACT TTC GGT CTA AAA CCA ATG AAC TGT CCA GGT CAT TGT TTG
TaqI BstNI
Met Phe Lys Ser Arg Gky Arg Ser Tyr Arg Glu Leu Pro Trp Arg Val Ala Asp Phe Gly Val Ile His Arg Asn
+1276 ATG TTC AAG TCT AGA GAA CGT TCT TAT AGA GAA TTG CCA TGG AGA GTT GCA GAC TTC GGT GTT ATC CAC AGA AAT
XbaI

```
        Glu Phe Ser Gly Ala Leu Ser Gly Leu Thr Arg Val Arg Arg Phe Gln Gln Asp Asp Ala His Ile Phe Cys Thr
+1351   GAA TTT TCT GGT GCC TTG TCT GGT TTG ACT CGT GTC AGA AGA TTC CAA CAA GAT GAT GCT CAT ATC TTC TGT ACC
                                          HinfI           HinfI                                   RsaI

        His Asp Gln Ile Glu Ser Glu Ile Glu Asn Ile Phe Asn Phe Leu Gln Tyr Ile Tyr Gly Val Phe Gly Phe Glu
+1426   CAT GAC CAA ATT GAA TCT GAA ATT GAA AAC ATT TTC AAC TTT TTG CAA TAC ATT TAC GGC GTT TTT GGA TTT GAA
                      HinfI

        Phe Glu Met Glu Leu Ser Thr Arg Pro Glu Lys Tyr Val Gly Lys Ile Glu Thr Trp Asp Ala Ala Glu Ser Lys
+1501   TTT GAA ATG GAA TTG TCC ACT AGA CCA GAA AAG TAC GTT GGA AAG ATC GAA ACC TGG GAT GCC GCT GAA TCA AAA
                                          RsaI              TaqI   BstNI

        Leu Glu Ser Ala Leu Lys Lys Trp Gly Gly Asn Trp Glu Ile Asn Ala Gly Asp Gly Ala Phe Tyr Gly Pro Lys
+1576   TTA GAA TCT GCC TTA AAG AAA TGG GGT GGT AAC TGG GAG ATC AAT GCT GGT GAT GGT GCT TTC TAC GGT CCA AAG

        Ile Asp Ile Met Ile Ser Asp Ala Leu Arg Arg Trp His Gln Cys Ala Thr Ile Gln Leu Asp Phe Gln Leu Pro
+1651   ATT GAC ATT ATG ATT TCT GAC GCT TTA AGA AGA TGG CAT CAA TGT GCC ACC ATC CAA TTA GAT TTC CAA TTG CCA

        Asn Arg Phe Glu Leu Glu Phe Lys Ser Lys Asp Gln Asp Ser Glu Ser Tyr Glu Arg Pro Val Met Ile His Arg
+1726   AAC AGG TTC GAA TTG GAA TTT AAA TCT AAA GAT CAA GAT AGC GAG AGT TAC GAA AGA CCG GTC ATG ATC CAT CGT
              TaqI

        Ala Ile Leu Gly Ser Val Glu Arg Met Thr Ala Ile Leu Thr Glu His Phe Ala Gly Lys Trp Pro Phe Trp Leu
+1801   GCC ATT TTA GGG TCT GTT GAA AGA ATG ACT GCC ATT TTG ACC GAG CAT TTT GCT GGT AAA TGG CCA TTT TGG TTA
                                                                                    HaeIII

        Ser Pro Arg Gln Val Leu Val Val Pro Val Gly Val Lys Tyr Gln Gly Tyr Ala Glu Asp Val Arg Asn Lys Leu
+1876   TCA CCA CGT CAA GTT TTG GTT GTA CCA GTT GGT GTC AAG TAC CAA GGG TAT GCT GAA GAC GTC CGT AAC AAA TTG
                                  RsaI

        His Asp Ala Gly Phe Tyr Ala Asp Val Asp Leu Thr Gly Asn Thr Leu Gln Lys Lys Val Arg Asn Gly Gln Met
+1951   CAC GAT GCT GGC TTC TAT GCC GAT GTC GAT TTG ACT GGT AAC ACT CTG CAA AAG AAG GTC AGA AAC GGG CAA ATG
                                          TaqI

        Leu Lys Tyr Asn Phe Ile Phe Ile Val Gly Glu Gln Glu Met Asn Glu Lys Ser Val Asn Ile Arg Asn Arg Asp
+2026   CTA AAA TAT AAC TTC ATT TTT ATT GTT GGT GAA CAA GAA ATG AAT GAA AAA TCT GTT AAC ATT AGA AAC AGA GAC

        Val Met Glu Gln Gln Gly Lys Asn Ala Thr Val Ser Val Glu Glu Val Leu Lys Gln Leu Arg Asn Leu Lys Asp
+2101   GTT ATG GAA CAA CAG GGT AAA AAT GCT ACT GTT TCT GTT GAA GAG GTT TTG AAA CAG TTG CGT AAC TTG AAA GAT

        Glu Lys Arg Gly Asp Asn Val Leu Ala Och
+2176   GAA AAG AGA GGT GAC AAC GTC TTA GCT TAA TGAGATTTT ATGTAGTTAAATTTTGACTTATTTAATTTATGTTTTGTAAAGAATAAAAT
                                  DdeI

+2264   GAGTATGATGAAGGAATCTTACCTCTATTAACTATTTCTTCTTCTTTTT CCTTCGATAAGATGTGCGCGAATTC
                  HinfI                                   TaqI              EcoRI
```

Fig. 4. Nucleotide sequence of the pUC18/THS1 insert. The sequence shown is that of the sense strand. The amino acid sequence of the reading frame from nucleotides +1 to +2202 is shown above the DNA sequence.

approximately 550 bp larger. The second larger *Eco*RI fragment is consistent with the replacement of one wild-type copy of *THS1* with the disrupted gene. This was confirmed by the results of the *Eco*RI plus *Pst*I digest. Since *Pst*I cleaves the *HIS3* but not the *THS1* gene, the larger 4.2 kb *Eco*RI fragment in the transformant should be reduced to a 2.5 kb *Eco*RI-PstI fragment. Lane 5 shows that genomic DNA from the transformant has the expected 3.6 kb *Eco*RI fragment and in addition the smaller 2.5 kb *Eco*RI-*Pst*I fragment.

The diploid strain W303Vths1 with the disrupted threonyl-tRNA synthetase gene was sporulated and tetrads were dissected. If *THS1* codes for a cytoplasmic synthetase, each tetrad should have only two viable spores, both being auxotrophic for histidine. If, on the other hand, the gene product is a mitochondrial aminoacyl-tRNA synthetase, all four meiotic spore progeny should be viable, but the two prototrophic for histidine would be expected to be respiratory deficient. Out of 8 tetrads dissected, 5 had 2 viable spores. The other 3 tetrads yielded only one viable spore each. All the spores, however,

TABLE I

Codon Usage in *THS1*

UUU	Phe	18	UCU	Ser	28	UAU	Tyr	8	UGU	Cys	9
UUC	Phe	21	UCC	Ser	7	UAC	Tyr	16	UGC	Cys	3
UUA	Leu	22	UCA	Ser	3	UAA	Ter	1	UGA	Ter	0
UUG	Leu	30	UCG	Ser	2	UAG	Ter	0	UGG	Trp	13
CUU	Leu	2	CCU	Pro	4	CAU	His	13	CGU	Arg	10
CUC	Leu	0	CCC	Pro	0	CAC	His	6	CGC	Arg	0
CUA	Leu	5	CCA	Pro	22	CAA	Gln	23	CGA	Arg	0
CUG	Leu	3	CCG	Pro	2	CAG	Gln	4	CGG	Arg	0
AUU	Ile	21	ACU	Thr	19	AAU	Asn	11	AGU	Ser	6
AUC	Ile	13	ACC	Thr	8	AAC	Asn	22	AGC	Ser	1
AUA	Ile	1	ACA	Thr	2	AAA	Lys	32	AGA	Arg	27
AUG	Met	22	ACG	Thr	0	AAG	Lys	32	AGG	Arg	2
GUU	Val	24	GCU	Ala	18	GAU	Asp	30	GGU	Gly	31
GUC	Val	18	GCC	Ala	17	GAC	Asp	14	GGC	Gly	5
GUA	Val	4	GCA	Ala	3	GAA	Glu	56	GGA	Gly	6
GUG	Val	0	GCG	Ala	2	GAG	Glu	8	GGG	Gly	5

were auxotrophic for histidine and were respiratory competent. These results and our inability to demonstrate disruption of *THS1* in the haploid strains of W303 indicate that the product of *THS1* is essential for cell viability and therefore functions in cytoplasmic protein synthesis. Whether this synthetase also functions in charging of the mitochondrial $tRNA_2^{Thr}$ is unclear at present. All the currently known genes coding for mitochondrial tRNA synthetases (those coding for the $threonyl_1$-, tryptophanyl- , aspartyl-, α subunit of phenylalanyl-, and tyrosyl-tRNA synthetases) are unique and, when inactivated, result in a subsequent loss of the corresponding mitochondrial synthetase alone (6,20; A. Gampel, J. Hill, T. J. Koerner, A. M. Myers, and A. Tzagoloff, unpublished data).

Homology of the *THS1* product to the mitochondrial and *E. coli* threonyl-tRNA synthetases

The sequence of the protein encoded in *THS1* is homologous to both *E. coli* threonyl-tRNA synthetase (21) and to the mitochondrial $tRNA_1^{Thr}$ synthetase (6). This is illustrated in the two dot matrices shown in Fig. 6. The homology with the bacterial enzyme extends over almost the entire lengths of the two polypeptide chains. The two proteins share 246 identical residues with a total of only 8 gaps in the alignment. This represents 38% homology. The yeast protein, however, has an extra 49 amino terminal residues for which there is no match in the sequence of the *E. coli* synthetase. It also has an extra 12 amino acids at the carboxyl terminal end. The yeast methionyl-tRNA synthetase also has additional residues at its amino terminus not present in the *E. coli* counterpart; 191 amino acids in this case (22).

Previous studies indicated that the $tRNA_1^{Thr}$ of yeast mitochondria is considerably shorter than the *E. coli* threonyl-tRNA synthetase (6). The discrepancy in the sizes of the two proteins is due mainly to the absence in the mitochondrial synthetase of some 200 amino terminal residues. This is also evident in the dot matrix of the two yeast proteins. Even though there is a clear line of homology, the $tRNA_1^{Thr}$ synthetase is 300 residues shorter at the amino terminal end. An alignment of the two sequences shows that the first unambiguous homology starts at residue 46 of $tRNA_1^{Thr}$ synthetase and residue 337 of the *THS1* product (Fig. 7).

A more informative comparison of the three protein sequences is presented in Fig. 7 where the

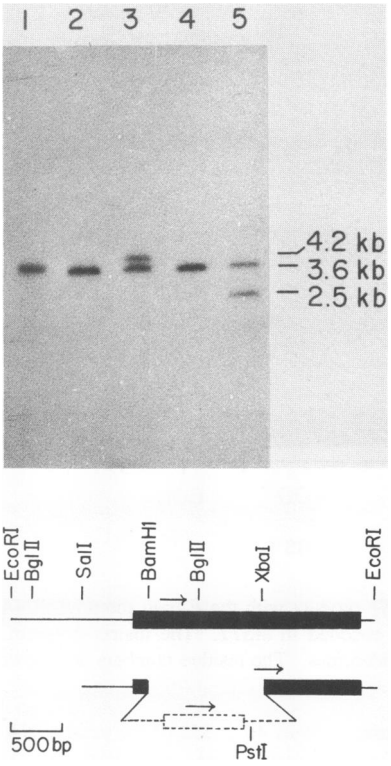


Fig. 5. Southern hybridization analysis of genomic DNA from the wild-type a/α and aW303 strains and from the a/α W303∇ths1 transformant. Purified chromosomal DNAs from the three strains digested with *Eco*RI alone and with *Eco*RI plus *Pst*I were separated electrophoretically on a 1% agarose gel. Following transfer to nitrocellulose (13), the Southern blot was hybridized with a nick-translated 1.3 kb *Eco*RI-*Bam*HI fragment that included 5' non-coding sequences of the *THS1* gene and 127 bp of coding sequences. The hybridization was carried out overnight at 65° C in 6x SSC, 1% sodium lauroyl sarcosyl, 0.1 mg/ml salmon sperm DNA. The blot was washed once with 2x SSC, 0.1% SDS and three times with 5 mM Tris-Cl pH 8 at room temperature. Lane 1: a/α W303 digested with *Eco*RI; lane 2: aW303 digested with *Eco*RI; lane 3: a/α W303∇ths1 digested with EcoRI; lane 4: aW303 digested with *Eco*RI and *Pst*I; lane 5: a/αW303∇ths1 digested with EcoRI and *Pst*I. lane 5: a/α W303∇ths1 digested with *Eco*RI and *Pst*I. The sizes of the *Eco*RI and EcoRI-*Pst*I fragments containing the *THS1* gene are marked in the margin.

The lower half of the figure illustrates the constructed ∇*ths1* gene. The *THS1* gene and flanking regions are drawn with a solid line. The inserted fragment of DNA with the *HIS3* gene is depicted by a broken line. The direction of transcription of the genes is indicated by the arrows. The probe used for the hybridization was the 5' *Eco*RI-*Bam*HI fragment.

alignment has been optimized based on identical residues. A visual inspection of the aligned sequences indicates that the most conserved regions occur in the central part of the proteins where there are numerous blocks of identical residues. In order to minimize errors introduced by faulty alignments, the analysis of the protein sequences has been confined to the regions spanning residues 251 to 544 of the

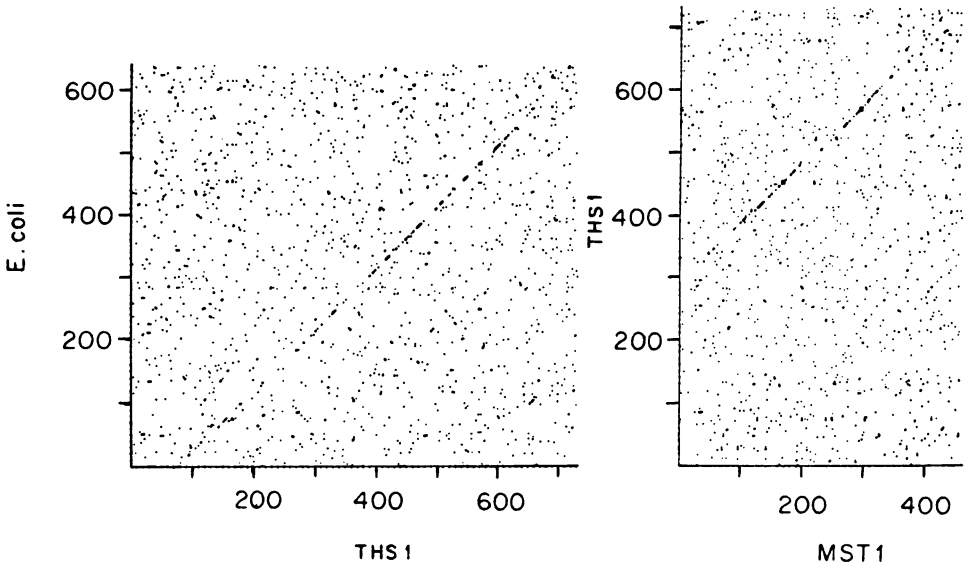


Fig. 6. Comparison of the *THS1* product with the *E. coli* threonyl-tRNA synthetase and with the mitochondrial $tRNA_1^{Thr}$ synthetase encoded in *MST1*. The matrix program used for the analysis scores a dot for every two consecutive identities. The residue numbers are shown on the abscissa and ordinate.

E. coli threonyl-tRNA synthetase, residues 46 to 352 of the yeast $tRNA_1^{Thr}$ synthetase and residues 337 to 633 of the *THS1* product. As can be seen in Fig. 7 these are the most homologous domains of the synthetases requiring the fewest deletions/insertions for alignment. The number of identities amongst the three proteins is summarized in Table II. Of the two yeast synthetases, the product of *THS1* is more homologous to the bacterial synthetase (47.6% identities) than the $tRNA_1^{Thr}$ synthetase (44.2% identities). The *THS1* product and the bacterial synthetase share 140 identical amino acids and differ in length by only two residues. In the same region the mitochondrial $tRNA_1^{Thr}$ synthetase had 130 identical amino acids with the *E. coli* synthetase. In addition, the mitochondrial protein is 13 residues longer necessitating 4 gaps in the alignment of the two sequences. An even more striking homology (52.2%) exists between the two yeast proteins.

The greater divergence of the $tRNA_1^{Thr}$ synthetase is not unexpected. The $tRNA_1^{Thr}$ has several unconventional features not seen in other threonine tRNAs. In addition to its leucine anticodon, it also has an extra nucleotide in the anticodon loop. Since the anticodon is one of the important factors in determining the specificity with which synthetases recognize their tRNA substrates (24,25), part of the domain involved in binding of $tRNA_1^{Thr}$ must be different in this synthetase. A comparison of the three sequences reveals a number of regions that are unique to the $tRNA_1^{Thr}$ synthetase. In the absence of information about the tRNA binding domains of synthetases in general, their significance is no clear.

In vitro aminoacylation of cytoplasmic tRNAs with wild-type and transformant extracts

Further confirmation that *THS1* codes for the cytoplasmic threonyl-tRNA synthetase was obtained by measuring the *in vitro* acylating activity of wild type extract and of an extract obtained from a

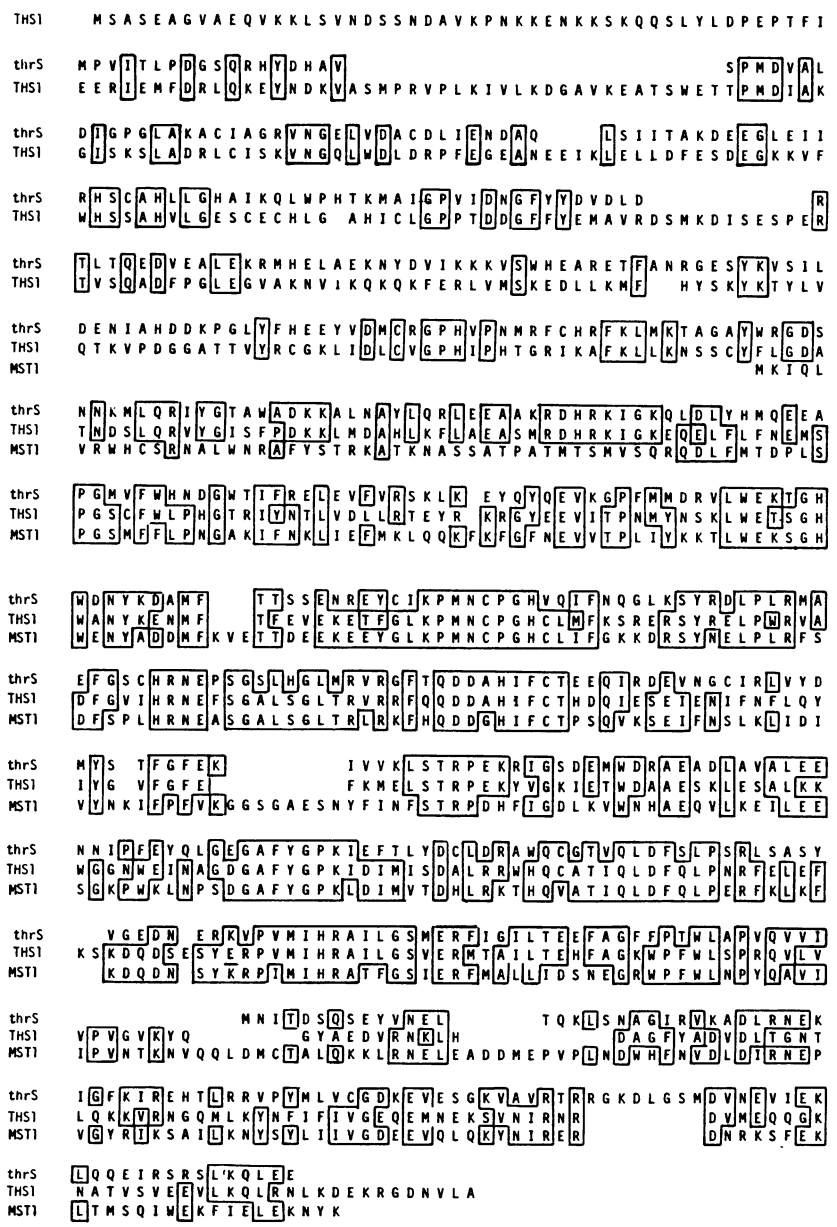


Fig. 7. Protein sequence homologies of the two yeast threonyl-tRNA synthetases and of *E. coli* threonyl-tRNA synthetase. The alignments of the three proteins are based on the MFALGO alignment program (23). Identical residues in the synthetases encoded by the two yeast genes (*MST1*, *THS1*), and by the *E. coli* gene (*thrS*) have been boxed.

TABLE II

Primary Sequence Homologies Between the Two Yeast and the *E. coli* Threonyl-tRNA Synthetases

Comparison	No. of Identities	% Homology
E. coli/THS1	140	47.6
MST1/THS1	158	52.2
E. coli/MST1	130	44.2
E. coli/MST1/THS1	108	36.7

The data in this table are based on alignments of the sequences included between residues 251-544 of *E. coli* threonyl-tRNA synthetase, residues 337-534 of the *THS1* encoded synthetase and residus 46-352 of the synthetase encoded in *MST1*.

strain with the *THS1* gene on an autonomously replicating plasmid. The 3.6 kb *Eco*RI fragment containing the *THS1* gene was transferred into the vector YEp351 (J. Hill, T. J. Koerner and A. Tzagoloff, personal communication). This construct, designated YEp351/THS1, was introduced into a wild-type strain auxotrophic for leucine, αW303-1B. A leucine independent transformant verified to harbor YEp351 was used to prepare total cytoplasmic aminoacyl-tRNA synthetases. A similar synthetase preparation was obtained from the parental strain, αW303-1B. Cytoplasmic tRNAs, isolated from a ρ^0 strain and hence devoid of any mitochondrial tRNAs, were acylated in the presence of L-[^{3}H] threonine with varying concentrations of the wild-type αW303-1B extract and with the comparative extract obtained from the transformant with the *THS1* gene on the multiple copy vector YEp351. At low concentrations (1-3 μg protein per 50 μl reaction), the extract obtained from the transformant was consistently more active in acylating threonine tRNA than the wild-type extract (Table III). Saturating

TABLE III

In vitro Aminoacylation of Cytoplasmic tRNAs with Extracts from W303-1B and W303/T1

μg protein	Extract: W303-1B	W303/T1*	Ratio of T1/wild-type
Experiment I			
0	146	248	-
1	1,624	6,344	3.9
2	2,426	10,628	4.4
3	4,694	14,318	3.1
5	7,106	11,706	1.6
7.5	10,542	18,208	1.7
Experiment II			
0	63	43	-
3	3,448	9,404	2.7
10	10,342	13,472	1.3
25	15,228	13,513	0.9
50	13,951	10,555	0.8

*W303/T1 designates the αW303-1B transformant harboring the multi-copy plasmid YEp351/THS1.
After 30 min at 30° C in a 50 μl reaction mix consisting of 0.1 M Tris-Cl pH 7.5, 5 mM ATP, 15 mM $MgCl_2$, 0.2 mM DTT, 50 μg cytoplasmic tRNAs, 0-50 μg protein, and 2.5 μCi [^{3}H] threonine (ICN, 20 Ci/mmol), the tRNAs were precipitated in 6% TCA, washed in 5% TCA and 80% ethanol, and counted. Table reports cpm recovered in the TCA precipitation.

levels of the cytoplasmic threonyl-tRNA synthetase were reached at a lower protein concentration in the transformant. These results lend further support for the assignment of *THS1* as the gene that codes for the cytoplasmic threonyl-tRNA synthetase.

ACKNOWLEDGEMENTS

We wish to thank Michael Levine for helpful suggestions pertaining to heterologous hybridization experiments. This research was supported by a National Science Foundation Grant PCM-8116680.

REFERENCES

1. Hartwell, L. H. and McLaughlin (1968) Proc. Natl. Acad. Sci. U.S.A. *59*, 422-428.
2. McLaughlin, C. S. and Hartwell, L. H. (1969) Genetics *61*, 557-566.
3. Nass, G. and Poralla, K. (1976) Molec. gen. Genet. *147*, 39-43.
4. Ramos, F. and Wiame, F.-M. (1979), Eur. J. Biochem. *94*, 409-417.
5. Mitchell, A. P. and Ludmerer, S. W. (1984) J. Bacteriol. *158*, 530-534.
6. Pape, L. K., Koerner, T. J. and Tzagoloff, A. (1985), J. Biol. Chem., in press.
7. Myers, A. M., Pape, L. K. and Tzagoloff, A. (1985) EMBO J., in press.
8. Southern, E. (1976) Methods in Enzymol. *68*, 152-176.
9. Norrander, J., Kempe, T. and Messing, J. (1983) Gene *26*, 101-106.
10. Grunstein, M. and Hogness, D. S. (1975) Proc. Natl. Acad. Sci. U.S.A. *72*, 3961-3965.
11. McGinnis, W., Levine, M. S., Hagen, E., Kuroiwa, A., and Gehring, W. H. (1984) Nature *308*, 428-433.
12. Beckmann, J. S., Johnson, P. F., and Abelson, J. (1977) Science *196*, 205-208.
13. Southern, E. M. (1975) J. Mol. Biol. *98*, 503-517.
14. Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. *74*, 504-564.
15. Macino, G. and Tzagoloff, A. (1979) Molec. gen. Genet. *169*, 183-188.
16. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall R. J. (1951) J. Biol. Chem. *193*, 265-275.
17. Rothstein, R. J. (1983) Methods in Enzymol. *101*, 202-211.
18. Vieira, J. and Messing, J. (1982) Gene *19*, 259-268.
19. Struhl, K., Stinchcomb, D. T., Scherer, S., and Davis, R. W. (1979) Proc. Natl. Acad. Sci. U.S.A. *76*, 1035-1039.
20. Myers, A. M. and Tzagoloff, A. (1985), J. Biol. Chem., in press.
21. Mayaux, J.-F., Fayat, G., Fromant, M., Springer, M., Grunberg-Manago, M., and Blanquet, S. (1983) Proc. Natl. Acad. Sci. U.S.A. *80*, 6152-6156.
22. Walter, P., Gangloff, J., Bonnet, J., Boulanger, Y., Ebel, J.-P., and Fasiolo, F. (1983) Proc. Natl. Acad. Sci. U.S.A. *80*, 2437-2441.
23. Doolittle, R. F. (1981) Science *214*, 149-159.
24. Bruce, A. G. and Uhlenbeck, O. C. (1982) Biochem. *21*, 3921-3926.
25. Schulman, L. H. and Pelka, H. (1984) Fed. Proc. *43*, 2977-2980.