Yeast omnipotent suppressor SUP1 (SUP45): nucleotide sequence of the wildtype and a mutant gene

Peter Breining and Wolfgang Piepersberg

Lehrstuhl für Mikrobiologie der Universität, Maria-Ward-Str. 1a, D-8000 München 19 and Fachgebiet Biochemie der Technischen Hochschule Darmstadt, Petersenstr. 22, D-6100 Darmstadt, FRG

Received 29 April 1986; Revised 12 June 1986; Accepted 18 June 1986

ABSTRACT
The primary structures of the yeast recessive omnipotent suppressor gene SUP1 (SUP45) and one of its mutant alleles (sup1-ts36) was determined. The gene codes for a protein of 49 kD. The mutant protein differs from the wildtype form in one amino acid residue (Ser instead of Leu) in the N-terminal part. The codon usage differs significantly from that of yeast ribosomal protein genes. However, an upstream element resembling a conserved oligonucleotide in the region 5' to ribosomal protein genes in S. cerevisiae has been found. A DNA probe internal to the SUP1 gene does not exhibit detectable homology to genomic DNA neither from higher eucaryotes nor from eu- or archaeabacteria. The hypothetical function of this protein in control of translational fidelity is discussed.

INTRODUCTION
The phenomenon of omnipotent recessive suppression in Saccharomyces cerevisiae has been known for long years (1), but was never elucidated on a biochemical basis. Mutations resulting in this phenotype were mapped at two genetic loci, SUP1 (SUP45) and SUP2 (SUP35), on chromosomes IIR and IVR, respectively, and induce pleiotropic effects on various cellular functions (2,3). Similarities with certain bacterial (Escherichia coli) mutants altered in ribosomal proteins S4 (ramA) and S5 (ramC) (4), and with another omnipotent but dominant suppressor mutation in yeast (SUP46), altered in a small ribosomal subunit protein (5), suggested that SUP1 and SUP2 genes may code for ribosomal proteins. Recently the SUP1 structural gene has been cloned from S. cerevisiae (6), and a 1.6 kb transcript and a 54 kd protein have been identified (7). Sequence determination revealed an unsplit gene coding for a protein of 49 kd molecular weight and deviating in codon usage significantly from the one typical for...
ribosomal protein genes and other highly expressed genes in yeast (8, 9). The DNA sequence of a mutant gene (sup1-ts36) (6) differed in several nucleotides from that of the wildtype only one of which led to an amino acid exchange in the N-terminal segment of the sup1 protein.

MATERIALS AND METHODS
Organisms, plasmids and transformation
Yeast strain 7B-D244, Escherichia coli HB101 (as a plasmid re-
ceptor), the derivation of plasmids pPBM10 (SUP1) and pPBM25
(sup1-ts36), and the conditions for transformation of DNA were
described earlier (6). Plasmids pPBM102 and pPBM1022 were con-
structed by replacing the small SalI/XbaI fragment of pACYC184
(10) by the 2.6 kb SalI/XbaI fragments from pPBM25 and pPBM10,
respectively. As a yeast/E. coli shuttle vector YEpl3 (11) was
used.

Preparation and manipulation of DNA and RNA
Preparations of DNA, total cellular RNA and polyA+ mRNA were
performed as given elsewhere (6,12). Restriction endonuclease
cleavage, ligation and other enzymatic treatment was carried out
as recommended by the suppliers of the enzymes (Boehringer-Mann-
heim, New England Biolabs). Hybridization of radioactively
labelled DNA probes was done according to Southern (13). Samples
of genomic DNA were from the following sources: Methanococcus
vannielii DSM 1224, human placenta, mouse spleen (line C57B1/6)
and wheat were kindly provided by H. Auer and J. Hauber, Munich.

Sequence analysis and S1 mapping
The nucleotide sequences of the cloned SUP1 alleles were deter-
mined using the chemical modification and cleavage method (14).
Nuclease S1 mapping of the 5' end of the SUP1 mRNA was deter-
mined and interpreted as described by others (12, 15-17).

In vitro translation
Translation of mRNA selected as hybrids towards the 1060 bp
BglII/BamHI fragment, internal to the SUP1 coding region, was
carried out in a rabbit reticulocyte system (Amersham-Buchler)
according to published procedures (12).

Computer analysis
The protein sequence data base was that of the PIR, release 4.0,
Fig. 1. Strategy for sequencing the gene SUP1 (single-lined arrows) and its mutant allele (sup1-ts36) (double-lined arrows) from S. cerevisiae. The bar represents the SUP1 coding region (see Fig. 2). Restriction endonuclease cleavage sites are symbolized as follows: Ac = AccI, B = BamHI, Bg = BglII, E = EcoRI, EV = EcoRV, Hp = HpaI, Mb = MboI, X = XhoI, Xb = XbaI; brackets mean that not all of the sites for the respective enzyme are shown; the numbered vertical arrows mark the location of HindIII sites introduced into the wildtype sequence (see Materials and Methods).

National Biomedical Research Foundation (Washington, D.C., USA). All comparisons were run on a VAX/VMS.

RESULTS AND DISCUSSION
The cloning, subcloning and identification of the SUP1 (SUP45) structural gene of S. cerevisiae and of its mutant allele sup1-ts36 was described earlier (6). Because of lacking information about the nature of the gene product, we decided to sequence the gene localized on a 2.6 kb BamHI/XbaI fragment. The strategies for sequencing both the wildtype and mutant (sup1-ts36) alleles are illustrated in Fig. 1. The sequence derived contained only a single long open reading frame coding for a hypothetical polypeptide of 438 amino acids and about 49kD molecular weight (Fig. 2). This is in agreement with the finding of a single RNA band of 1.5 to 1.6 kb approximate size hybridizing against the coding segment (7; our own observations). In vitro a protein of 50 kD monomer size was translated from
hybrid selected polyA+ mRNA (not shown), and is probably identi-
cal to the protein (54 kD) shown by others to be translated
from a transcript read from the same chromosomal segment (7).
The stop codon (TAA) is the same as in most other yeast genes.
It is concluded that the open reading frame corresponds to the
SUP1 structural gene.

Most yeast ribosomal protein genes (15 out of 17 sequenced) con-
tain introns close to their 5' ends (8, 9). Especially a 100%
conserved oligonucleotide (TACTAAC) is present in all introns.
However, the SUP1 gene is unsplit, and no sequences resembling
those characteristic for introns in S. cerevisiae are found in
the whole sequence. The amino acid composition (Table 1) reveals
a more acidic nature of the SUP1 protein unlike most ribosomal
proteins. Also, the codon usage deviates considerably from the
one typically used in yeast ribosomal protein genes (Table 2),
or other highly expressed yeast genes (10). Together with the
observation that the SUP1 (SUP45) gene is expressed at much
lower rate as cytoplasmic ribosomal proteins (7) and the fact
that the protein translated is larger than all known ribosomal
proteins in S. cerevisiae, all this suggests that the earlier
hypotheses (13, 19) concerning the nature of the SUP1 gene pro-
duct have to be modified: If the protein is a translational
factor interacting with cytoplasmic ribosomes, it is not an
integral and stoichiometric component of the organelle. The
presumed SUP1 (SUP45) transcript was roughly mapped by others
to start between the HpaI and the BglII sites (7; cf. Fig.1).
S1 mapping, however, revealed that the transcript started left
of the BglII site (Fig.3) and thereby proved that the first
ATG in the open reading frame is most likely the gene start.

Fig. 2. Nucleotide and deduced amino acid sequence of the SUP1
gene of S. cerevisiae. The location of relevant restriction
sites is indicated. The arrows mark the transcription initi-
ation sites within the purin-rich block (36 nucleotides) 5' to
the initiation codon. The oligonucleotide resembling the con-
served element HOMOL1, found upstream of several ribosomal
protein genes of yeast (8, 9), is marked by a broken line.
Nucleotide exchanges found in the sup1-ts36 sequence are
written over the respective position of the wildtype sequence,
one of which results in an amino acid exchange from Leu to Ser
(pos. 34 of the protein sequence).
There are other features of DNA primary structure in the region upstream of the SUP1 gene supporting this conclusion: (i) the transcript starts in a poly-purin stretch, and (ii) its initiation site is preceded by a typical poly-pyrimidin block (pos. -37 to -74; Fig. 2). (iii) A TATA-box (around pos. -100) is

Table 1. Calculated amino acid composition of the yeast SUP1 protein

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>29</td>
<td>6.62</td>
</tr>
<tr>
<td>Arg</td>
<td>13</td>
<td>2.97</td>
</tr>
<tr>
<td>Asn</td>
<td>22</td>
<td>5.02</td>
</tr>
<tr>
<td>Asp</td>
<td>32</td>
<td>7.31</td>
</tr>
<tr>
<td>Cys</td>
<td>4</td>
<td>0.91</td>
</tr>
<tr>
<td>Gln</td>
<td>14</td>
<td>3.20</td>
</tr>
<tr>
<td>Glu</td>
<td>39</td>
<td>8.90</td>
</tr>
<tr>
<td>Gly</td>
<td>31</td>
<td>7.08</td>
</tr>
<tr>
<td>His</td>
<td>4</td>
<td>0.91</td>
</tr>
<tr>
<td>Ile</td>
<td>27</td>
<td>6.16</td>
</tr>
<tr>
<td>Lys</td>
<td>41</td>
<td>9.36</td>
</tr>
<tr>
<td>Leu</td>
<td>42</td>
<td>9.59</td>
</tr>
<tr>
<td>Met</td>
<td>6</td>
<td>1.37</td>
</tr>
<tr>
<td>Phe</td>
<td>23</td>
<td>5.25</td>
</tr>
<tr>
<td>Pro</td>
<td>10</td>
<td>2.28</td>
</tr>
<tr>
<td>Ser</td>
<td>25</td>
<td>5.71</td>
</tr>
<tr>
<td>Thr</td>
<td>24</td>
<td>5.48</td>
</tr>
<tr>
<td>Trp</td>
<td>2</td>
<td>0.46</td>
</tr>
<tr>
<td>Tyr</td>
<td>17</td>
<td>3.88</td>
</tr>
<tr>
<td>Val</td>
<td>32</td>
<td>7.31</td>
</tr>
<tr>
<td>total</td>
<td>438</td>
<td>99.77</td>
</tr>
</tbody>
</table>

Table 2. Codon usage of the yeast SUP1 gene and yeast ribosomal protein genes

<table>
<thead>
<tr>
<th>SUP1</th>
<th>YRP(a)</th>
<th>SUP1</th>
<th>YRP</th>
<th>SUP1</th>
<th>YRP</th>
<th>SUP1</th>
<th>YRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>UUU</td>
<td>Ser</td>
<td>UCU</td>
<td>Tyr</td>
<td>UAU</td>
<td>Cys</td>
<td>UGU</td>
</tr>
<tr>
<td>Phe</td>
<td>UUC</td>
<td>Ser</td>
<td>UCC</td>
<td>Tyr</td>
<td>UAC</td>
<td>Cys</td>
<td>UGC</td>
</tr>
<tr>
<td>Leu</td>
<td>UUA</td>
<td>Ser</td>
<td>UCA</td>
<td>Leu</td>
<td>UUC</td>
<td>Trp</td>
<td>UGG</td>
</tr>
<tr>
<td>Leu</td>
<td>UUG</td>
<td>Ser</td>
<td>UCG</td>
<td>Leu</td>
<td>UUG</td>
<td>Trp</td>
<td>UGG</td>
</tr>
<tr>
<td>Leu</td>
<td>CUU</td>
<td>Pro</td>
<td>CCA</td>
<td>His</td>
<td>CAU</td>
<td>Arg</td>
<td>CGU</td>
</tr>
<tr>
<td>Leu</td>
<td>CUC</td>
<td>Pro</td>
<td>CCC</td>
<td>His</td>
<td>CAC</td>
<td>Arg</td>
<td>CGC</td>
</tr>
<tr>
<td>Leu</td>
<td>CUA</td>
<td>Pro</td>
<td>CCA</td>
<td>Gln</td>
<td>CAA</td>
<td>Arg</td>
<td>CGA</td>
</tr>
<tr>
<td>Leu</td>
<td>CUG</td>
<td>Pro</td>
<td>CCG</td>
<td>Gln</td>
<td>CAG</td>
<td>Arg</td>
<td>CGG</td>
</tr>
<tr>
<td>Ile</td>
<td>AUU</td>
<td>Thr</td>
<td>ACU</td>
<td>Asn</td>
<td>AAU</td>
<td>Ser</td>
<td>AGU</td>
</tr>
<tr>
<td>Ile</td>
<td>AUG</td>
<td>Thr</td>
<td>ACC</td>
<td>Asn</td>
<td>AAC</td>
<td>Ser</td>
<td>AGC</td>
</tr>
<tr>
<td>Ile</td>
<td>AUA</td>
<td>Thr</td>
<td>ACA</td>
<td>Lys</td>
<td>AAA</td>
<td>Arg</td>
<td>AGA</td>
</tr>
<tr>
<td>Met</td>
<td>AUG</td>
<td>Thr</td>
<td>ACC</td>
<td>Lys</td>
<td>AAG</td>
<td>Arg</td>
<td>AGG</td>
</tr>
<tr>
<td>Val</td>
<td>GUU</td>
<td>Ala</td>
<td>GCU</td>
<td>Asp</td>
<td>GAA</td>
<td>Gly</td>
<td>GGU</td>
</tr>
<tr>
<td>Val</td>
<td>GUC</td>
<td>Ala</td>
<td>GCC</td>
<td>Asp</td>
<td>GAC</td>
<td>Gly</td>
<td>GGC</td>
</tr>
<tr>
<td>Val</td>
<td>GUA</td>
<td>Ala</td>
<td>GCA</td>
<td>Glu</td>
<td>GAA</td>
<td>Gly</td>
<td>GGA</td>
</tr>
<tr>
<td>Val</td>
<td>GUG</td>
<td>Ala</td>
<td>GCG</td>
<td>Glu</td>
<td>GAG</td>
<td>Gly</td>
<td>GGG</td>
</tr>
</tbody>
</table>

a) from 11 yeast ribosomal protein genes, compiled from Ref. 8.
Fig. 3. S1 nuclease mapping of the 5' end of the SUP1 mRNA. For duplex formation the 254 bp HaeIII/TaqI fragment from position -84 through +170 (see Fig. 2) was used. The sequencing reaction was started at the TaqI site; only the relevant section is shown.

present in a distance where this element is also found ahead of many other yeast genes, which is thought to have its signifi-
cance in the initiation of transcription by RNA polymerase II, and which gave rise to controversial discussions (20, 21). (iv)
In a distance (around pos. -240), where upstream of ribosomal protein genes a conserved oligo-nucleotide, called HOMOL1, was observed (3, 9), a very similar sequence was found (Fig. 2). The last finding also could mean that the gene may be recognized by factors modulating transcription of components of the transla-
tional apparatus, and, therefore, is somehow co-regulated to-
gether with ribosomal protein genes.
The mutant allele (sup1-ts36) from strain 73-D244 when sequenced (cf. Fig. 1, 2) showed several nucleotide exchanges in and out-
side the coding segment, which probably reflects that the allele was selected in a strain of S. cerevisiae differing in deri-
vation from the one (S233C) form which the wildtype allele was cloned (6; M.D. Ter-Avanesyan, personal communication). Only one nucleotide exchange (T to C; pos. +101 in Fig. 2), however, was non-conservative and led to an amino acid exchange (Leu to Ser) in the N-terminal part of the SUP1 protein. Additional prove, that the mutation was located in this segment of the gene,
came from a fragment exchange experiment (Fig. 4): Complementation in strain 7B-D244 of the sup1-ts36 mutation, resulting in a temperature-resistant and adenine-requiring phenotype (6), could only be achieved, when the BamHI/XbaI fragment from pPBM1025 was introduced into vector YEp13 (BamHI/XbaI) and transformed. The same fragment derived from pPBM1020, however, was not able to relieve the suppressor phenotype of the mutant allele. Therefore, the conclusion seems justified, that this segment within the SUP1 protein is of major importance for its function, and, when slightly altered, exerts a severe effect on various other cellular functions. Alterations immediately at the C-terminus, on the other hand, for instance fusions at the single BamHI site (replacing the 5 terminal amino acids) as in the sup1-ts36 complementing clone pPBM8 (6) do not seem to interfere with the proteins function.

The function of the SUP1 protein remains a mystery. Some elucidation could come from comparison of the nucleotide and amino acid sequences with the library of sequence data available. Interesting, though weak, homologies were found with some aminacyl-tRNA synthetase sequences: (i) Methionyl-tRNA synthetase
Fig. 5. Comparison of the SUP1 sequence with sequences of aminoacyl-tRNA synthetases. (A) Area of homology to the MetRS gene from S. cerevisiae (22). Asterisks mark homologous base pairs; amino acid homologies are boxed; the numbers refer to the amino acid residues within the protein. (B) Protein homology between TyrRS from Bacillus stearothermophilus (23) and E. coli (24) with SUP1. Homologous amino acids are indicated by asterisks; the overlined peptide indicates an area highly conserved in several aminoacyl-tRNA synthetases (23).

(MetRS) from yeast (22) contains a stretch of accordance in both nucleotide and polypeptide sequence, where the respective rading frames are in phase with each other (Fig.5A).(ii) Homology on the protein level exists between SUP1 sequence and a segment from tyrosyl-tRNA synthetases (TyrRS) from B. stearothermophilus (23) and E. coli (24), which is highly conserved between those two enzymes in primary and probably tertiary structure(25), and close to a sequence conserved in several tRNA synthetases (23) (Fig.5B). Also the predicted secondary structure of the SUP1 protein in this small segment is consistent with the one described for the TyrRS polypeptides, where this section could be part of the ATP binding site (26). (iii)Another weak homology exists between the N-terminal portions (when the first 100 amino acids were compared) of both the ras2 gene product (27) and the SUP1 protein. Again the N-terminal half is highly conserved between ras-like factors (27). All this could indicate that SUP1 protein is a nucleotide or tRNA binding factor controlling, directly or indirectly, the codon-anticodon fitting during cytoplasmic translation. Though present probably in sub-stoichiometric amounts relative to the 80S ribosomes (7), it would at least part of the time interact with nascent or mature ribosomes, as can be judged from the pleiotropic effects on ribosome structure and function it exerts (3, 18, 19).
Since factor SUP1 is an essential protein in yeast it could be conserved also in other organisms. However, when a DNA fragment internal to the coding region is used as a probe for hybridization against total genomic DNA, from human, mouse, plant, eubacterial (E. coli) or archaebacterial (Methanococcus vanniellii) origin, digested by various restriction enzymes, no homology was found (not shown). Therefore the factor is either unique to lower eucaryotes, or it was not highly conserved during evolution.

The gene isolation and the derived structure of the coded protein of yeast SUP1 factor will make it possible in future to express the gene product in bacteria, to purify and test it for its effects on in vitro translation in a highly purified yeast system. Also it will be possible now to test for its relationship to the gene product(s) encoded by SUP2 (SUP35) since this gene was cloned recently (A.P. Surguchov, personal communication).

ACKNOWLEDGEMENTS
We are grateful to U. Pilz for expert technical assistance. We thank H. Feldmann und P. Foertsch for help with the computer analysis. Many stimulating discussions with A.P. Surguchov are gratefully acknowledged. This work was supported by a grant from BMFT (PTB).

REFERENCES