The extended promoter of the gene encoding ribosomal protein S33 in yeast consists of multiple protein binding elements

Martien H.Herruer, Willem H.Mager, Thijs M.Doorenbosch, Peter L.M.Wessels, Trudy M.Wassenaar and Rudi J.Planta*

Biochemisch Laboratorium, Vrije Universiteit, de Boelelaan 1083, 1081 HV Amsterdam, The Netherlands

Received July 6, 1989; Revised and Accepted August 11, 1989

ABSTRACT:

At least 4 different, protein binding *cis*-acting elements are present in the upstream region of the S33-gene. The major protein binding site is situated between positions -148 and -163 relative to the ATG start codon. It binds a *trans*-acting factor designated SUF (S33 Upstream Factor). When yeast cells are growing on glucose, deletion of this site results in a decrease of transcription of 50%. Using ethanol as a carbon-source, deletion of the SUF-responsive site lowers the transcription as much as 80%. A second protein binding site is found between positions -85 and -105. Only extracts from glucose-grown cells contain a factor that is able to bind to this site in vitro. A third protein binding site was found using a protein extract from ethanol-grown cells. This site, which is located quite close to the transcriptional start site, is probably responsible for the 20% residual transcription when the SUF binding site is removed. Finally, a site far upstream was found, which binds a protein from both glucose-grown and ethanol-grown cells. This site may function as an upstream repression site which is only functional when a non-fermentable carbon-source is used. Taking these findings into account, we present a model for the carbon-source dependent transcription activation of the gene encoding S33.

INTRODUCTION

Ribosomal protein synthesis (rp-synthesis) in yeast is a highly regulated process encompassing the coordinate expression of over a hundred genes (for recent reviews see refs 1,2). The balanced supply of ribosomal proteins is primarily controlled at the level of transcription (1,2). For instance upon addition of glucose to a yeast culture growing on ethanol as a carbon-source, the transcriptional yield of all rp-genes is elevated about four-fold (3,4). The equimolar production of ribosomal proteins is maintained during such changes in nutritional conditions.

The majority of yeast rp-genes harbours one or two tandemly arranged transcription activation sites, RPG-boxes, at a distance of 250-450 bp upstream of the respective genes.(1,2). RPG-boxes have been identified as the recognition site for a ubiquitous protein factor TUF (5), also designated RAP1 (6) or GRF1 (7). It has been postulated that the coordinate transcriptional activation of yeast rp-genes is mediated through this common *trans*-acting factor interacting with the common *cis*-acting element (8). Evidence is accumulating, however, that additional factors may be involved. First of all, the RPG-box/TUF complex formation is not exclusive for rp-genes. Transcription of many other genes is promoted through the RPG-box in a cell-growth-



Fig. 1. In frame fusions between several parts of the S33-upstream region and the galK reporter gene.

In panel A the S33 gene is shown including its 5' flanking region. ----> indicates the direction of transcription.Fragments HE, PE, RE and XE were isolated from pBMCY76-3 and cloned in M13mp8. In panel B a detailed picture is given from the S33 upstream region between -163 and -58. The italic numbers above the sequence refer to end points of the Bal31 deletions which were started from the Xbal site, followed by blunt ligation with the HincII site in M13mp8. Using the M13-derived BgIII site and the EcoRI site, all fragments were cloned in frame with the coding region of the *E. coli* galK reporter gene on the yeast vector YCpR1-1 (panel C), cut with BamHI and EcoRI.

dependent fashion (for refs. see 9). Moreover, an RPG-box has been identified as part of the silencer of the mating type locus and the factor TUF also binds to telomeres, indicating that the actual role of TUF may vary depending on the nature of the sequences surrounding its binding site and on additional factors (6,10,11).

A second argument against the involvement of TUF exclusively in regulating rp-gene transcription, is the fact that at least two rp-genes do not contain an RPG-box (12). One is the L3 gene (TCM1) where a so-called TAF-responsive site present in the 5' flanking sequence has been shown to mediate transcriptional activation of this gene (13). This paper deals with the second example, namely the S33-gene for which also binding of a different factor (designated SUF) has been demonstrated (8). Evidence is available that TAF and SUF may be identical proteins (19). The present study is an analysis of the function of the SUF-binding site in the transcriptional regulation of the S33 gene and of other *cis*-acting elements and corresponding factors, that may play a part in this process.

MATERIALS AND METHODS

Recombinant plasmids

Plasmid pBMCY76-3 consists of a HindIII-generated fragment carrying the gene for ribosomal protein S33, cloned in pBR322 (17).

The CEN containing plasmid YCpR1, kindly provided by Dr. R.S. Zitomer is a *E. coli*-yeast shuttle vector carrying the *E. coli* galK gene under control of the cyc-1 promoter (15). In YCpR1-1 an EcoRI site originally located immediately downstream of the galK gene has been destroyed (H. Klein, this laboratory). S33-galK fusion genes used in this study are depicted in Fig. 1.

Carbon-source upshift

Transformants of the yeast strain HR-2 (tryp, leu, his, ura) were grown in a medium containing 0.67% yeast nitrogen base, 0.05% yeast extract, 0.04% glucose and 2% ethanol supplemented with 20 mg/l uracil, 50 mg/l histidine and 40 mg/l leucine until an OD₆₆₀ of 0.8 was reached. Then 0.1 volume of 20% glucose was added and rapidly mixed. Samples were taken at 0, 10, 20, 30 and 60 min after this carbon-source upshift. Cells were immediately frozen in liquid nitrogen and stored at -20°C until use.

Northern analysis of transcripts

RNA from transformed cells was analyzed as described previously (4) using an EcoRIgenerated fragment from YCpR6 (*i.e.* galK; Ref. 15) and a (HindIII + PvuII)- generated fragment of pBMCY76-3 (*i.e.* S33; Ref. 17) as probes.

Preparation of S100 extracts

Yeast transformants were grown until an OD_{660} of 0.8 was reached. Cells were broken with glass beads in a buffer containing 0.2 M Tris-HCl pH 8.0, 0.01 M EDTA, 0.1 M MgCl₂, 30 mM (NH₄)₂SO₄, 10% v/v glycerol, 2.5 mM β -mercaptoethanol and 0.01 M phenylmethylsulfonyl-



Fig. 2. In vivo analysis of fusion genes using glucose and ethanol as a carbon-source. Fusion constructs PE, RE, XE, 28, 40, 24, 7, 3, 11 and 8 were transformed to the yeast strain HR-2 and grown on glucose (panel a) and ethanol (panel b) as a carbon-source. RNA was extracted from logarithmic growing cells (see Materials and Methods) and a Northern hybridization was performed using galK- and S33-gene specific probes.

In panel c construct HE is compared with constructs PE and RE for cells growing on glucose and ethanol.

In panel d a graphical representation is shown of the galK mRNA levels relative to the S33mRNA internal standard.

fluoride. After removal of the glass beads the extract was centrifuged for 1 hr at 100,000 x g and the supernatant was frozen at -20° C until use.

Gel retardation and DNAse I footprint analysis

These analyses were performed as described by Huet et al. 1985 (5).

Bal31 nuclease digestion

Construct PE cut with XbaI was treated with Bal31 (Boehringer; 0.25 U per 10 μ g DNA) for 10 min at room temperature. After digestion with HincII, blunt end ligation was performed using T4-DNA ligase. The position of the deletions was determined by sequencing according to Sanger (16).

<u>RESULTS</u>

Deletion analysis of the S33 promoter.

Transcription activation of most, but not all, ribosomal protein genes in yeast is mediated through the ubiquitous protein factor TUF, interacting with its responsive element, the RPG-box (14).

Gel mobility shift assays and footprint analyses have revealed that the gene encoding ribosomal protein S33 does not contain a functional RPG-box (8). Another protein-binding DNA element with the nucleotide sequence 5'CCGTCTAGAGTGACCA3', however, is located between positions -148 and -163 (8). The respective protein factor was provisionally called SUF, for S33 upstream factor. Methylation interference studies indicated that the actual SUF-recognition sequence consists of a dual motif: CGTN₅GTGAC (19). To investigate the function of this DNA element in vivo, fusion genes were constructed consisting of parts of the upstream region of the S33-gene coupled to the galK-gene of E.coli (see Fig. 1). We used the EcoRI-site within the S33-gene for in-frame fusions with the galK-reporter gene, whereas upstream deletions were generated using four restriction sites: HindIII (HE), PvuII (PE), RsaI (RE) and XbaI (XE). Deletions downstream from the XbaI-site were made by Bal31 treatment of DNA cut with XbaI. Constructs XE, 28, 40 and 24 carry deletions within the SUF binding area. In constructs 7, 3, 11 and 8 also a pyrimidine-rich region (T-stretch) is deleted. Using Northern hybridisation with gene-specific probes we measured the steady state concentrations of the hybrid galK-mRNAs relative to rp-mRNA levels in the respective transformed cells grown on either glucose or ethanol as a carbon source. The results are presented in Fig. 2.

In panel a we present the results of a deletion experiment in yeast transformants growing on glucose as a carbon source. Under such nutritional conditions removal of the footprinting area has only a slight and gradual effect. A graphical representation is shown in panel d. Fusion gene 24 still gives rise to a considerable galK-mRNA level. Surprisingly, deletion of the distal part of the SUF binding site (XE) results only in a 30-40% decrease of the cellular galK-mRNA level. Removal of the complete binding site gives a decrease of more than 70%. This nucleotide element, thus, represents a *cis*-acting element, although a basal level of transcription remains after deletion of this element. Further deletion mutants in which also the T-stretch has been

Nucleic Acids Research



Fig. 3. Band shift assay using upstream fragments of the S33-gene. An S100 protein extract from logarithmic growing cells was incubated with the following probes for 15' at 25°C; lanes 1-3 HE, lanes 4-6 TT, lanes 7-9 PE, lanes 10-12 RE, lanes 13-15 XE, lanes 16-17 the S33 part of construct 40.

In lanes 1, 4, 7, 10 and 13 no protein was added. In lanes 2, 5, 8, 11, 14 and 16 protein from cells grown on glucose was added. In lanes 3, 6, 9, 12, 15 and 17 protein from cells grown on ethanol was added.

removed display a residual transcription activity of about 10%, whereas for construct 7 no transcription can be measured.

Panel b shows the results obtained in similar deletion experiments using transformed cells grown on ethanol. A graphical representation of these results is shown in panel d. A sharp drop in galK-mRNA concentration can be observed for those deletions in which only one half of the protein binding site is removed. However, the transcriptional level does not decrease significantly below the value observed with construct HE (panel c). Therefore, removal of the SUF-binding site has only a drastic effect on the galK-mRNA concentration for cells growing on ethanol when the upstream region (-320 to -1000) is absent. When the upstream region is intact, the SUF-binding site can not be identified as a genuine UAS. On the other hand the basal transcription activity under ethanol growth conditions remains more or less unaffected, even for deletion constructs 3 and 11.

Panel c shows the results obtained with constructs HE, PE and RE in cells growing on glucose and ethanol, respectively. Obviously, deletion of the far upstream sequence between -1000 and -320, gives rise to a strongly elevated transcription activity when ethanol is used as a carbon-



Fig. 4. Footprint analysis of the glucose dependent upstream factor-binding site.

The S33 fragment of construct 40 was used as a probe. The conditions for binding were optimized (see Results). After mild DNAse treatment the mixture was subjected to electrophoresis for a band shift assay. Free probe (F) and complex (C) bands were eluted from the gel and electrophoresed in a denaturing sequencing gel.

Lanes 1 and 2 represent Maxam and Gilbert sequencing reactions. In lanes 3 and 4 DNA is incubated with the partly purified protein extract, in lane 5 DNA is incubated without protein extract. The sequence protected from DNAse I digestion is indicated, the numbering referring to the translation start site.



Fig. 5.Band shift assay with a GDUF-binding oligonucleotide.

A synthetic oligo encompassing the DNAse I protected sequence (see Fig. 4) was incubated with a yeast protein extract that was fractionated over a heparin-sepharose column. The numbers 10-34 refer to the fractions obtained from this column.



source. This phenomenon might reflect the presence of an upstream repressor site in this region of the promoter.

Protein factors binding to the S33 gene promoter.

To account for the *in vivo* results we have to suppose that the S33-promoter consists of multiple cis-acting elements which may serve as binding sites for different transcription factors. Therefore we used the upstream DNA fragments present in the fusion constructs HE, TT, PE, RE, XE and 40 (cf. Fig. 1) for gel mobility shift assays with total yeast protein isolated from cells grown on glucose and ethanol, respectively. The results are shown in Fig. 3. The first lane of each set contains the free probe; each second lane represents the binding assay using protein from cells grown on glucose, whereas in each third lane protein from cells grown on ethanol is used. The most striking conclusion to be drawn from the results presented in Fig.3 is that the binding in vitro of proteins to the upstream DNA of this ribosomal protein gene is carbonsource dependent. The upper complex band is due to binding of SUF; this complex is formed with protein isolated from glucose-grown as well as from ethanol-grown cells. Apparently, SUF-binding capacity is lost from construct XE on. Two additional complex bands can be observed using protein isolated from cells grown on ethanol. One of these is designated EDUF (Ethanol Dependent Upstream Factor), and is formed with all probes used, suggesting that the binding site is located downstream of the SUF-binding site. The second one (X) is also formed with protein isolated from glucose grown cells. This complex band is only present when HE is used as a probe, suggesting that the responsive DNA element lies upstream of the SUF-binding site. We postulate that the factors identified in this way are involved in transcriptional regulation under ethanol growth conditions. Another complex is present only in those lanes where protein from glucose-grown cells is used. It is, therefore, designated GDUF (Glucose Dependent Upstream Factor). This complex, too, is present with all probes used and therefore the responsive DNA element is supposed to lie downstream of the SUF binding site and may be part of the glucose-specific site in front of the S33 gene. This glucose-specific protein binding site was investigated in more detail.

Footprinting of the glucose-specific protein binding site.

The gel mobility shift assay described above indicated that the binding site for the glucosespecific factor is located downstream of the SUF-binding site. In the initial experiments the pertinent protein-DNA interaction was rather weak. However we were able to increase complex

Fig. 6. Effect of a nutritional upshift on S33-galK fusion gene-expression.

Cellular levels of galK and \$33 mRNAs were measured during a carbon-source upshift. RNA was isolated from cells growing on ethanol as well as 10', 20', 30' and 60 min after glucose addition. Northern hybridization using gene-specific probes was performed as described in Materials and Methods.

In the graphical representation the galK-S33 ratio was plotted against the incubation time. In this way we were able to compare the carbon-source upshift response of several deletion mutants. A horizontal line reflects a response of galK fusion gene transcription which is similar to that of the S33-gene.

formation by omitting $MgCl_2$ from the incubation mixture and raising the $CaCl_2$ concentration to 5 mM. Using these conditions it became possible to carry out an indirect footprint analysis (Fig. 4) which showed a region within the sequence TACTTGGAATAAAGAAAATGAA to be protected. A synthetic oligonucleotide encompassing this sequence was able to bind a low molecular weight protein factor that was extracted from cells grown on glucose and subsequently fractionated using a heparin sepharose column, as is shown in Fig. 5.

Response of \$33-gene expression to a carbon-source upshift.

In an earlier report (4) we presented evidence that the RPG-box mediates the sudden rise in transcription of rp-genes upon a nutritional upshift. It was also shown then that the S33-gene displays the same kinetics of transcription activation observed for other rp-genes. The majority of rp-genes in yeast actually share a tripartite UAS, consisting of two adjacent RPG-boxes and a T-stretch further downstream (1,2). The latter element also occurs in the S33-promoter. Sofar we have not been able to detect a protein factor binding to this DNA element. To investigate the separate contributions of the SUF binding site, the T-stretch and the glucose-specific protein binding site to the transcriptional regulation of the S33-gene, we performed a carbon-source upshift experiment, using the constructs HE, PE, RE, XE, 28, 40, 24 and 7 (cf. Fig. 1). In these experiments we observed a characteristic carbon-source upshift response only in those cases where the SUF-protein binding site is still active (see Fig 6, constructs HE, PE and RE; giving rise to an almost horizontal line in the graphical representation). After removal of the SUF binding site the observed response is only 50% of the control level (constructs XE, 28, 40 and 24). This reduced carbon-source upshift response might be mediated by the glucose-specific protein binding site, possibly in combination with the T-stretch. Therefore we hypothesize that the normal response could be a cooperative effect of two or more *trans*-acting factors, one of which is SUF.

DISCUSSION

Deletion analysis of the 5' flanking sequences of the gene encoding ribosomal protein S33 in yeast revealed a rather complicated promoter structure. The binding site for the previously identified protein factor SUF appears to play an important part in promoting S33-gene transcription, but it certainly is not an UAS element with the same dominant role as the RPG-boxes located in front of most other rp-genes.

We have recently reported evidence that factor SUF, binding to the S33-gene promoter, and factor TAF, identified as a *trans*-activator of the gene encoding ribosomal protein L3 (TCM1) are very likely the same protein (19). Moreover, SUF and TAF are probably identical to the factors GF1, SBF-B and ABF-1, which were identified as protein factors binding to the upstream sequences of many nuclear genes encoding mitochondrial proteins (18), to the silencer region of the mating type locus (20) and to ARS-associated sequences (7), respectively.

For the interpretation of the role the SUF/TAF factor plays in the regulation of rp-gene transcription, it is relevant to compare the deletion analysis of the S33-gene promoter with that

factor			se	quence			
SUF TAF	-143 -236	GTGCG <u>TG</u> TAGTA	GTCA A <u>TCG</u>	CTCTAG TTTTGT	ACG ACG	GCCGCGTCTGTAC TTTTTCAAGAA	-175 -208
Consensus			RTCR	NNNNN	ACG		

Fig. 7. Comparison of nucleotide elements binding SUF and TAF. Comparison of the TAF-binding site and the SUF-binding site showing a consensus with a

Comparison of the TAF-binding site and the SUF-binding site showing a consensus with a bipartite motif. Competition experiments revealed that TAF and SUF likely are the same protein (19). Footprint areas are underlined.

reported for the L3 gene (13). The respective footprint areas are aligned in Fig.7 which also depicts the bipartite protein binding nucleotide motif identified by methylation interference studies. A proper comparison of the effect of deleting this site on the transcription activity, however, is hampered by the fact that the SUF binding site is present in the reverse orientation relative to the TAF-binding site. In addition, available data for the L3 gene are limited to glucose-growth conditions.

Extending the upstream deletions of the L3-gene promoter from -236 to -202 - which eliminates both parts of the split binding motif (cf. Fig. 7) - results in an almost 20-fold reduction of transcription activity. This deletion not only includes the TAF-binding site, however, but also encompasses 25 nucleotides further upstream. It can not be excluded, therefore, that apart from the TAF-site also an adjacent protein binding site has been removed. Deleting upstream sequences of the S33-gene down to -103 (construct 7) similarly reduces transcription activity to about 5%. We present evidence that in the region closely upstream of the SUF-binding site a binding site for an additional factor is present that may promote transcription activity in glucose-grown cells (designated GDUF). Only limited sequence homology exists between the GDUF binding area and sequences adjacent to the TAF-binding site. Furthermore, the detailed deletion analyses within the SUF-binding area strongly suggest that SUF may be dispensable for growth on glucose. In deletion construct 40, SUF is not able to bind in vitro but the corresponding decrease in transcription in vivo is only about 30%. These data indicate only a slight stimulatory effect of SUF under such steady state growth conditions. Further deletions of the SUF binding site cause additional loss of transcription activity to about 30% of the normal value. It remains to be seen whether this effect reflects the involvement of a so far undiscovered factor.

The abovementioned GDUF protein is proposed to mediate the residual transcription activity, whereas the T-stretch might also play a role. Transcription activation under ethanol growth conditions is clearly distinct from that in glucose-grown cells. We show evidence that the potential activating effect of SUF binding to its responsive site probably is suppressed by the action of a (so far unidentified) element located between -320 and -1000 (further upstream).

Nucleic Acids Research

Basal transcription under ethanol steady-state growth conditions is likely to be mediated by (a) carbon-source specific factor(s) binding downstream of the SUF-site.

On the other hand, the nutritional upshift experiments do suggest that SUF is required to achieve the rapid transcriptional response (Fig. 6). Band shift assays using S100 extracts from glucose- or ethanol-grown cells show that the cellular level of SUF may be the same under both conditions. In this respect SUF differs from the *trans*-acting factor TUF mediating transcription activity of the majority of rp-genes, of which the concentration appears to be growth-rate dependent (this laboratory, unpublished results). The data presented in this paper suggest that SUF, as a general primary DNA-binding protein, might facilitate the efficient assembly of a transcription initiation complex with RNA polymerase II during changing nutritional conditions. Additional, carbon-source-specific, protein factors likely are responsible for the growth-rate dependent transcription activity of these rp-genes. We postulate that these additional factors are also involved in the regulated transcription activation of the rp-genes under primary control of the TUF-RPG box combination.

ACKNOWLEDGEMENTS

The authors are indebted to their coworkers at the Biochemisch Laboratorium for stimulating discussions. We thank Mrs. A.C. van der Kuyl and Mrs. H.M. Kuiper for their technical assistance and Mrs. P.G. Brink for the finishing touch of the typescript.

This work was supported in part by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for Scientific Research (N.W.O.).

*To whom correspondence should be addressed

REFERENCES

- 1. Mager, W.H. (1988) Biochim. Biophys. Acta, 949, 1-15.
- 2. Planta, R.J. and Raué, H.A. (1988) Trends in Gen., 4, 64-68.
- 3. Donovan, D.M. and Pearson, N.J. (1986) Mol. Cell. Biol., 6, 2429-2435.
- 4. Herruer, M.H., Mager, W.H., Woudt, L.P., Nieuwint, R.T.M., Wassenaar, G.M., Groeneveld, P. and Planta, R.J. (1987) Nucl. Acids. Res., 15, 10133-10144.
- 5. Huet, J., Cottrelle, P., Cool, M., Vignais, M-L., Thiele, D., Marck, C., Bühler, J.M., Sentenac, A. and Fromageot, P. (1985) The EMBO J., 4, 3539-3547.
- 6. Shore, D. and Nasmyth, K. (1987) Cell, 51, 721-732.
- 7. Buchman, A.R., Kimmerly, W.J., Rine, J. and Kornberg, R.D. (1988) Mol. Cell. Biol., 8, 210-225.
- 8. Vignais, M-L., Woudt, L.P., Wassenaar, G.M., Mager, W.H., Sentenac, A. and Planta, R.J. (1987) The EMBO J., 6, 1451-1457.
- 9. Nieuwint, R.T.M., Mager, W.H., Maurer, C.T.C. and Planta, R.J. (1989) Curr.Genet, in press.
- 10. Kimmerly,W.J., Buchman,A.R., Kornberg,R.D. and Rine,J. (1988) The EMBO J.,7, 2241-2253.
- 11. Berman, J., Tachibana, C.Y. and Tye, B.K. (1986) Proc. Natl. Acad. Sci. USA, 83, 3713-3717.
- 12. Leer,R.J., van Raamsdonk-Duin,M.M.C., Mager,W.H. and Planta,R.J. (1985) Curr. Genet., 9, 273-277.
- 13. Hamil, K.G., Nam, H.G. and Fried, H.M. (1988) Mol. Cell. Biol., 8, 4328-4341.

- 14. Woudt,L.P., Smit,A.B., Mager,W.H. and Planta,R.J. (1986) The EMBO J.,5, 1037-1040. 15. Rymond,B.C., Zitomer,R.S., Schümperli,D. and Rosenberg,M. (1983) Gene,25, 249-262.
- Sanger, F., Nicklen, S. and Coulsen, A.R. (1979) Proc. Nat. Acad. Sci. USA, 74, 5463-5467.
 Leer, R.J., van Raamsdonk-Duin, M.M.C., Schopping, P.J., Cornelissen, M.T.E., Cohen, L.H., Mager, W.H. and Planta, R.J. (1983) Nucl. Acids Res., 11, 7759-7768.
- 18. Dorsman, J.C., van Heeswijk, W.C. and Grivell, L.A. (1988) Nucl. Acids Res., 16, 7287-7301.
- 19. Dorsman, J.C., Doorenbosch, M.M., Maurer, C.T.C, de Winde, J.H., Mager, W.H., Planta, R.J. and Grivell, L.A. (1989) Nucl. Acids Res. in press.
- 20. Shore, D., Stillman, D.J., Brand, A.H. and Nasmyth, K. (1987) EMBO J., 6, 461-467.