Volume 11 Number 22 1983 Nucleic Acids Research

Yeast ribosomal protein S33 is encoded by an unsplit gene

R.J.Leer, M.M.C.van Raamsdonk-Duin, P.J.Schoppink, M.T.E.Cornelissen, L.H.Cohen*, W.H.Mager and R.J.Planta

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

Received 23 September 1983; Revised and Accepted 25 October 1983

ABSTRACT

The structure of the gene coding for ribosomal protein S33, - a protein which escapes the coordinate control of ribosomal protein synthesis in rna 2 mutant cells -, was determined by sequence analysis. The gene comprises an uninterrupted coding region of 204 nucleotides encoding a protein of 8.9 kD. Like for other yeast ribosomal protein genes that have been sequenced so far, a relatively strong codon bias was observed. By S1 nuclease mapping the 5' end of the S33 mRNA was shown to be located at 11 to 15 nucleotides upstream from the initiation codon.

INTRODUCTION

Upon a temperature-shock ribosomal protein synthesis in yeast declines to about one third after 15 min (1,2) due to a temporary block of transcription (3). In wild-type cells the synthesis of ribosomal proteins proceeds within an hour at normal rate. However, in the temperature-sensitive mutant rna 2 no recovery of the ribosomal protein synthesis takes place. Ample evidence has been obtained that this is due to a lack of proper maturation of most of the ribosomal protein precursor mRNAs at the restrictive temperature (4-7). This phenomenon can be related to the fact that most yeast ribosomal genes studied so far contain an intervening sequence (4,6,8). Therefore the permanent decrease in the concentration of mature mRNAs for most ribosomal proteins in rna 2 cells can be attributed to a defect splicing mechanism leading to the accumulation of intron-containing pre-mRNAs at the restrictive temperature (3,4,6,7).

In this paper we describe the structure of a gene coding for a ribosomal protein, S33, the synthesis of which escapes the coordinate control. Determination of the S33 coding sequence revealed the unsplit character of this ribosomal protein gene. Due to the absence of an intron this gene is of particular interest, since its expression may be regulated in a different way.

EXPERIMENTAL

Isolation of nucleic acids

Recombinant plasmids pBMCY76, 76-3, 89 and 113 (described in Ref. 9) were purified from Triton-treated bacterial spheroplasts by CsCl-ethidium bromide density gradient centrifugation (10).

Poly A-containing RNA was isolated from the yeast *Saccharomyces carlsbergensis* S74 and the yeast *Saccharomyces cerevisiae rna 2* (ts 368^-) according to a procedure described previously (11).

Hybrid-selected in vitro translation

Single-stranded recombinant DNAs were bound to nitrocellulose filter papers as described by Kafatos *et al.* (12). Hybridization with poly A-containing RNA was performed overnight as described (13,14). Hybrid-selected mRNAs were melted off by incubation of the filters in 2 mM EDTA during 90 sec at 100°C and were then precipitated with ethanol. *In vitro* translation was performed in a rabbit reticulocyte lysate as described previously (9). Analysis of the translation products was carried out by electrophoresis in 10% polyacrylamide gels in the presence of SDS (9) or in mini-two-dimensional gels (15). Immunoprecipitation was performed as described recently (9).

Restriction site mapping

Restriction endonucleases were obtained from Bethesda Research Laboratories and New England Laboratories and used as recommended by the supplier.

Northern blot analysis

Poly A-containing RNA was fractionated on 1.6% agarose gels and transferred to DBM-paper (16), which was then hybridized with recombinant DNAs labelled *in vitro* by nick-translation (17).

DNA sequence analysis

DNA sequence analysis was performed using the chain termination method (18). Single-stranded templates were obtained by transforming JM101 cells with recombinant bacteriophage M13RF DNA (19). The M13 vectors used were M13mp8 or M13mp9 (20).

S1 nuclease analysis (21,22)

A single-stranded EcoRI *plus* XbaI-generated fragment (251 nucleotides; see Fig. 2) was 5'-end labelled at the EcoRI site using (γ-^{32}P)-ATP and T4 polynucleotide kinase (Boehringer). 30 µg of poly A-containing RNA and 0.5 pmol probe (40.000 cpm) were annealed at 48°C overnight in 16 µl 80% formamide, 40 mM PIPES (pH 6.5), 400 mM NaCl, 1 mM EDTA. The mixture was then diluted 15-fold into 0.28 M NaCl, 5 mM $ZnSO_4$, 50 mM sodium acetate (pH 4.5),

containing S1-nuclease amounts of 12-200 U (Sigma). Digestion was performed for 1½ h at 25°C. Reactions were terminated by the addition of 60 µl 4 M ammonium acetate, 0.3 M EDTA (pH 8.0) followed by ethanol precipitation. The pellets were washed with 96% ethanol, air-dried, dissolved in formamide/dye solution, boiled for 2 min and then analyzed on a 8% polyacrylamide/urea sequencing gel.

RESULTS AND DISCUSSION

In cells of the temperature-sensitive yeast mutant *rna 2* (ts 368⁻) ribosomal protein synthesis is coordinately depressed at the restrictive temperature (1,2). Two-dimensional gel analysis of proteins synthesized *in vitro* on mRNA isolated from the *rna 2* mutant at both 23°C and 36°C revealed, however, that the messenger RNAs for ribosomal proteins L16 and S33 are not reduced in concentration at the restrictive temperature (see Fig. 1). Since the temperature-induced decrease of the synthesis of most ribosomal proteins in mutant cells was shown to be caused by a splicing defect (4-7), most probably no splicing is necessary for the maturation of both L16 and S33 mRNAs. Initial studies of the structure of the respective genes were performed by electron microscopic R-loop analysis. The gene coding for ribosomal protein L16 was shown to be contiguous with the mRNA (8). Concerning the gene coding for ribosomal protein S33, we previously described the isolation and preliminary characterization of a cloned HindIII-generated yeast DNA fragment, pBMCY76, carrying the gene for S33 (9). Electron microscopic analysis of

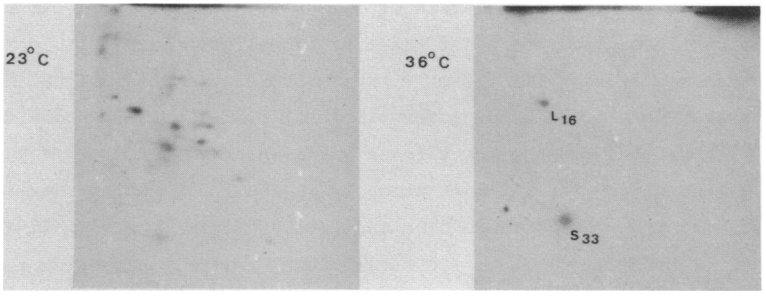

Fig.1. Translation product analysis of messenger RNA from *rna 2* cells. Poly A-containing RNA was isolated from *rna 2* (ts 368⁻) mutant cells grown at both 23° and 36°C, and used to direct the *in vitro* synthesis of proteins in a reticulocyte lysate in the presence of ^{35}S-methionine (9). After TCA-precipitation two-dimensional gel electrophoresis was carried out according to the procedure described by Bollen *et al.* (15). The pictures show the respective fluorograms.

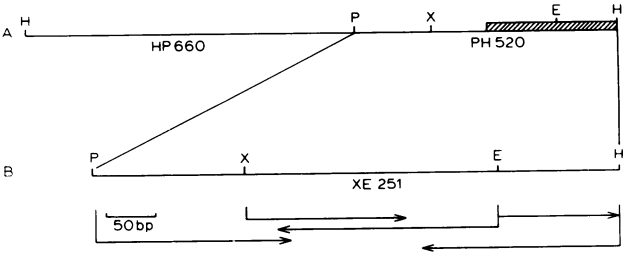

Fig. 2. Map of the insert of pBMCY76-3 and extent of DNA sequencing. The position of some restriction enzyme sites and the location of the R-loop were published in a previous paper (8). The sequence strategy is shown; the arrows give the direction and extent of nucleotide analysis. Each sequence determination was performed at least twice. H = HindIII; E = EcoRI; P = PvuII; X = XbaI.

R-loop structures formed between this fragment and yeast mRNA demonstrated that it contains two genes, one of which possesses an intron (8). Hybrid-selected *in vitro* translation (9) performed with several subclones of pBMCY76 subsequently learnt that S33 is encoded by the unsplit gene (result not shown; *vide infra*). The pertinent subcloned DNA fragment was designated pBMCY76-3 and its physical map is shown in Fig. 2. Sequence analysis was carried out according to the strategy outlined in Fig. 2.

The DNA sequence revealed three open reading frames: a short one near the left-handed HindIII-site and two others overlapping in opposite strands near the right-handed HindIII-site. In order to discriminate between these possible S33-coding sequences single-stranded recombinant DNAs were used for hybrid-selected *in vitro* translation experiments. From Fig. 3 it can be concluded that the gene coding for yeast ribosomal protein S33 is located on the right-handed side of pBMCY76-3 with its translation stop signal near the HindIII site. The sequence coding for ribosomal protein S33 is shown in Fig.4. Obviously, this gene does not contain an intervening sequence.

The mRNA coding for S33 most probably is about 400 nucleotides long as can be deduced from the Northern blot analyses shown in Fig. 5 (pY76-3; P). The longer transcript (of about 900 nucleotides) may originate from the unidentified open reading frame at the left-handed side of the cloned DNA fragment. Moreover from hybridization experiments with RNAs isolated from *rna 2* cells grown at both 23° and 36°C it is clear that at the restrictive temperature no accumulation of precursor mRNA for S33 occurs (pY76-3; R), consistent with the finding for L16 mRNA (pY89) but different from the result with respect to the S10 mRNA (pY113). Definite proof for the absence

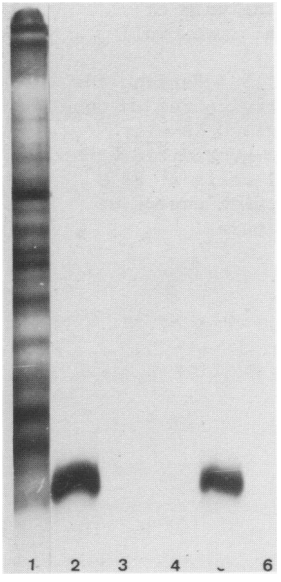

Fig. 3. Hybrid-selected in vitro translation of subcloned fragments of pBMCY76-3. The following DNA fragments were used (cf. Fig. 2): ss HP660 (5' → 3') cloned in M13mp9 (lane 3), ss HP660 (3' → 5') cloned in M13mp8 (lane 4), ss PH520 (3' → 5') cloned in M13mp9 (lane 5) and ss PH520 (5' → 3') cloned in M13mp8 (lane 6). Undigested pBMCY76-3 DNA was used as a control (lane 2). After protein synthesis in a cell-free system (9) immunoprecipitation was performed using a mixture of antisera directed against yeast ribosomal protein (9). The immunoprecipitated proteins were analyzed on 10% SDS-containing polyacrylamide gels. Lane 1 shows the immunoprecipitated ribosomal proteins synthesized in vitro on total poly A-containing mRNA.

```
                    -140                -120                -100
TCTAGAGTGACCACGCACTTTTTTGATAAATTTTTTTTTCTTGGTCGTTGAACTACTTGGAATAAAGAAAATGAAAT

 -80                -60                 -40                 -20
TTCAGTCTAATAGATGATTTATTGTAAATTACAGTTGTGTTCGTTTTCGATTCTTCTCAAAAGTAGAAAACCAAGCT
                                                               ↑

   -1 1                        20                          40
AGCAATC ATG GAT AAC AAA ACC CCA GTC ACT TTA GCC AAG GTC ATC AAA GTT TTA GGA
      (met) asp asn lys thr pro val thr leu ala lys val ile lys val leu gly

        60                          80                          100
AGA ACC GGT TCT CGT GGT GGT GTC ACC CAA GTC CGT GTC GAA TTC TTG GAA GAC ACT
arg thr gly ser arg gly gly val thr gln val arg val glu phe leu glu asp thr

            120                         140                         160
TCC AGA ACT ATT GTC AGA AAC GTG AAG GGC CCA GTT AGA GAA AAC GAC ATT TTG GTT
ser arg thr ile val arg asn val lys gly pro val arg glu asn asp ile leu val

        180                                 200
CTA ATG GAA TCT GAA CGT GAA GCT CGT CGT TTG CGT TAG GAAGCTT
leu met glu ser glu arg glu ala arg arg leu arg end
```

Fig. 4. DNA sequence of the gene for ribosomal protein S33 and aminoacid sequence of the S33 protein.
The DNA sequence of the XbaI plus HindIII-generated fragment (see Fig. 2) is shown. The nucleotides are numbered with italics starting from the first nucleotide of the initiation codon. The position of the 5' end of the S33 mRNA is indicated by an arrow.

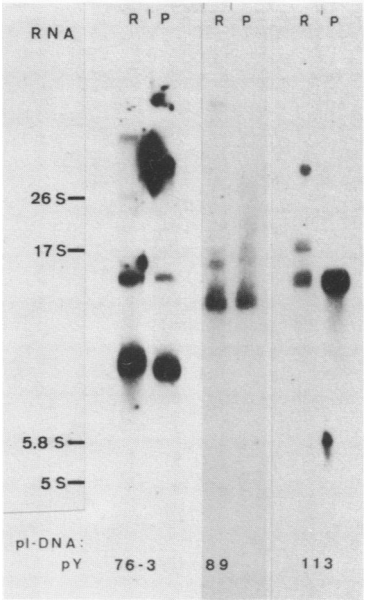

Fig. 5. Northern blot analysis of transcripts from several ribosomal protein genes.
Recombinant DNAs pBMCY113 (carrying the S10 gene), pBMCY89 (carrying the L16 gene) and pBMCY76-3 were labelled in vitro by nick-translation and hybridized with RNA isolated from rna 2 cells at 23°C (P, permissive temperature) and 36°C (R, restrictive temperature).

of an intervening sequence within the gene for ribosomal protein S33 was obtained by S1 nuclease analysis. Figure 6 demonstrates that the S1-nuclease resistant hybrids are contiguous with the coding sequence at the 5'-part of the gene (from the EcoRI site to the left). Also on the adjacent EcoRI plus HindIII-generated fragment the coding sequence turned out to be continuous (result not shown). About 108 nucleotides of the used DNA probe (XbaI-EcoRI) were protected by the S33 mRNA, which allows the mapping of the 5'-end of

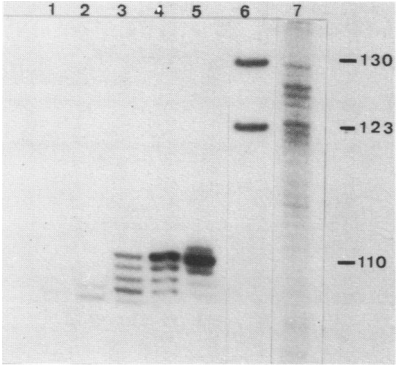

Fig. 6. S1-nuclease analysis of the gene coding for ribosomal protein S33.
The single-stranded XbaI plus EcoRI-generated fragment (see Fig. 2) was hybridized with yeast poly A-containing RNA. Digestion was performed with S1-nuclease amounts of 208, 104, 52, 26 and 13U respectively (lanes 1-5). After S1-nuclease treatment the [^{32}P] DNA-RNA hybrids were electrophoresed on 8% polyacrylamide/urea gels. A digest of the end-labelled input DNA fragment obtained with TaqI was used as size-marker (lane 6). Lane 7 represents a randomly terminated sequence ladder.

TABLE I. Codon usage for the genes coding for ribosomal proteins S33 (this paper) and S10 (7) and several other (highly expressed) genes (Y; Ref. 30) from yeast.

		S33	S10	Y			S33	S10	Y			S33	S10	Y			S33	S10	Y
Phe	UUU	-	-	5	Ser	UCU	2	8	79	Tyr	UAU	-	-	2	Cys	UGU	-	1	20
Phe	UUC	1	9	63	Ser	UCC	1	2	73	Tyr	UAC	-	5	73	Cys	UGC	-	-	2
Leu	UUA	2	-	14	Ser	UCA	-	-	1										
Leu	UUG	3	19	148	Ser	UCG	-	-	1						Trp	UGG	-	-	26
Leu	CUU	-	-	-	Pro	CCU	-	-	6	His	CAU	-	-	3	Arg	CGU	6	3	2
Leu	CUC	-	-	1	Pro	CCC	-	-	1	His	CAC	-	2	56	Arg	CGC	-	-	-
Leu	CUA	1	1	6	Pro	CCA	2	8	71	Gln	CAA	1	12	44	Arg	CGA	-	-	-
Leu	CUG	-	-	-	Pro	CCG	-	-	-	Gln	CAG	-	1	-	Arg	CGG	-	-	-
Ile	AUU	2	5	58	Thr	ACU	3	6	49	Asn	AAU	-	-	3	Ser	AGU	-	-	1
Ile	AUC	1	8	66	Thr	ACC	3	4	66	Asn	AAC	3	7	86	Ser	AGC	-	-	1
Ile	AUA	-	-	1	Thr	ACA	-	-	-	Lys	AAA	2	1	21	Arg	AGA	4	24	63
Met	AUG	2	2	39	Thr	ACG	-	-	-	Lys	AAG	2	27	151	Arg	AGG	-	-	1
Val	GUU	3	8	112	Ala	GCU	1	14	182	Asp	GAU	1	8	37	Gly	GGU	3	15	164
Val	GUC	6	10	93	Ala	GCC	1	4	62	Asp	GAC	2	4	100	Gly	GGC	1	1	11
Val	GUA	-	-	3	Ala	GCA	-	-	3	Glu	GAA	6	17	67	Gly	GGA	1	-	1
Val	GUG	1	-	1	Ala	GCG	-	-	4	Glu	GAG	-	-	3	Gly	GGG	-	1	1

TABLE II. Aminoacid composition of yeast ribosomal protein S33.

aminoacid	number of residues	
	found[a]	reported[b]
Asp	6[c]	6,2[c]
Thr	6	5,2
Ser	3	3,4
Glu	7[c]	7,1[c]
Pro	2	2,0
Gly	5	5,0
Ala	2	2,2
Val	10	8,3
Met	2[d]	1,0[d]
Ile	3	2,7
Leu	6	6,0
Tyr	0	0,1
Phe	1	1,1
His	0	0,1
Lys	4	4,0
Arg	10	9,8

a) Number deduced from the nucleotide sequence reported in this paper.
b) Number determined for yeast protein YP70 (= S33; Ref. 15) calculated from an aminoacid analysis (33) assuming that the protein has a molecular mass of 8 kilodalton.
c) Sum of acid and amide forms.
d) The mature S33 ribosomal protein apparently does not contain the first methionine.

this mRNA. Obviously the leader sequence of the S33 mRNA extends to nucleotides 11-15 upstream from the initiation codon (depending upon the amount of S1 nuclease used). Both the 5'-terminal cap-structure of the mRNA and the A-rich region at the 3' end of the probe may hinder an accurate determination (cf. Ref. 22). The 5'-end of the S33 mRNA maps at or near the tetranucleotide sequence CAAG which is in agreement with the consensus sequence PyAAPu proposed for yeast polymerase B transcriptional start sites (23,24).

In the DNA region upstream from the 5'-end of the gene no perfect TATA or CAAT boxes (see Ref. 25) can be observed, consistent with findings concerning other yeast ribosomal protein genes that have been sequenced so far (7,26). Putative TATA sequences are present at nucleotides -60 (TAAATTA) and -101 (AATAAA). Moreover a pyrimidine-rich block (-27 to -45) precedes the transcription start site at a distance of 13 bp. It has been proposed that the occurrence of a TC-rich block 10-20 bp downstream from the CAAG sequence is correlated with a high level of expression (23,24). The pyrimidine block may also function in mediating the level of transcription (27).

The leader region of the S33 mRNA is able to anneal weakly with the 3'-end of yeast 17S rRNA (28). Apart from that no obvious regulatory structures can be observed which may play a role in translational control of the S33 synthesis. Otherwise, from Fig. 1 it seems that yeast rna 2 cells do not have a mechanism to compensate for the overproduction of ribosomal proteins encoded by unsplit genes.

The preferred codon usage within the gene for ribosomal protein S33 is consistent with the selection pattern proposed by Bennetzen and Hall (29) and allows a rather efficient translation of this mRNA (see Table I). A striking deviation from the general codon usage of highly expressed genes is observed with respect to the arginine-codons: in 6 out of 10 cases CGU is used, 4 of which are present in the very 3'-part of the coding sequence.

The S33 gene encodes a protein of 8.9 kD built up of 67 aminoacids (see Fig. 3). This molecular weight is rather low as compared with the value deduced from migration of S33 into SDS-containing gels (viz. 11 kD; Ref. 31, 32). This discrepancy may be due to an aberrant electrophoretic behaviour of S33 or to the lack of suitable marker proteins in the pertinent region of the gel (cf. also Ref. 33). The aminoacid composition of the S33 protein as deduced from the nucleotide sequence is in good agreement with data reported by Higo and Otaka (33) for the small subunit protein YP70 (= YS27) (see Table II).

ACKNOWLEDGEMENTS

The authors are indebted to Mrs. L.H.P. Caro and Mr. P. Kraakman for skilful technical assistance during part of the work, to Drs C.M.T. Molenaar and L.P. Woudt for stimulating discussions and to Mrs. P.G. Brink for preparing 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 the Advancement of Pure Research (Z.W.O.).

*Present address: Gaubius Instituut, Gezondheidsorganisatie TNO, Leiden, The Netherlands

REFERENCES

1. Gorenstein, C.G. and Warner, J.R. (1976) Proc. Natl. Acad. Sci. USA 73, 1547-1551
2. Warner, J.R. and Gorenstein, C.G. (1977) Cell 11, 201-212
3. Kim, C.H. and Warner, J.R. (1983) Mol. Cell Biol. 3, 457-465

4. Rosbash, M., Harris, P.K.W., Woolford, J.L. and Teem, J.L. (1981) Cell 24, 679-686
5. Bromley, S., Hereford, L. and Rosbash, M. (1982) Mol. Cell Biol. 2, 1205-1211
6. Fried, H.M., Pearson, N.J., Kim, C.H. and Warner, J.R. (1981) J. Biol. Chem. 256, 10176-10183
7. Leer, R.J., Van Raamsdonk-Duin, M.M.C., Molenaar, C.M.T., Cohen, L.H., Mager, W.H. and Planta, R.J. (1982) Nucl. Acids Res. 10, 5869-5878
8. Bollen, G.H.P.M., Molenaar, C.M.T., Cohen, L.H., Van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.J. (1982) Gene 18, 29-37
9. Bollen, G.H.P.M., Cohen, L.H., Mager, W.H., Klaassen, A.W. and Planta, R.J. (1981) Gene 14, 279-287
10. Clewell, D.B. and Helinski, D.R. (1969) Proc. Natl. Acad. Sci. USA 62, 1159-1166
11. Bollen, G.H.P.M., Mager, W.H., Jenneskens, L.W. and Planta, R.J. (1980) Eur. J. Biochem. 105, 75-80
12. Kafatos, F.C., Jones, C.W. and Efstratiadis, A. (1979) Nucl. Acids Res. 7, 1541-1552
13. Vlak, J.M., Smith, G.E. and Summers, M.D. (1981) J. Virol. 40, 762-771
14. Smith, G.E., Vlak, J.M. and Summers, M.D. (1982) J. Virol. 44, 199-208
15. Bollen, G.H.P.M., Mager, W.H. and Planta, R.J. (1981) Mol. Biol. Rep. 8, 37-44
16. Stark, G.R. and Williams, J.G. (1979) Nucl. Acids Res. 6, 195-203
17. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251
18. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467
19. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143, 161-178
20. Messing, J. (1982) in Genetic engineering/Principles and Methods (Setlow, J.K. and Hollaender, A., eds) Plenum Press, New York, London, Vol. 4, pp. 19-34
21. Berk, A.J. and Sharp, P.A. (1978) Proc. Natl. Acad. Sci. USA 75, 1274-1278
22. Weaver, R.F. and Weissmann, C. (1979) Nucl. Acids Res. 7, 1175-1193
23. Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J., Kingsman, S.M., Perkins, R.E., Conroy, S.C., Dunbar, B. and Fothergrill, L.A. (1982) Nucl. Acids Res. 10, 2625-2637
24. Burke, R.L., Tekamp-Olson, P. and Najarian, R. (1983) J. Biol. Chem. 258, 2193-2201
25. Breatnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-384
26. Käufer, N.F., Fried, H.M., Schwindinger, W.T., Jasin, M. and Warner, J.R. (1983) Nucl. Acids Res. 11, 3123-3135
27. Montgomery, D.L., Leung, D.W., Smith, M., Shalit, P., Faye, G. and Hall, B.D. (1980) Proc. Natl. Acad. Sci. USA 77, 541-545
28. Azad, A.A. and Deacon, N.J. (1983) Nucl. Acids Res. 8, 4365-4376
29. Bennetzen, J.L. and Hall, B.D. (1982) J. Biol. Chem. 257, 3026-3031
30. Ammerer, G., Hitzeman, R., Hagie, F., Barta, A. and Hall, B.D. (1981) in Recombinant DNA (Walton, A.G., ed.) Elsevier Scientific Publishing Co., Amsterdam, pp. 185-197
31. Otaka, E. and Kobata, K. (1978) Mol. Gen. Genet. 165, 259-268
32. Kruiswijk, T., Planta, R.J. and Mager, W.H. (1978) Eur. J. Biochem. 83, 245-252
33. Higo, K. and Otaka, E. (1979) Biochemistry 18, 4191-4196