Volume 11 Number 16 1983 Nucleic Acids Research

A comparison of the promoter strengths of two eukaryotic genes *in vitro* reveals a region of DNA that can influence the rate of transcription in cis over long distances

Stephen P.Gregory and Peter H.W.Butterworth

Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, UK

Received 15 July 1983; Accepted 1 August 1983

ABSTRACT

We have compared the strength of a trout protamine gene promoter with that of the mouse β^{major}-globin gene by analysing the relative levels of run-off transcripts produced in a single mammalian in vitro transcription reaction. When the promoters are introduced on separate recombinant plasmids, the protamine transcripts are synthesised with much greater efficiency than those originating from the globin cap site. This enhanced transcription of the protamine gene is again observed when the promoters are applied as separate DNA fragments derived from the same recombinant plasmid. However, when the promoters are linked on a DNA fragment that includes 7 kb of DNA separating the initiation sites, then there is a marked reduction in the protamine signal relative to the globin. Deletion of a region of this fragment that contains the sequences flanking the globin gene at positions -335 to -1400 restores the enhanced protamine gene expression to the levels observed when the promoters are carried on separate DNA fragments.

INTRODUCTION

In recent years, it has become apparent that cloned genes vary in the efficiency with which they are expressed in eukaryotic transcription systems (1-5). The requirement of the cell for a particular gene product is likely to influence the characteristics of the promoter which ultimately determine the efficiency ('strength') of that gene's transcription. Yet our understanding of what constitutes a strong or weak promoter sequence within a given expression system is still very incomplete.

We have recently cloned a rainbow trout protamine gene (TPG-3) and found that this gene is transcribed faithfully at high levels in a whole cell extract derived from Hela cells (6). In this particular in vitro transcription system, the major functional element of the RNA polymerase II promoter is thought to be the Golberg/Hogness or ATA-box located approximately 30 bp upstream from the transcriptional start site (7,8). Sequences further upstream from this region have been shown to have an affect on the rate of transcription in the Hela extract (6,9,10) but the influence of upstream

© IRL Press Limited, Oxford, England.

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

sequences is often more evident in expression systems utilising intact cells (11-14).

We have observed that the ATA-box region of TPG-3 shares an 11 bp homology with the corresponding regions of the Adenovirus-2 (Ad-2) major late and the chicken conalbumin promoters (6), both of which are also known to be strongly transcribed in vitro (1). It would appear then that this sequence feature is particularly effective in promoting transcription in the in vitro transcription system. To assess how far this statement is true, we have chosen to compare the TPG-3 promoter with that of the mouse β^{major}-globin gene since the latter has an ATA-box sequence that differs considerably (four out of eleven base pairs) from the observed promoter homology (see Fig. 1B). This gene is also known to be transcribed faithfully in the Hela whole cell extract system (15), though how efficiently has still to be quantified.

When applied to the same transcription mix in equimolar quantities, we have found that the TPG-3 promoter is indeed more efficient than the globin promoter. Surprisingly, however, this enhanced transcription of the protamine gene is reduced when TPG-3 is carried on the same DNA fragment as the 5'-end of the globin gene which includes 1400 bp of globin 5'-flanking sequences. Removal of the distal 1000 bp of these flanking sequences restores protamine gene expression to its elevated levels.

MATERIALS AND METHODS

Recombinant plasmids

pTPG-1 : a 2500 bp fragment of trout genomic DNA cloned into the Eco R1-Bam H1 sites of pAT 153; the insert contains the entire TPG-3 gene, 2000 bp of 5'-flanking sequence and 275 bp of 3'-flanking sequence. This construct has previously been described as pTPG-3 Δ -2000 (6).

pSPG-1 : a subclone of pMβg (16) that contains 1070 bp of the 5'-end of the mouse β^{major}-globin gene plus1385 bp of flanking sequence (-1385 to +1070) cloned into the Eco R1-Pst 1 sites of pBR 322.

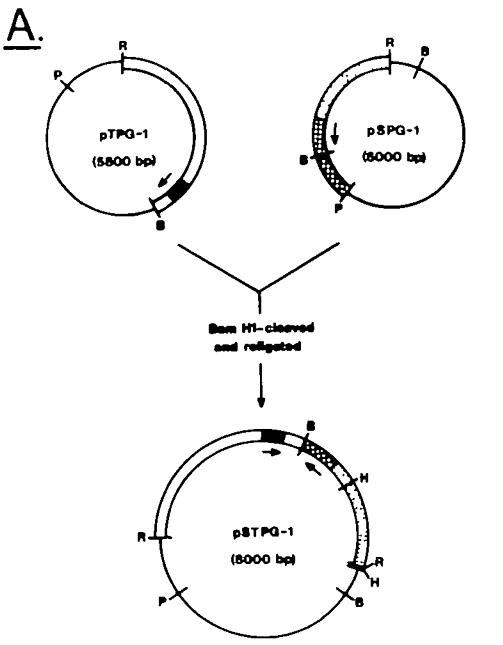
pSTPG-1 : a hybrid derived from the above two plasmids as described in the legend to Fig. 1A. The entire pAT 153 vector sequence is present together with the same eukaryotic sequences as the parental plasmids with the exception of those globin sequences from +465 to +1070.

pSTPG-2 : a derivative of pSTPG-1 in which the Bam H1 site in the vector tetracycline gene has been destroyed by filling-in and religation.

pSTPG-3 : a subclone of pSTPG-2 which is lacking the minor Hind III fragment (25 bp of plasmid sequence plus 1050 bp of sequence flanking the globin gene, see Figs 1A and 3).

pTPG-3A24 ΔTc-Bam : a derivative of pTPG-3A24 (6) in which the Bam H1-site in the tetracycline resistance gene of the vector has been destroyed by

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024



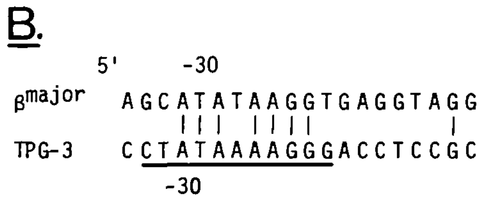


Fig. 1. A. The scheme employed for the construction of the hybrid recombinant plasmid pSTPG-1. The large Bam H1-fragment of pSPG-1 was inserted into the Bam H1-site of pTPG-1. The colonies so derived were screened for resistance to ampicillin and tetracycline. Only those colonies carrying the pSTPG-1 plasmid were viable. The location of the relevant restriction enzyme cleavage sites are indicated as follows: B = Bam H1; H = Hind III; P = Pst 1; R = Eco R1. ——— , vector sequences; ▬ , the TPG-3 trout protamine gene; ▭, trout genomic sequences flanking the TPG-3 gene; ▩, the 5'-region of the mouse β^{major}-globin gene; ⬚, mouse genomic sequences flanking the 5'-end of the β^{major}-globin gene.

B. Comparison of the ATA-box regions of the TPG-3 trout protamine and mouse β^{major} globin genes. The sequences have been aligned to produce the maximum homology in this region (vertical lines). The 11 bp homology shared between the TPG-3 protamine, chicken conalbumin and Ad-2 major late promoters is underscored (see ref. 6).

filling-in and religation. This plasmid contains only the TPG-3 protamine gene and its flanking sequences (2 kb either side) but is comparable in size to pSTPG-1 and pSTPG-2.

In vitro transcription

The *in vitro* transcriptions and subsequent analyses of the products were carried out as described previously (6) using the Hela whole cell extract system devised by Manley et al (17). CsCl-purified templates were applied as Bam H1-cleaved linear molecules at final concentrations ranging from 20 to 100 $\mu g.ml^{-1}$. The radioactive substrate used as tracer was α-^{32}P-GTP. To quantitate the relative intensity of radioactive bands, autoradiographic exposures were scanned in a Joyce-Loebl Chromoscan 3 spectrophotometer and the area under the peaks determined.

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

RESULTS AND DISCUSSION

In devising a stringent protocol to compare the promoter strengths of the trout protamine and mouse β^{major}-globin genes, we have sought to eliminate those parameters which might generate variability in the analysis. These occur at four levels: the template, the reaction, the processing of the transcripts and the final gel analysis.

1. The form in which the promoters are presented to the reaction is important because the response of the extract to increasing DNA concentrations is not a linear one. A threshold DNA concentration has to be reached before any specific transcription is observed (presumably to absorb non-specific binding factors); too much DNA leads to non-specific initiations being favoured over genuine initiations (17). It is critical then to have identical DNA and promoter concentrations when comparing promoter strengths. To this end, we have constructed recombinant plasmids (pTPG-1 and pSPG-1, see Fig. 1) that are of the same size but contain only one copy of the relevant promoter per plasmid. This assumes that the promoters lie in the 2000 bp flanking the TPG-3 gene (6) and the 1400 bp flanking the β^{major}-globin gene (3). Equal quantities of these plasmids applied to separate transcription reactions will thus contain equivalent promoter concentrations.

2. To minimise potential sources of variation, we have chosen to carry out the transcription of both genes in the same reaction mix and, where possible, with both genes on or derived from the same recombinant plasmid. The final DNA concentrations were kept within the optimal range for specific transcription. The rest of the components were prepared in the form of a pre-mix so that parallel incubations were as near identical in constitution as possible. Only the nature and concentration of the template was varied.

3. Some processing of the transcripts is desirable before eventually loading the gels for analysis but, providing the extraction volumes are kept relatively high and sufficient carrier tRNA is present at each precipitation, the differential losses which might be incurred, both within and between samples, are kept to a minimum.

4. The form of the product analysis is again of paramount importance since this is the tool by which one finally identifies differences in promoter strengths. The run-off analysis we have used here is the most powerful form of analysis since it looks at the concentration of product directly rather than relying on hybridisation and a secondary enzymatic reaction to quantitate the levels of transcript produced. Correct initiation at the *in vivo* cap sites has been demonstrated for both these linearised genes

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

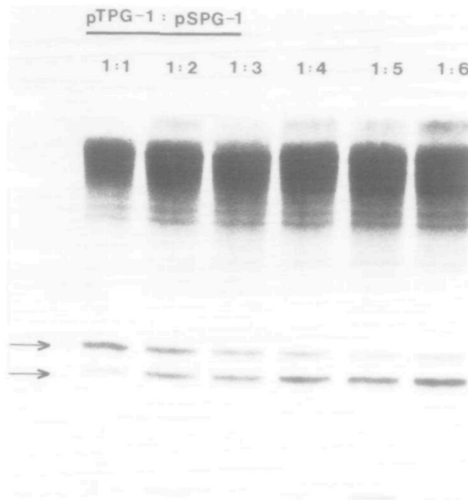


Fig. 2. Competition between pTPG-1 and pSPG-1 in the same reaction mix. The final DNA concentration was kept constant at 60 $\mu g.ml^{-1}$ whilst the ratio of pTPG-1 to pSPG-1 was progressively increased in successive reactions (see lane headings for absolute ratios). The upper arrow signifies the 505 base run-off from the protamine gene in pTPG-1; the lower arrow signifies the 469 base run-off from the globin gene in pSPG-1. It should be noted that polyadenylation does not occur in this system so there is no termination of TPG-3 transcription at the natural polyadenylation site.

in this in vitro transcription system (6,15) so the run-off assay is a valid approach to use here to study promoter strength. To minimise the chance of premature termination events, the run-off transcripts should be reasonably short but not of such a size that they comigrate with other transcription products from the plasmid (RNA polymerase III transcripts are particularly prevalent, see ref. 6). Finally, for a valid comparison to be made, the run-off products should migrate close together, thus reducing artefacts encountered during the gel separation and subsequent autoradiography and thereby ensuring a similar background for scanning determinations of band intensities. In this respect, it is fortunate that there is a single cleavage site for Bam H1 which is located 500 and 464 bp downstream from the protamine and globin cap sites respectively.

The criteria established above have provided us with a rigorous protocol for comparing the promoter strengths of the protamine and globin genes. Competition experiments were carried out in which the two templates were transcribed in a single reaction mix containing α-^{32}P-GTP (see Fig. 2). Allowing for the fact that the globin transcript is enriched in G-residues (126 compared with 101 for the TPG-3 gene run-off), quantitative analysis revealed that the protamine promoter was more than 3-fold stronger than that of the globin gene.

The analysis still depends, however, on pipetting small volumes of template-containing solutions accurately. To circumvent this problem, we

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

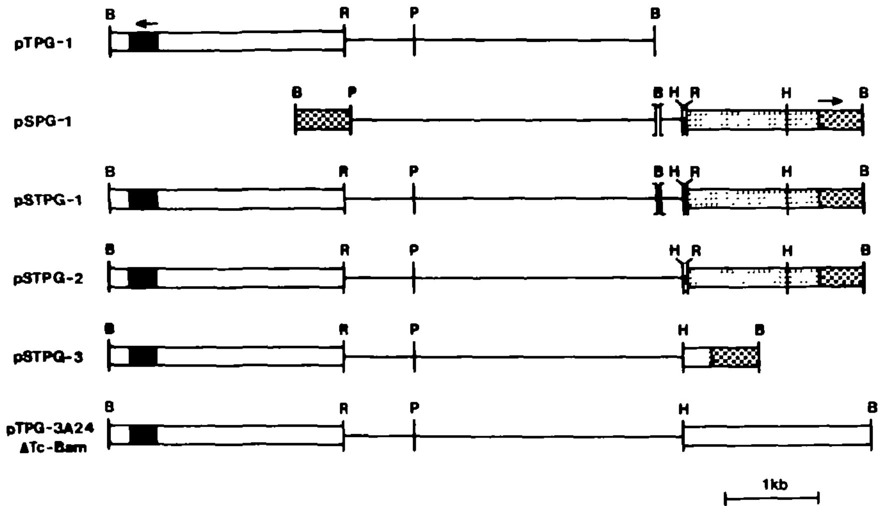


Fig. 3. Bam H1-cleavage products of templates containing protamine and/or globin genes. Common restriction sites have been aligned where possible: B = Bam H1; R = Eco R1; H = Hind III; P = Pst 1. The construction of these recombinants is described in Materials and Methods and in the legend to Fig. 1A. The key to the identification of the origin of the sequences shown is as follows: ——— , plasmid vector sequences; ■ , the TPG-3 trout protamine gene; □, trout genomic sequences flanking the TPG-3 gene; ▦ , the 5'-region of the mouse β^{major}-globin gene; ▦ , mouse genomic sequences flanking the 5'-end of the mouse β^{major}-globin gene.

have cloned these two gene sequences into the same hybrid plasmid (see Fig. Fig. 1A) so ensuring an equimolar concentration of promoter in each reaction mix. The original construct, pSTPG-1, generated two fragments on Bam H1-cleavage with one gene on each fragment (see Fig. 3). The *in vitro* transcription products from these templates confirmed the results obtained from the previous competition experiments, namely that the TPG-3 promoter was much more efficient that the globin promoter (see pSTPG 1-containing lanes of Fig. 4). Complete separation of eukaryotic and plasmid sequences by incorporation of an Eco R1 digest prior to the Bam H1-cleavage shown in Fig. 3 revealed no change in this pattern of transcription (data not shown). Thus neither the plasmid sequences nor the overall target size of the DNA template was contributing in any way to the differential transcription of the two genes. The conclusion must be that this reflects a genuine difference in promoter strengths of these two genes.

To complete the analysis, the internal Bam H1-site in the pSTPG-1 vector was eliminated generating pSTPG-2 (see Fig. 3) so that protamine and globin

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

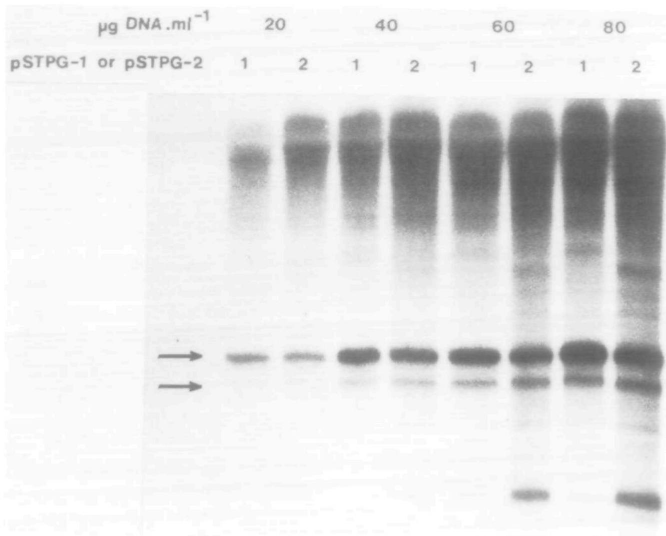


Fig. 4. The effect of varying the DNA concentration on the transcription of Bam H1-cleaved pSTPG-1 and pSTPG-2. Final DNA concentration, expressed in $\mu g.ml^{-1}$, and the nature of the template are displayed above the relevant lanes. The upper arrow shows the position of the run-off from the TPG-3 gene, the lower arrow shows the position of the run-off from the mouse β^{major} globin gene. The bands at the bottom of the gel are RNA polymerase III-transcribed products (see ref. 6).

genes are retained, one at either end of the linear molecule arising from Bam H1-cleavage. Therefore run-off transcripts will be produced from either end of the linearised template, 7 kb separating the two initiation sites. To our surprise, we observed a decrease in the relative intensity of TPG-3 transcription from this template. This reproducible phenomenon occurred to approximately the same degree over a wide range of DNA concentrations (see Fig. 4). The protamine gene product was still synthesised at higher levels than that from the globin gene but the ratio fell dramatically, principally due to a decrease in the efficiency of the TPG-3 promoter (see Table 1). Table 1 presents quantitative data concerning the relative intensities of these run-off bands. In each experiment, the products from reactions containing equivalent concentrations of pSTPG-1 or pSTPG-2 were run in adjacent lanes on polyacrylamide gels; autoradiograms were developed until the TPG-3 run-off from pSTPG-1 was of moderate intensity. The non-linear response of X-ray film to β-emissions creates variation in the apparent degree of enhanced protamine transcription (different exposures give

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

Table 1. Quantitative estimation of the relative amount of the globin and protamine gene products synthesised from either pSTPG-1 or pSTPG-2. Incubations contained a final DNA concentration of 40-80 μg.ml^{-1}. Data has been corrected to take into account the imbalance in the proportion of G-residues in the run-off transcripts of the two genes (see text).

Expt	% transcription from pSTPG-2 relative to that from pSTPG-1*		Ratio of transcription of Protamine : Globin		
	Protamine	Globin	pSTPG-1	pSTPG-2	% decrease
1	87	142	6.12	3.76	49
2	76	145	8.23	4.28	48
3	31	62	9.70	4.92	49
4	83	131	10.34	6.52	37
5	31	60	4.30	2.20	49
6	60	110	6.72	3.68	45

* expressed as a percentage of transcription of the relevant gene from pSTPG-2 relative to that derived from a parallel incubation containing pSTPG-1.

different comparative results). However, within a single exposure a valid comparison can be drawn between the relative transcription from pSTPG-1 and pSPTG-2. These results reveal a consistent fall in the ratio of TPG-3 : globin transcription of 40-50%.

The decrease in TPG-3 transcription is not simply due to the promoter being carried on a larger 8 kb fragment of DNA since data in Fig. 5 shows that pTPG-3A24 ΔTc-Bam (see Fig. 3) generated the same intensity product as smaller fragments of TPG-3 DNA. The inhibitory effect appears instead to be mediated by sequences flanking the globin gene. Removal of the small Hind-III fragment of pSTPG-2 to yield pSTPG-3 (see Fig. 3) restored the enhanced TPG-3 transcription to those levels observed in parallel pSTPG-1 (Fig. 5) or Eco R1-digested pSTPG-2 reactions (data not shown). There was no corresponding dramatic increase in globin transcription. The deleted piece of DNA contains 1000 bp of globin 5'-flanking sequence located at -335 to -1400 plus the 37 bp Eco R1 - Hind III fragment of pAT 153. The plasmid sequences are unlikely to be responsible for this reduction in protamine transcription for they are present in other TPG-3 constructs that exhibit no reduction in the efficiency of transcription (ref. 6 and unpublished data). Thus we are forced to invoke the presence of a *cis*-acting inhibitory element, probably unidirectional, that can influence the rate of transcription over large distances (>5 kb in the case of the protamine gene).

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

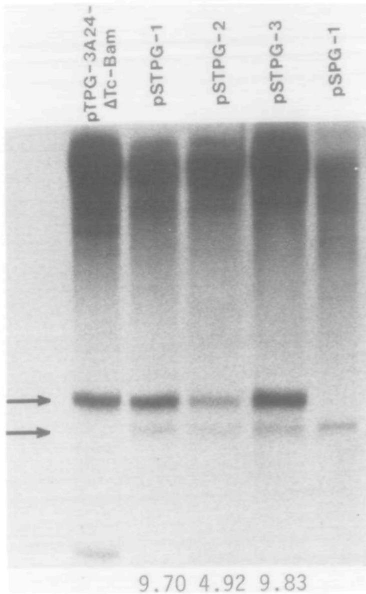


Fig. 5. Comparison of the levels of run-off production from pSTPG plasmids. A final DNA concentration of 60 $\mu g.ml^{-1}$ was used for pSTPG-1 and pSTPG-2. The concentration of the other plasmids was adjusted accordingly to achieve equimolar concentrations of promoters. The upper arrow shows the position of the TPG-3 run-off, the lower arrow shows the position of the globin run-off product.

Ratio of protamine : globin transcripts

CONCLUSIONS

The strategy outlined here was designed to provide a direct comparison of the overall promoter efficiency of two eukaryotic genes. The results show unequivocally that the trout TPG-3 promoter is considerably stronger than the mouse β^{major}-globin gene in the Hela *in vitro* whole cell extract transcription system. Whether this is solely due to the ATA-box differences or whether some other element of the promoter is involved we cannot say. We favour the latter because it has been reported that the Ad-2 major late promoter is twice as strong *in vitro* as the chicken conalbumin promoter (1) although they share a 12 bp homology in the ATA-box region (18). Similarly, in the related Hela S100 *in vitro* transcription system, the β^{major}-globin gene is transcribed at levels 3-fold higher than the mouse α_1-globin gene, even though these two genes share a 9 bp homology (GCATATAAG) in their ATA-box regions (3). We have presented evidence which suggests that TPG-3 sequences upstream from the ATA-box may play a role in determining the rate of expression of the protamine gene in the Hela extract transcription system (6).

The identification of a 'reducer' region that inhibits transcription over long distances is intriguing. Elements that stimulate transcription at long range ('enhancers') have now been identified *in vivo* (4,19-23) so one might expect to find analogous inhibitory elements, especially if

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024

enhancers function by inducing changes in chromatin structure. It is our intention to confirm that this segment truly possesses inhibitory properties by linking it to other eukaryotic genes, in both orientations and at variable distances and by testing the constructs in *in vitro* and *in vivo* expression systems.

ACKNOWLEDGEMENTS

We would like to thank Ian Brener and Ian Zachary for their help in creating pSTPG-1, Phil Leder for supplying us with pMβg and the Imperial Cancer Research Fund for allowing us to use their Joyce-Loebl Chromoscan 3. This work was supported by the Science and Engineering Research Council and the Wellcome Trust.

REFERENCES

1. Wasylyk, B., Kedinger, C., Corden, J., Brison, O. & Chambon, P. (1980) Nature 285, 367-373.
2. Proudfoot, N.J., Shander, M.H.M., Manley, J.L., Gefter, M.L. & Maniatis, T. (1980) Science 209, 1329-1336.
3. Talkington, C.A. & Leder, P. in "Promoters: Structure and Function" (ed. Rodriguez, R.L. & Chamberlin, M.J., Praeger Publishers, New York) 1982, 307-331.
4. Humphries, R.K., Ley, T., Turner, P., Moulton, A.D. & Nienhuis, A.W. (1982) Cell 30, 173-183.
5. Dynan, W.S. & Tjian, R. (1983) Cell 32, 669-680.
6. Gregory, S.P., Dillon, N.O. & BUtterworth, P.H.W. (1982) Nuc. Acids Res. 10, 7581-7592.
7. Hu, S.-L. & Manley, J.L. (1981) Proc. Natl Acad. Sci. USA 78, 820-824.
8. Grosveld, G.C., Shewmaker, C.K., Jat, P. & Flavell, R.A. (1981) Cell 25, 215-226.
9. Hen, R., Sassone-Corsi, P., Corden, J., Gaub, M.P. & Chambon, P. (1982) Proc. Natl Acad. Sci. USA 79, 7132-7136.
10. Myers, R.M., Rio, D.C., Robbins, A.K. & Tjian, R. (1981) Cell 25, 373-384.
11. Grosschedl, R. & Birnsteil, M.L. (1980) Proc. Natl Acad. Sci. USA 77, 7102-7106.
12. McKnight, S.L. & Kingsbury, R. (1982) Science 217, 316-324.
13. Grosveld, G.C., Rosenthal, A. & Flavell, R.A. (1982) Nuc. Acids Res. 10, 4951-4971.
14. Dierks, P., Van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. & Weissmann, C. (1983) Cell 32, 695-706.
15. Hagenbüchle, O. & Schibler, U. (1981) Proc. Natl Acad. Sci. USA 78, 2283-2286.
16. Konkel, D.A., Tilghman, S.M. & Leder, P. (1978) Cell 15, 1125-1132.
17. Manley, J.L., Fire, A., Cano, A., Sharp, P.A. & Gefter, M.L. (1980) Proc. Natl Acad. Sci. USA 77, 3855-3859.
18. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. & Chambon, P. (1980) Science 209, 1406-1414.
19. Banerji, J., Rusconi, S. & Schaffner, W. (1981) Cell 27, 299-308.
20. De Villiers, J. & Schaffner, W. (1981) Nuc. Acids Res. 9, 6251-6264.
21. Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M.P. & Chambon, P. (1981) Nuc. Acids Res. 9, 6047-6068.
22. Gruss, P., Dhar, R. & Khoury, G. (1981) Proc. Natl Acad. Sci. USA 78, 943-947.
23. Boss, M.A. (1983) Nature 303, 281-282.

Downloaded from https://academic.oup.com/nar/article/11/16/5317/2379205 by guest on 19 April 2024