Volume 16 Number 9 1988 Nucleic Acids Research

An ACCC-containing protein-binding sequence in the neighbourhood of the decanucleotide recognition site of the immunoglobulin gene promoter

Ralph Mocikat, Ger J.M.Pruijn[1], Peter C.van der Vliet[1] and Hans G.Zachau

Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, FRG and [1]Laboratorium voor Fysiologische Chemie, Rijksuniversiteit Utrecht, The Netherlands

Received February 9, 1988; Revised and Accepted March 29, 1988

ABSTRACT

The decanucleotide (dc) box TNATTTGCAT is an important functional element in the K light chain gene promoter and was shown to bind nuclear proteins. We now describe the binding of nuclear factor III to dc containing oligonucleotides and define a second protein binding site in the neighbourhood of dc which turned out to be recognized by a partially purified NFIII preparation but not by highly purified NFIII (abbreviations ref. 1). This site comprises an ACCC-containing sequence element. It was further characterized by the hydroxyl radical footprinting method which is applicable to 40 bp long oligonucleotides. The significance of the ACCC-containing sequence and of the previously described 15mer sequence pd is discussed.

INTRODUCTION

Several promoter and enhancer regions are known which were shown to bind transcription stimulating proteins. Such transcription factors are e.g. the Sp1 protein and the Drosophila heat shock transcription factor (reviewed in refs. 2-4). An important regulatory element in the upstream region of K light chain genes is the highly conserved decanucleotide (dc) TNATTTGCAT (5). Parslow et al. defined this element as an octanucleotide ATTTGCAT (6). dc sequences were found in several other gene regions too (summarized in ref. 7). In the K light chain promoter region there is another conserved sequence element which is called pentekaidecanucleotide (pd) the function of which is unknown (5,8).

In the past two years several groups have described proteins which bind to the dc sequence in various gene regions and which are believed to act as transcription factors (9-16). One protein recognizing dc related sequences in several gene promoters is nuclear factor III (17) which had previously been shown to sti-

mulate replication of adenovirus DNA in vitro (18). This suggests that replication and transcription systems might be related with regard to the involved regulatory proteins. In this report we describe the binding of partially and highly purified NFIII to oligonucleotides containing the dc box as it was originally described by Falkner et al. (5). In a previous paper we had described an ACCC-containing sequence (termed "C-rich sequence") which seemed to be recognized by a dc binding protein (11). Since most of those experiments had been done with crude nuclear extracts it was not sure whether the two elements were recognized by the same or by different protein(s). For this reason we have repeated the experiments with partially and highly purified NFIII and included footprinting experiments.

MATERIALS AND METHODS

Preparation and purification of DNA probes

Synthetic oligonucleotides were prepared on a Biosearch 3810 DNA synthesizer. The purification procedure was exactly as described in ref. 11. The adenovirus 4 probe was a 85 bp EcoRI-XbaI fragment which contained the terminal 78 bp of the Ad4 inverted terminal repeat (17). The double stranded DNAs were 5'-endlabeled with $-^{32}$P-ATP by polynucleotide kinase (Boehringer Mannheim). Oligonucleotides used in footprint experiments were labeled in single stranded form prior to reassociation with the unlabeled complementary strand.

Cell culture, preparation of protein extracts and purification of NFIII

Raji cells were propagated as described in ref. 11. The nuclear protein extracts were prepared following the procedure of Dignam et al. (19) with the modification of ref. 11.

Nuclear extract from Raji cells was chromatographed on a Sephadex G 200 column (11) and further purified on a Heparin-Sepharose column according to ref. 20 except that buffer A contained 50 mM KCl. 60 fractions were collected and tested for the presence of binding activity.

NFIII was purified from a nuclear extract from HeLa cells. Extraction of the nuclei and the first purification steps (DEAE-cellulose and phoshocellulose column chromatography) were per-

formed as described (18). NFIII containing fractions were further purified by pKB67-88 DNA-cellulose chromatography as described by Rosenfeld and Kelly (21). After washing this column with 0.2 M NaCl in buffer B (18) NFIII was eluted with 0.5 M NaCl in buffer B. Finally, these fractions were chromatographed on a denatured calf thymus DNA-cellulose column. This column was washed with 0.1 M NaCl followed by step elutions at 0.35 M NaCl and 1 M NaCl. NFIII eluted with the 0.35 M NaCl step.

Gel retardation assay

The gel electrophoretic binding test and competition experiments were performed as outlined previously (11,17).

Footprinting experiments

For the DNAase I protection experiments a standard binding reaction was performed for 15 min at room temperature. The reaction mixture contained 20-40 fmoles of endlabeled oligonucleotide, 200 ng poly (dI-dC).poly (dI-dC) and 2-4 µl NFIII (phosphocellulose fraction) or nuclear extract in standard binding buffer which contained 50 mM NaCl. The digestion was performed in the presence of 5 mM $MgCl_2$ and 5 units DNAase I (Pharmacia) and was stopped after 2 min by adding 1 µl 0.5 M EDTA. The mixture was then separated on a retention gel. The complex band was excised, and the DNA was eluted in 10 mM Tris, pH 8.0, 1 mM EDTA. After extraction with phenol/Tris, pH 7.9, and chloroform the DNA was concentrated by centrifuging in a Centricon 10 microconcentrator tube (Amicon GmbH, Witten an der Ruhr) and precipitated in the presence of 10 µg carrier-tRNA, 50 mM ammonium carbonate and 3 volumes of ethanol. The fragments were analyzed by electrophoresis on 10-20% polyacrylamide gels containing 6 M urea.

The hydroxyl radical footprinting procedure was a modification of the method described in ref. 22. The binding reaction was performed as described above with the exception that the binding buffer did not contain glycerol. The breakdown reaction was initiated by mixing 1.5 µl of a freshly prepared solution of 2 mM $(NH_4)_2Fe(SO_4)_2$, 4 mM EDTA, 1.5 µl of 6% H_2O_2 and 1.5 µl of 200 mM sodium ascorbate on the inner wall of the tube and centrifuging into the binding mixture. After 5 min the reaction was stopped by adding the retention gel loading buffer which contained 20% glycerol. Further handling of the probes was as outlined above.

Transfection experiments

Stable transformants of the cell line X63Ag8.653 (ref. 23) were obtained by a modification of the calcium phosphate coprecipitation technique (8) and selection with mycophenolic acid. RNA was isolated from the clones according to ref. 24 and analyzed by Northern blotting.

RESULTS AND DISCUSSION

Binding properties of NFIII to dc containing oligonucleotides

Nuclear factor III binds to a sequence in the adenovirus inverted terminal repeat but also to a conserved consensus sequence of the H2B, the V_K and the V_H promoters and of the IgH, U2 and SV40 enhancers (17). This consensus sequence forms part of the dc box, as it was described by Falkner et al. (5). We studied the binding of partially and highly purified NFIII to oligonucleotides which were synthesized according to the sequence upstream of the V_K gene T1 (ref. 25); this sequence contains a canonical dc box.

The oligonucleotides used are compiled in Fig. 1. The identity of NFIII with the dc binding activity in crude nuclear extracts was confirmed by gel retention assays with ^{32}P-labeled oligonucleotide A. Nuclear extract of Raji cells and NFIII (phosphocellulose fraction as well as a highly purified fraction) yielded complexes with the same electrophoretic mobilities (Fig. 2a,b). In a competition assay with labeled Ad 4-fragment (17), unlabeled oligonucleotide A competed for binding, but less efficiently than did the Ad 4-fragment (Fig. 3a). This difference is perhaps due to nucleotides 3' adjacent to dc (GCT in A, ATT in Ad 4) and/or to the fact that the Ad4 probe and the oligonucleotide A are different in length.

A dc sequence with changes of the first and second position (oligonucleotide A') is recognized by NFIII (phosphocellulose fraction) with a slightly lower affinity than the original dc box (Fig 3b). Affinity for A' was 74% in comparison to affinity for A. This finding supports the contention that the two positions 5' of the consensus sequence defined by Pruijn et al. (17) are important for the binding affinity. In agreement with this, the 70Z 3 gene promoter fragment which also contains a C in the first position is recognized by NFIII less efficiently (17).

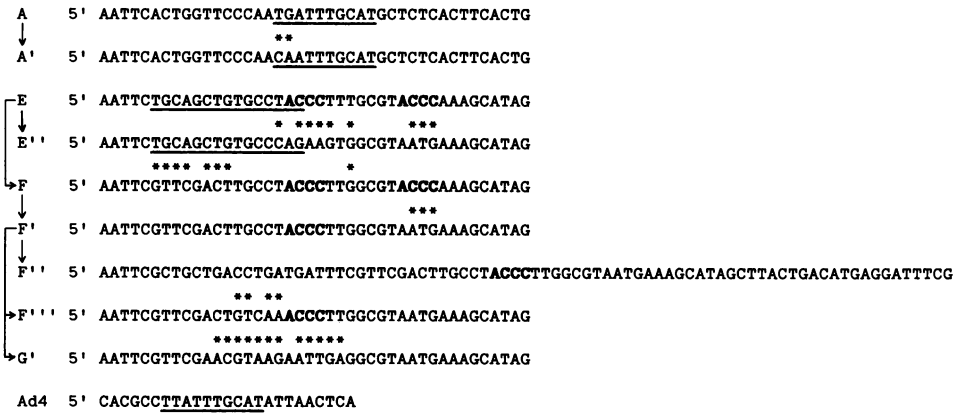

Figure 1. Compilation of the synthetic oligonucleotides used in this study. Oligonucleotides derived from each other are marked by arrows. Asterisks indicate mutated nucleotides. The dc and pd sequences (5,8) are underlined, the ACCC motif (see text) is written in bold letters. The sequences which extend oligonucleotide F' to F'' are derived from the upstream region of the V_K gene K2 (31).

Another protein binding site is located near dc

In a previous paper we described an ACCC-containing sequence occurring in the neighbourhood of dc which binds a protein independently of the dc box (11). Since these experiments were carried out with crude nuclear extract or fractions from a gel filtration column, we could not draw any conclusions concerning the identity of the protein(s) binding to this C-rich sequence. We now found that also in experiments with the phosphocellulose fraction of NFIII the dc sequence is competed for by the C-rich sequence (Fig. 3c). In order to define the boundaries of protein binding to the C-rich sequence we performed footprinting studies with Raji nuclear extract or with the phosphocellulose fraction of NFIII. As it was not possible to obtain a good DNAase I footprint with a 42 bp oligonucleotide, we used the oligonucleotide F'' which was derived from F by mutating the ACCC motif in the right half and by adding about 20 bp at each side. The DNAase I footprint shows a protected region of 24 nucleotides and 21-24 nucleotides on the complementary strand (Figs. 4a, 5). The protected region encompasses the pd box (which is partly mutated in oligonucleotides F, F' and F'') and the C-rich sequence.

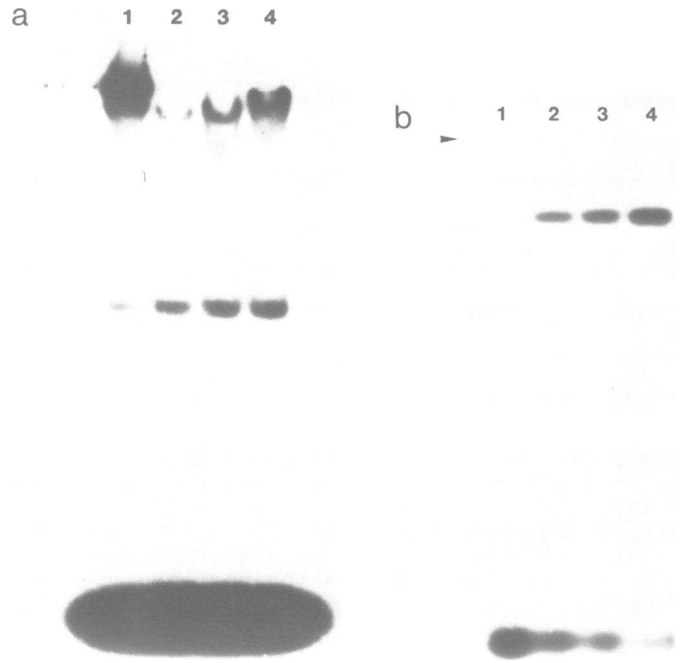

Figure 2. Gel retardation assays with ^{32}P-labeled oligonucleotide A. (a) Lane 1: 5 µg of crude nuclear extract of Raji cells; lanes 2-4: 1, 2 and 4 µl of NFIII phosphocellulose fraction. The gel retardation assay was done as described in ref. 11. (b) Lane 1: no protein; lane 2: NFIII phosphocellulose fraction; lanes 3, 4: highly purified NFIII from different isolations. This experiment was performed as described in ref. 17. The arrowhead designates the migration start point.

In order to define the binding sequence more precisely we used a modification of the hydroxyl radical footprinting method described in ref. 22. This technique is based on the degradation of the desoxyribose moiety by the hydroxyl radical, a reaction which can be prevented by proteins bound to the DNA. The advantage of the method is that the modifying agent is a small molecule producing rather short protected regions and that the cleavage is largely sequence unspecific. The hydroxyl radical footprint experiments were done with oligonucleotide F'' as well as with the 42 bp oligonucleotide F'. Fig. 4b shows a protected region of 8-12 nucleotides. This DNA stretch contains the ACCC motif and the sequence TGCCT which forms part of the 3' end of

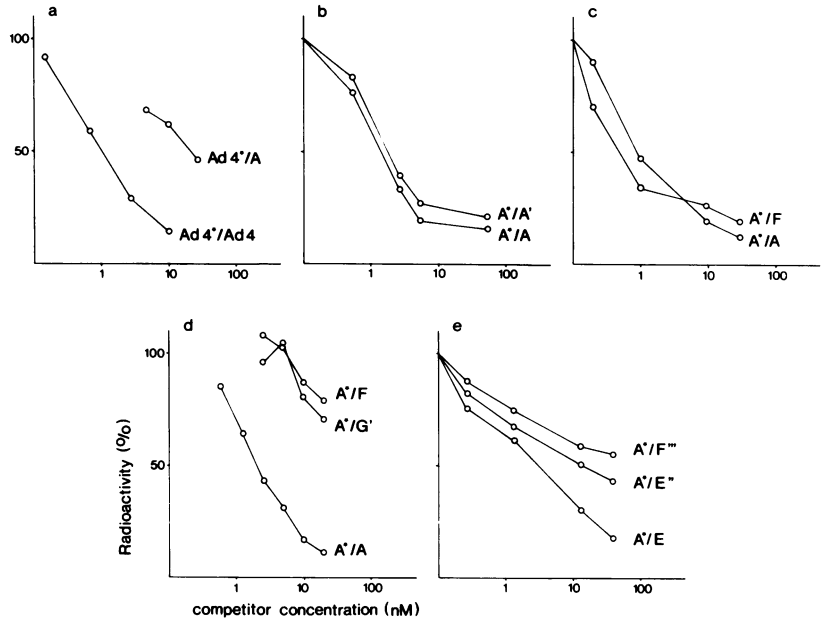

Figure 3. Quantitation of competition assays in gel retardation electrophoresis. The complex bands were excised from the gel and measured by liquid scintillation counting. Radioactivity was expressed in per cent of the complex without competitor and plotted against the competitor concentration. The radioactive oligonucleotides (marked by asterisks) and the competing unlabeled oligonucleotides are designated by the capital letters. The experiments in panel a and d are performed with highly purified NFIII, those in panel b and c with NFIII phosphocellulose fraction and those in panel e with a fraction purified on a Sephadex G200 and on a heparin-sepharose column.

the pd sequence (Fig. 5). Footprint reactions with oligonucleotide F yielded the analogous result (not shown). A technical aspect of these experiments can probably be generalized: chemical agents as the hydroxyl radical allow footprinting experiments with smaller oligonucleotides than does DNAase I.

Different proteins may be involved in binding to dc and the ACCC-containing sequence

As mentioned above, binding activity for the C-rich sequence was not only found in crude extract or in gel filtration fractions containing the dc binding factor(s) but also in the NFIII phosphocellulose fraction (Fig. 3c) and in an extract partially puri-

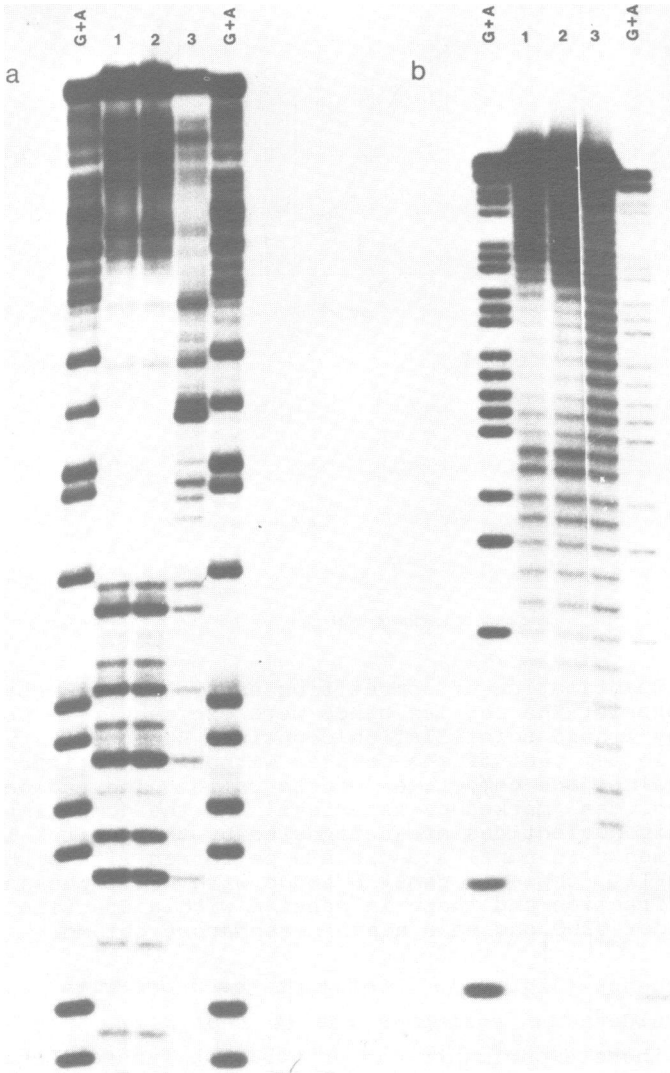

Figure 4. Footprint experiments by DNAase I digestion (a) and by the hydroxyl radical method (b). The procedures are described in Materials and Methods. Lane 1: footprint with NFIII phosphocellulose fraction; lane 2: footprint with crude nuclear extract from Raji cells; lane 3: no protein; G+A: chemical cleavage ladder (32). The weaker bands in the lower part of the gel in Fig. 4b do not represent a second footprint region but are due to lower recoveries of the short cleavage fragments of the 42 bp oligonucleotide after the concentration and precipitation steps. In panel b, the footprint on the complementary strand of oligonucleotide F' is shown.

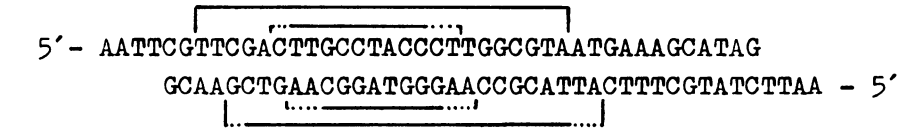

Figure 5. Protected region on oligonucleotide F' and F'', respectively. Large bracket: DNAase I footprint: small bracket: hydroxyl radical footprint. The figure shows only the part of oligonucleotide F'' that corresponds to F'.

fied by gel filtration and a heparin-sepharose column (not shown). In order to test the possibility that NFIII might recognize two different sequences we performed competition assays with highly purified NFIII. In these experiments oligonucleotide F did not compete for complex formation with oligonucleotide A (Fig. 3d). There was no difference between oligonucleotide F and the control oligonucleotide G'. This result suggests that the C-rich region is not recognized by NFIII and may be recognized by another protein which is present in partially purified NFIII preparations. An alternative explanation would be the existence of a factor in the partially purified preparations that does not bind to DNA but recognizes NFIII thereby influencing its binding behaviour. This factor might be eliminated in the last steps of the purification procedure. The factor alone would not be detected in the retention assay.

The role of the pd region and the ACCC-containing sequence in protein binding

The hydroxyl radical footprint experiments showed a protected region which contains the C-rich sequence and a part of pd. In order to evaluate the importance of the intact 3' part of pd for protein binding we synthesized oligonucleotides with mutated pd (F''') and mutated ACCC (E'') and performed competition assays with nuclear extract partially purified on a gel filtration and on a heparin-sepharose column. Fig. 3e shows that binding affinity for both mutated oligonucleotides is markedly decreased and that affinity for E'' is slightly higher than for F'''. Obviously both, the 5' and the 3' part of the footprint sequence are used for protein binding under the assay conditions. Only oligonucleotide G' which has alterations in both regions does not bind at all.

The ACCC-containing sequence alone is not sufficient to promote transcription.

The finding of the C-rich protein binding site in the K gene promoter raises the question whether the sequence acts as a promoter in vivo. Therefore we prepared a construct by excising the 780 bp EcoRI-SphI fragment including the dc element from the recombinant p870-gpt which was described in ref. 5. The deleted K promoter was substituted by oligonucleotide E which contained an intact pd region, the C-rich sequence, but no dc box. The resulting construct pE-gpt was transfected into X63Ag8.653 cells (23) by the calcium phosphate coprecipitation technique modified as described in ref. 8. After selection with mycophenolic acid and subcloning RNA was isolated (24), glyoxylated (26), electrophoresed on 1.5% agarose gels and blotted on to nitrocellulose filters which were then hybridized with a ^{32}P-labeled C_K probe. No K specific transcript was detected while in a control experiment with the intact p870-gpt a 1.2 kb K mRNA was clearly seen (data not shown).

Concluding remarks

In this study we have defined a novel protein binding site in the neighbourhood of the decamer box of the immunoglobulin K gene promoter. We found with crude nuclear extracts and partially purified NFIII preparations a protected region of 12 bp including the C-rich sequence. Interestingly, elements similar to the footprint sequence occur near the dc region of most human V_K genes of the four subgroups as well as in the upstream region of some V_H genes. A similar sequence was found in the c-myc oncogene promoter and was shown to bind a nuclear protein (27). Also in the SV40 enhancer (28), in the BPV I origin of replication (which also acts as a promoter, ref. 29) and in the ß-globin gene promoter (30) there are sequence elements which show homology to our footprint sequence. The functional significance of the described region is unclear. Certainly it is dispensable for the expression of K genes but it might facilitate binding of NFIII and other factors to the gene promoter.

ACKNOWLEDGEMENTS

We thank C. Benthake for excellent technical assistance. The work

was supported by Bundesministerium für Forschung und Technologie and Fonds der Chemischen Industrie (München).

In Utrecht, excellent technical assistance by R.T. van Miltenburg and the gift of plasmid pKB67-88 by P.J. Rosenfeld and T.J. Kelly are gratefully acknowledged. The work was supported in part by the Netherlands Foundation for Chemical Research (SON) with financial support from the Netherlands Organization for the Advancement of Pure Research (ZWO).

REFERENCES

1. Abbreviations: dc, pd, decanucleotide and pentekaidecanucleotide sequences in immunoglobulin gene promoters; NFIII, nuclear factor III; Ad4, adenovirus 4; V_K gene, variable gene segment of an immunoglobulin gene of the K type.
2. Dynan, W.S. and Tjian, R. (1985) Nature 316, 774-778.
3. Serfling, E., Jasin, M. and Schaffner, W. (1985) Trends Gen. 1, 224-230.
4. Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) Trends Biochem. 11, 20-23.
5. Falkner, F.G. and Zachau, H.G. (1984) Nature 310, 71-74.
6. Parslow, T.G., Blair, D.L., Murphy, W.J. and Granner, D.K. (1984) Proc. Natl. Acad. Sci. USA 81, 2650-2654.
7. Falkner, F.G., Mocikat, R. and Zachau, H.G. (1986) Nucl. Acids Res. 14, 8819-8827.
8. Falkner, F.G., Neumann, E. and Zachau, H.G. (1984) Hoppe-Seyler's Z. Physiol. Chem. 365, 1331-1343.
9. Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) Nature 319, 154-158.
10. Hromas, R. and van Ness, B. (1986) Nucl. Acids Res. 14, 4837-4848.
11. Mocikat, R., Falkner, F.G., Mertz, R. and Zachau, H.G. (1986) Nucl. Acids Res. 14, 8829-8844.
12. Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A. and Baltimore, D. (1986) Nature 323, 640-643.
13. Peterson, C.L., Orth, K. and Calame, K.L. (1986) Mol. Cell. Biol. 6, 4168-4178.
14. Sive, H.L. and Roeder, R.G. (1986) Proc. Natl. Acad. Sci. USA 83, 6382-6386.
15. Bohmann, D., Keller, W., Dale, T., Schöler, H.R., Tebb, G. and Mattaj, I.W. (1987) Nature 325, 268-272.
16. Rosales, R., Vigneron, M., Macchi, M., Davidson, I., Xiao, J.H. and Chambon, P. (1987) EMBO J. 6, 3015-3025.
17. Pruijn, G.J.M., van Driel, W., van Miltenburg, R.T. and van der Vliet, P.C. (1987) EMBO J. 6, 3771-3778.
18. Pruijn, G.J.M., van Driel, W. and van der Vliet, P.C. (1986) Nature 322, 656-659.
19. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucl. Acids Res. 11, 1475-1489.
20. Davison, B.L., Egly, J.M., Mulvihill, E.R. and Chambon, P. (1983) Nature 301, 680-686.
21. Rosenfeld, P.J. and Kelly, T.J. (1986) J. Biol. Chem. 261, 1398-1408.

22. Tullius, T.D. and Dombroski, B.A. (1986) Proc. Natl. Acad. Sci. USA 83, 5469-5473.
23. Kearney, J.F., Radbruch, A., Liesegang, B. and Rajewsky, K. (1979) J. Immunol. 123, 1548-1550.
24. Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
25. Altenburger, W., Steinmetz, M. and Zachau, H.G. (1980) Nature 287, 603-607.
26. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
27. Thalmeier, K., Winnacker, E.-L. and Lipp, M. (1987) Abstracts of the Annual Meeting of Gesellschaft für Genetik, May 28-30, München, p. 18.
28. Weiher, H., König, M. and Gruss, P. (1983) Science 219, 626-631.
29. Stenlund, A., Bream, G.L. and Botchan, M.R. (1987) Science 236, 1666-1671.
30. Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and Weissmann, C. (1983) Cell 32, 695-706.
31. Nishioka, Y. and Leder, P. (1980) J. Biol. Chem. 255, 3691-3694.
32. Maxam, A. and Gilbert, W. (1980) Meth. Enzym. 65, 499-525.