Volume 8 Number 7 1980 Nucleic Acids Research

A mouse immunoglobulin heavy chain deletion mutant: isolation of a cDNA clone and sequence analysis of the mRNA

W.Dunnick, T.H.Rabbitts and C.Milstein

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Received 11 February 1980

ABSTRACT

The mouse cell line IF2 secretes an immunoglobulin heavy chain lacking the C_H1 domain. We have isolated and characterised a recombinant plasmid containing cDNA copies of the IF2 mutant mRNA. The cloned sequence extends from the nucleotides coding for amino acid 96 in the variable region through 100 nucleotides of untranslated region at the 3' end. The sequence of the cDNA insert reveals no discontinuity at the variable-hinge region junction, the site of the C_H1 deletion. Experiments employing direct priming on the poly(A) tail of the IF2 heavy chain mRNA suggest that the 3' end of the cDNA clone (sequence C-C-C-T-G-C) is also the 3' end of the mRNA.

INTRODUCTION

Immunoglobulin light and heavy polypeptide chains are subdivided into regions (domains) of primary sequence and three-dimensional structural homology. Several heavy chains have been identified in which one or more domains are missing from the protein [1]. It has been shown that domains are encoded by exons [2,3], suggesting that the study of mutants involving deletion of domains may further our understanding of RNA processing. A well characterised example of such a mutant is the heavy chain secreted by the IF2 cell line. This cell line was derived as a spontaneous mutant of the mouse myeloma cell line P3. The mutant IF2 $\gamma1$ heavy chain lacks the C_H1 domain [4]. The IF2 heavy chain mRNA is shorter than the corresponding "wild type" mRNA [5], an observation consistent with deletion of the coding sequence for the C_H1 domain from this mRNA. However, direct evidence for deletion of the C_H1 coding sequence from the IF2 mRNA is lacking. Furthermore, the mutation is not understood at the gene level. To further characterise the molecular genotype of this mutation, we have cloned cDNA copies of the IF2 heavy chain mRNA in a bacterial plasmid. This paper describes the characteristics of one cDNA clone.

Understanding of the control mechanisms and possible RNA processing events

for any gene requires a precise definition of the limits of the mRNA product of that gene. Cloned copies of mRNA, which frequently lack one or both ends of the mRNA, can therefore fail to define the 5' and 3' ends of the mRNA. Hence, we also describe sequence data at the 3' end of the IF2 heavy chain mRNA by direct priming on the poly(A) tail of the mRNA in order to derive the site for polyadenylation of the γ1 heavy chain mRNA.

MATERIALS AND METHODS

Preparation of cDNA clones

Heavy chain mRNA was prepared from microsomes of IF2 tissue culture cells as described previously [5] and cDNA copies of the mRNA prepared as described by Matthyssens and Rabbitts [6]. 20 μg mRNA were added to a 0.5 ml incubation mixture containing 50 mM Tris-HCl (pH 8.3) at 42°C, 10 mM $MgCl_2$, 10 mM dithiothreitol (DTT), 50 mM KCl, 10 μg/ml oligo$(dT)_{12-18}$ (P.-L. Biochemicals), 150 units/ml AMV reverse transcriptase, 1 mM of the unlabelled deoxynucleoside triphosphates dCTP, dGTP, dATP and dTTP in the presence of 150 μCi α-^{32}P dATP (~300 Ci/mmole). After 60 min incubation at 42°C, the reaction mixture was boiled in 0.3 N NaOH for 3 min, neutralised, sodium dodecyl sulphate (SDS) added to 0.3% final and passed through a Sephadex SP50 column in 0.3 M NaCl, 0.01 M Na acetate pH 5. The DNA was EtOH-precipitated without addition of carrier. The cDNA was fractionated on a 5-20% sucrose gradient in 0.1 M NaOH, 0.5% SDS, and material greater than 1000 bases long was pooled and concentrated by ethanol precipitation. This cDNA fraction was made double-stranded by synthesis using the Klenow fragment of DNA polymerase [7]. The double-stranded cDNA was cleaved by the single-strand specific S_1-nuclease (to generate blunt-ended molecules) in 0.28 M NaCl, 4.5 mM $ZnSO_4$, 30 mM sodium acetate, pH 4.4 at 37°C [8,9], and concentrated by ethanol precipitation after phenol extraction. The double-stranded cDNA was dissolved in water and adjusted to 140 mM cacodylate (pH 7.6), 30 mM Tris, 100 μM DTT, 1 mM cobalt chloride, 100 μM dATP (^{3}H-dATP, 100 μCi/mmole), and 50-100 dA residues were added per 3' end using terminal transferase [10]. The bacterial plasmid pBR322 [11] was cleaved with BamHI (which cuts within the tetracycline resistance gene), and poly(dT) tails were added (about 70 residues per 3' end) using terminal transferase as described above. The dA-elongated cDNA and dT-elongated pBR322 were annealed in an equimolar ratio and the mixture used to transform *E. coli* χ1776 to ampicillin-resistant tetracycline-sensitive colonies according to G. Wahl and R. Padgett (personal communication). *E. coli* χ1776 cells (50 ml, A_{550} = 0.2) were harvested,

resuspended in 20 ml 70 mM $MnCl_2$, 30 mM $CaCl_2$, 10 mM sodium acetate (pH 5.5) and held on ice for 15 min. The cells were recovered and resuspended in 2.5 ml of the same buffer and the DNA for transformation added to the cells. The mixture of competent cells and DNA was incubated on ice for 15 min, at 25° C for 4 min, and again on ice for 30 min, and then spread on agar plates containing 20 μg/ml ampicillin, 50 mg/ml diaminopimelic acid and 200 μg/ml glucosamine.

Preliminary analysis of cDNA clones

Tetracycline-sensitive, ampicillin-resistant clones were grown on nitrocellulose filters and the filters were processed according to Grunstein and Hogness [12] and hybridised to high specific activity ^{32}P-cDNA (see below). The hybridisation solution (3 ml/filter) included 0.45 NaCl, 0.045 sodium citrate, 0.1% SDS, 50 μg/ml sonicated salmon sperm DNA, 10 μg/ml poly(rA) and 2 g/litre each bovine serum albumin, polyvinylpyrrilidone and Ficoll-400.

Hybrid-arrested translations were performed as described by Paterson et al. [13] using messenger-dependent reticulocyte lysates [14]. The products were analysed on discontinuous 12% acrylamide gels containing SDS.

Preparation of high specific activity ^{32}P-cDNA

Ten μg of an IF2 heavy chain mRNA sample was electrophoresed at 10 m amps for 19 hr in a 1.5% agarose (Miles LGT) gel containing 5 mM methylmercury [15]. The gel was stained with 1 μg/ml ethidium bromide, 10 mM DTT and photographed under ultraviolet light. The 16S mRNA band [5] was cut out, and the agarose melted at 60° C in 0.25 ml 0.5% SDS. The resulting solution was centrifuged in a Ti50 rotor for 2 hr at 25° C and 35,000 rpm. Seven μg yeast tRNA was added to the supernatant, and nucleic acids were precipitated with ethanol. Complementary DNA was synthesised using ^{32}P-α-dATP as described above, except that no cold dATP was added to the reaction mixture.

Sequencing of recombinant DNA

Two methods were used. In the first, kinase labelled restriction fragments were sequenced as described by Maxam and Gilbert [16]. The A>C reaction in 1.5 M NaOH, 0.1 mM EDTA and the alternative G reaction were used.

In the second, Sau 3A restriction fragments from pγ1/A5 were purified on a 6% acrylamide gel, cloned in a derivative of M13mp2 [17] containing a single BamHI restriction enzyme site and an amber mutation (G.P. Winter, unpublished) and sequenced as described by Schreier and Cortese [18].

Direct sequencing of mRNA

RNA sequencing was carried out on IF2 heavy chain mRNA preparations as described by Hamlyn et al. [19] using the primers T_{10}C-A, T_8A and T_5G-C.

Some experiments were designed to define the sequence adjacent to the poly(A) tail. For this purpose the length of the cDNA transcripts was limited by performing the enzymatic elongation of a dT_{10} primer in the absence of dTTP. Transcripts, labelled with ^{32}P-α-dATP, ^{32}P-α-dGTP or ^{32}P-α-dCTP (Amersham, 350 μCi/mole), were prepared in separate incubation mixtures containing (in 10 μl) ca. 0.05 μg dT_{10}, ca. 0.05 μg mRNA, 1 unit reverse transcriptase, 60 mM KCl, 20 mM DTT, 6 mM $MgCl_2$, 50 mM Tris (pH 8.3 at 37° C) and 2 μM of the other two nucleotide triphosphates. After 1 hr at 37° C, the products were fractionated on DEAE-cellulose thin layers by homochromatography [20]. The products were eluted and analysed as described by Proudfoot [20].

RESULTS AND DISCUSSION

Plasmids containing cDNA inserts and the IF2 deletion at the RNA level

Double-stranded cDNA was prepared from RNA samples enriched in IF2 heavy chain mRNA [5] and inserted into the BamHI site of pBR322 by the A:T tailing method. Approximately 240 ampicillin-resistant, tetracycline-sensitive clones were isolated. Since heavy chain mRNA preparations are a heterogeneous mixture of several species, we expected only a portion of these 240 clones to contain heavy chain cDNA inserts. Twelve of these clones gave positive signals in a Grunstein-Hogness filter hybridisation test using cDNA prepared from mRNA further purified by gel electrophoresis under denaturing conditions. Some of these clones were analysed in the hybrid-arrested translation assay (Fig. 1). One clone (pγ1/A5) arrested the translation of the IF2 heavy chain mRNA, indicating the presence of heavy chain cDNA sequences which annealed to the mRNA and prevented its translation. The structure of the plasmids which did not significantly arrest the translation of the heavy chain mRNA is less clear. Plasmids A31, A34 and B124 may not contain the heavy chain cDNA sequences, or the inserts may be either too small or anneal to the mRNA in a position which does not lead to inhibition of translation.

The recombinant plasmid pγ1/A5 was further analysed by sequencing experiments. These experiments showed that the IF2 heavy chain cDNA insert extends from the nucleotides coding for amino acid 96 in the variable region to the sequence A-C-C-C-T-G-C in the 3' untranslated region (Figs. 2 and 3). The nucleotide sequence of pγ1/A5 is colinear with the amino acid sequence of the IF2 mutant heavy chain. Specifically, the nucleotides coding for amino acids 121-214 of the normal γ1 heavy chain are deleted. Direct sequence analysis reveals no discontinuity at the site of the deletion, so that the boundary

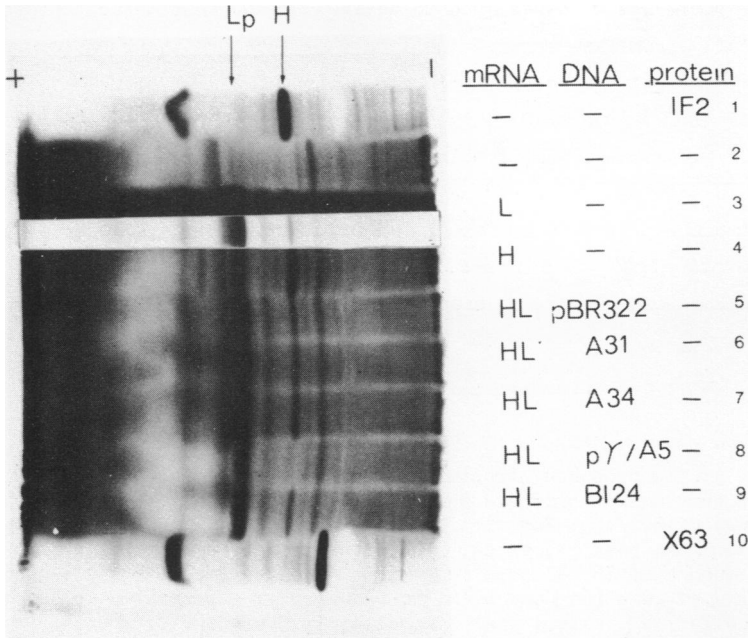

Figure 1. Hybrid-arrested translation using recombinant plasmids. Approximately 2 μg of various plasmids were hybridised to 0.1 μg of a mixture of IF2 light and heavy chain mRNA preparations. The nucleic acids were precipitated and translated (lanes 5-9). Included as controls of the translation reaction were no added mRNA (lane 2), light chain mRNA preparation alone (lane 3) and IF2 heavy chain mRNA preparation alone (lane 4). As the light chain mRNA gave very strong translation products, two autoradiographic exposures are shown for lane 3. Also shown for reference are the secreted proteins of IF2 cells (lane 1) and X63 cells (a subclone of the wild type P3 cells - lane 10).

between the C_H1 and hinge region coding sequences is replaced by a new V-region-hinge region boundary as predicted previously [4,5]. This deletion of C_H1 in both the protein and mRNA could reflect deletion of C_H1 in the gene. However, as the $\gamma1$ C_H1 is encoded by exactly one exon [2], it is also possible that RNA splicing signals could be altered by point mutations or removed by deletion. Such changes in RNA splicing signals could also result in the observed deletion in the mRNA and protein.

Sequence at the 3' end of the IF2 heavy chain mRNA

The methodology used to prepare the double-stranded DNA insert in cDNA plasmids does not guarantee the presence of the full untranslated regions in the resultant recombinant plasmids. To investigate the sequences in the mRNA

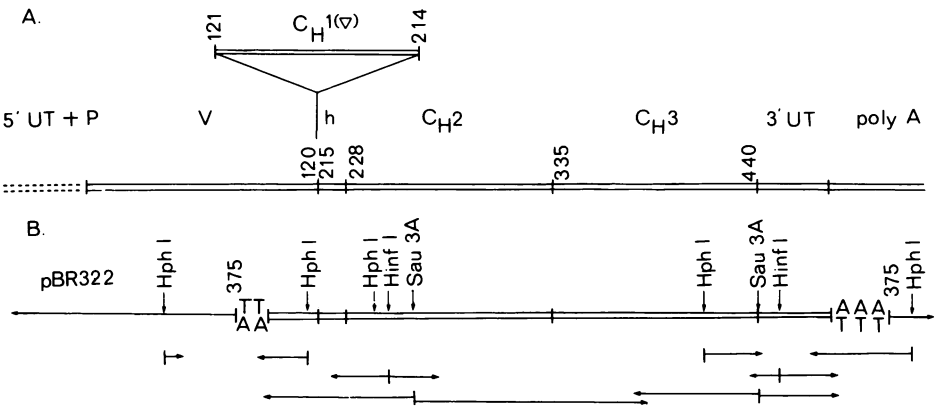

Figure 2. A: The proposed structure of the IF2 heavy chain mRNA. Residue numbers of some amino acids are shown and (above) the deleted C_H1 domain. B: Structure of pγ1/A5. Restriction sites used in sequencing are shown. The entire insert has been sequenced. The sequence agrees with the published partial sequence of the γ1 gene from MOPC 21 cells [2], with the sequence of the germline γ1 gene [21] and with the sequence of a partial γ1 cDNA clone from X63 cells [22], except that at position 915 we find G, whereas Rogers et al. [22] found A, and at positions 1177, 1182 and 1183 we found G, A and G, whereas Honjo et al. [21] found A, G and A, respectively. In addition to corrections in the amino acid sequence [23] previously noted [2,21,22] our nucleotide sequence indicates Trp at residue 99 rather than His, and His at residue 104 rather than Trp.

adjacent to the poly(A) tail, we have synthesised cDNA in the absence of dTTP using the primer dT_{10}. When dTTP is omitted from reverse transcription reactions in which pT_{10} is used as a primer, transcription must initiate next to the poly(A) tail and terminate at the first adenosine residue in the mRNA. IF2 heavy chain mRNA was reverse-transcribed under such conditions; the products were fractionated (by size) using homochromatography (Fig. 4) and sequenced by nearest neighbour and depurination analysis (Table 1). Three major products were T_{10}C-A-A-G, T_{10}A-A-A and T_{10}G-C-A-G-G-G. Sequencing experiments, using synthetic primers and dideoxynucleotides, were then performed to identify the origin of these products.

T_{10}C-A-A-G corresponds to the 3' end of the light chain mRNA [13] which, due to its aggregation, is a common impurity of the heavy chain mRNA. T_{10}C-A has been used as a light chain mRNA primer [19]; its use with IF2 heavy chain mRNA preparations demonstrated priming of a single major sequence corresponding to the light chain mRNA. It was concluded that T_{10}C-A-A-G is derived from light chain mRNA impurities.

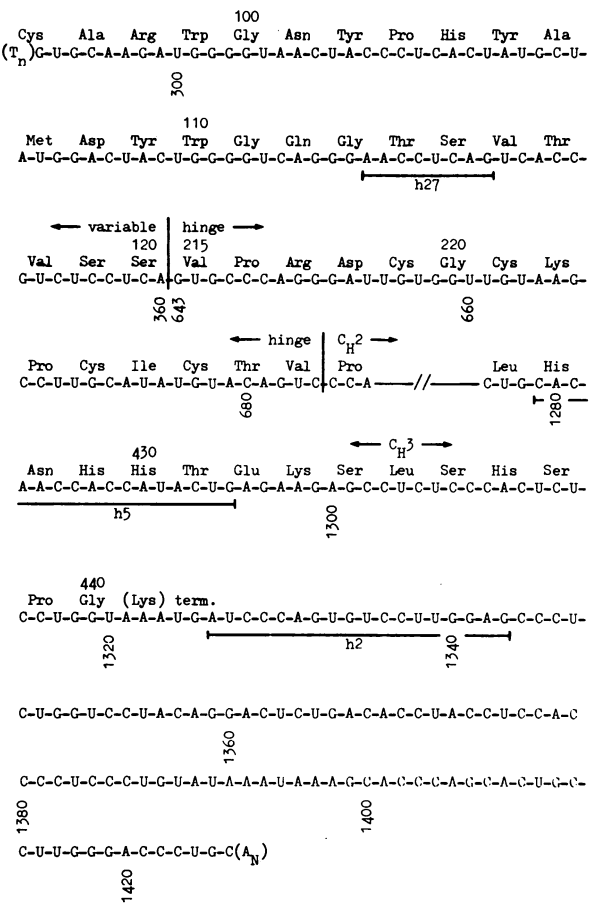

Figure 3. The nucleotide sequence of the C-terminus and 3' untranslated region of the IF2 heavy chain mRNA. This sequence was determined: by the method of Maxam and Gilbert [16] on HphI and HinfI restriction fragments of pγ1/A5; by the dideoxy method [18] on Sau 3A fragments cloned in phage M13; by the primed synthesis method [19] on mRNA (residues 1276-1352 only). Nucleotide number 1 is the G corresponding to the N-terminal Asp of the IF2 mutant mRNA. Three T1 oligonucleotides (h2, h5, h27) have been previously identified.

Experiments using T_8A and heavy chain mRNA fractions yielded a nucleotide sequence (about 150 residues) unrelated to either light or heavy chain mRNA. It seems that the T_{10}A-A-A product of dTTP limited synthesis is derived from an unidentified contaminant mRNA in the heavy chain mRNA preparation.

The third product of cDNA synthesis in the absence of dT (T_{10}G-C-A-G-G-G)

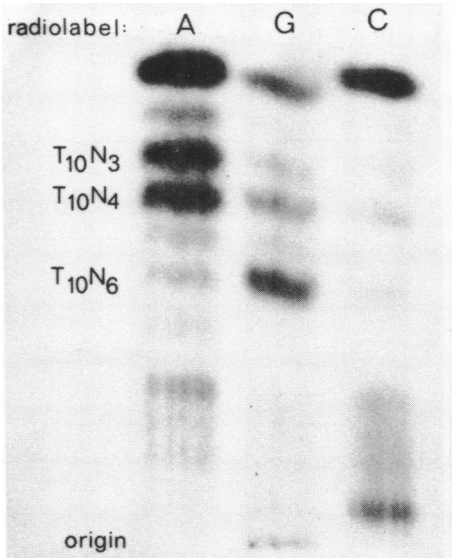

Figure 4. Fractionation of products of dT_{10} primed reverse transcription in the absence of dTTP. IF2 heavy chain mRNA fractions were reverse transcribed in the presence of ^{32}P-dATP, ^{32}P-dGTP or ^{32}P-dCTP as described in the Materials and Methods section. See Table 1 for determination of sequences. The product with the highest mobility in each slot ($T_{10}N$) is due to an artifactual template independent addition of one nucleotide to any oligonucleotide by reverse transcriptase. Several products are formed in low yield which are smaller than the major "build-up" products $T_{10}N_3$, $T_{10}N_4$ and $T_{10}N_6$. The sequences of these minor products are consistent with the hypothesis that they are premature termination products related to the major -dTTP products. Products longer than $T_{10}N_6$ were not analysed.

was investigated using the primer T_5G-C on the heavy chain mRNA. A sequence was established for transcripts longer than 50 residues and corresponded to residues 1276-1352 of the pγ1/A5 sequence (Fig. 3) and thus included both coding and 3' untranslated regions. These data indicate that sequence C-C-C-U-G-C poly(A) (complementary to the product T_{10}G-C-A-G-G-G) is the 3' end of the IF2 mRNA. We conclude, therefore, that the polyadenylation site is C-1426 (Fig. 3). However, two A residues follow C-1426 in the γ1 gene sequence [21] and thus either A could also be the polyadenylation site.

This conclusion is in contradiction to the suggestion of Honjo et al. [21] who proposed that the poly(A) addition site for the γ1 mRNA is G-1419. Our cDNA clone extends seven residues beyond G-1419 (Fig. 3) and therefore G-1419 cannot be the poly(A) addition site for the IF2 mRNA. It could be argued that the difference in apparent polyadenylation site is due to the IF2 mutation. However, the phenotype of the IF2 mutation involves the residues coding for the C_H1 domain, far removed from the 3' end of the mRNA. Also, the 3' untranslated sequences of the IF2 mRNA and the γ1 gene [21] are identical. We suggest that C-1426 or one of the two As following C-1426 is the polyadenylation site for both the γ1 wild type and the IF2 mutant.

Honjo et al. [21] have pointed out the homology between γ1 and γ2b 3'

Table 1. Sequences of products of limited synthesis. Nearest neighbour and depurination analyses were performed as described [20]. Some nearest neighbours were present in twice the yield of other products and this has been indicated in the analysis column. All reactions were minus T reactions.

Primer	Product	Analysis	Deduced sequence
T_{10}	$T_{10}N_6$	Nearest neighbours 1 G → T 1 C → G 1 A → C 1 G → A 2 G → G Depurination G → T_{10} A → C	T_nC-C-A-G-G-G(T)
T_5G-C	T_5G-C-N_4	Nearest neighbours 1 A → C 1 G → A 2 G → G	
T_{10}	$T_{10}N_4$	Nearest neighbours C → T A → C A → A G → A	T_nC-A-A-G(T)
T_{10}	$T_{10}N_3$	Nearest neighbours 2 A → A 1 A → T	T_nA-A-A(T)
T_8A	T_8A-N_2	Nearest neigbbours A → A	

untranslated regions. These sequences, however, have no significant homology with either the κ 3' untranslated region [19] or the mouse μ heavy chain 3' untranslated region [6].

Both our sequence of the IF2 heavy chain mRNA (whether derived from primed synthesis on the mRNA or from the cDNA clone), and that of the γ1 germline gene [21], show a lysine codon after the C-terminal glycine of the secreted protein and before the UGA terminator. It is probable that lysine is the C-terminus of the initial translation product, and is removed post-translationally. The post-translational removal of C-terminal lysine has also been suggested for a human epsilon heavy chain [24].

ACKNOWLEDGEMENTS

We thank Dr. J.W. Beard for supplying reverse transcriptase, Dr. H. Boyer for providing pBR322, Dr. R. Curtiss for providing χ1776, and we also thank Mr. A. Forster for technical assistance. W.D. was a fellow of the Jane Coffin Childs Memorial Fund for Medical Research, and this work was supported in part by a grant from that Fund.

REFERENCES

[1] Franklin, E.C. and Frangione, B. (1975) in "Contemporary Topics in Molecular Immunology" (ed. F.P. Inman), Plenum Press, New York, pp. 89-126.
[2] Sakano, H., Rogers, J.H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. and Tonegawa, S. (1979) Nature 277, 627-633.
[3] Early, P.W., Davis, M.M. Kaback, D.B., Davidson, N. and Hood, L. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 857-861.
[4] Adetugbo, K., Milstein, C. and Secher, D.S. (1977) Nature 265, 299-304.
[5] Cowan, N.J., Secher, D.S. and Milstein, C. (1976) Eur. J. Biochem. 61, 355-368.
[6] Matthyssens, G. and Rabbitts, T.H. (1980) Nucleic Acids Res., accepted for publication.
[7] Maniatis, T., Kee, S.G., Efstratiadis, A. and Kafatos, F.C. (1976) Cell 8, 163-182.
[8] Sutton, W.D. (1971) Biochim. Biophys. Acta 240, 522-531.
[9] Shenk, T.E., Rhodes, C., Rigby, P.W.J. and Berg, P. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 989-993.
[10] Roychoudhury, R., Jay, E. and Wu, R. (1976) Nucleic Acids Res. 3, 101-116.
[11] Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heneke, H.L. and Boyer, H.W. (1977) Gene 2, 95-113.
[12] Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 3961-3965.
[13] Paterson, R.P., Roberts, B.E. and Kuff, E.L. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4370-4374.
[14] Pelham, R.B. and Jackson, R.J. (1976) Eur. J. Biochem. 67, 247-256.
[15] Bailey, J.M. and Davidson, N. (1976) Anal. Biochem. 70, 75-85.
[16] Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 560-564.
[17] Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P.H. (177) Proc. Natl. Acad. Sci. U.S.A. 74, 3642-3646.
[18] Schreier, P.H. and Cortese, R. (1979) J. Mol. Biol. 129, 169-172.
[19] Hamlyn, P.H., Brownlee, G.G., Cheng, C.-C., Gait, M.J. and Milstein, C. (1978) Cell 15, 1067-1075.
[20] Proudfoot, N.J. (1976) J. Mol. Biol. 107, 491-525.
[21] Honjo, N. and Mano, Y. (1979) Cell 18, 559-568.
[22] Rogers, J., Clarke, P. and Salser, W. (1979) Nucleic Acids Res. 6, 3305-3321.
[23] Adetugbo, K. (1978) J. Biol. Chem. 253, 6068-6075.
[24] Bennich, H., Milstein, C. and Secher, D.S. (1973) FEBS Lett. 33, 49-53.