Isolation and analysis of genomic DNA clones encoding the third component of mouse complement

(genomic DNA library/promoter region/signal peptide/DNA sequence analysis)

KARIN WIEBAUER*, HORST DOMDEY†, HEIDI DIGGELMANN, AND GEORGE FEY‡

Department of Molecular Biology, Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland

Communicated by Hans J. Müller-Eberhard, July 15, 1982

ABSTRACT A gene library was constructed with DNA from strain A mice by using the phage λ vector λ 1059. By screening with cloned cDNA for the third component of mouse complement, C3, four different C3 genomic clones were isolated from this library. Two of the recombinant phages carry insertions of 14 and 18 kilobase pairs, respectively, which together cover one complete copy of the C3 gene and several hundred nucleotides of its 5' and 3' flanking sequences. The distance from the 5' end of the gene, which includes the hexanucleotide T-A-T-A-A-A and a translation initiation codon, to its 3' end as defined by the poly(A) attachment site is 24 kilobase pairs. From the genomic DNA sequence, a signal peptide of 24 amino acid residues is predicted at the NH₂ terminus of the initial translation product. The signal peptide and the next two amino acids are encoded by the first exon of this gene.

The third component of complement, C3, is a serum glycoprotein that plays an essential role in the defence of mammalian organisms against microbial and viral infections. It also participates in inflammatory and immunological reactions (1, 2). Qualitatively, the expression of the C3 gene is controlled in a tissue-specific manner: the protein is produced mainly by hepatocytes and, to a lesser extent, by monocytes, macrophages, splenic and lymphoid cells, by other tissues of the reticuloendothelial system, and by a few epithelial tissues (3, 4). In humans, C3 concentrations are significantly increased in maternal serum late in pregnancy and also during acute-phase inflammatory reactions (5, 6). The biosynthesis of C3 was reported to be increased several fold by hydrocortisone in a rat hepatoma cell line in tissue culture (7). These observations suggest that quantitative controls of the expression of the C3 gene exist and that glucocorticoids and possibly other steroid hormones might participate in its regulation.

The C3 protein has several distinguishable physiological functions mediated by specific peptides derived from it. The C3a peptide fragment of C3 is an anaphylatoxin which induces inflammatory responses. The C3d fragment harbors a reactive site, through which the protein covalently attaches to particle surfaces. Furthermore, the C3b and C3d polypeptides carry binding sites recognized by other complement proteins, by two specific regulatory proteins of the complement system, and by distinct receptor proteins present on the surfaces of a variety of effector cells (8–11).

With the aid of C3 genomic clones, we wish to correlate the exon-intron structure of the gene with known functional domains of the polypeptide and to study the control of its expression. In this report, we describe (i) the isolation and characterization of recombinant phages that contain a complete copy of the gene, (ii) the delineation of the promoter region and the poly(A) attachment site, and (iii) the detection of sequences cod-

ing for a signal peptide (12, 13). This signal peptide and the following two amino acids were found to be encoded by the first exon of the gene.

MATERIALS AND METHODS

Construction and Screening of a Mouse Gene Library. Liver DNA from adult strain A/Jax mice was prepared as described by Groner and Hynes (14). The DNA was partially digested with the restriction enzyme *Bam*HI, and fragments were separated by electrophoresis in a 0.6% agarose gel (low-gelling-temperature agarose; type VII, Sigma). Partial restriction fragments of 13–23 kilobases (kb) were recovered by repeated phenol extractions in the presence of 0.5 M NaCl (15) and were ligated into the *Bam*HI site of the replacement phage λ vector λ 1059 (16) with T4 DNA ligase. Infectious phage particles were reconstituted by *in vitro* packaging (17, 18) and used to infect *E. coli* Qu 358 cells (16). The resulting plaques were screened as described (19) with nick-translated (20) C3 cDNA radiolabeled probes (21, 22). Work with recombinant phages was carried out in the P3 facility of our Institute in accordance with the National Institutes of Health guidelines for recombinant DNA research.

Preparation and Analysis of Cloned DNA. Recombinant phage from large-scale lysates was purified by equilibrium centrifugation in cesium chloride density gradients, and DNA was extracted from phage particles by standard procedures (23). Transfer of DNA fragments from agarose gels to nitrocellulose was performed as described by Southern (24). Endonuclease restriction maps were constructed by established methods (25, 26), and DNA sequence determinations were carried out by the method of Maxam and Gilbert (27).

Preparation of Primer-Extended cDNA. The preparation of primer-extended cDNA was done essentially as described by Hagenbüchle *et al.* (28). Approximately 12 μ g of size-selected, polyadenylated mouse liver mRNA, sedimenting at 28 S in a sucrose gradient (21), and 540 ng of the 5'-proximal 650-base pair (bp) *Hind*III–*Bam*HI C3 cDNA fragment (22) were mixed and heated for 10 min at 50°C in 45 μ l of 90% formamide. Five μ l of 0.1 M Pipes, pH 6.4/4 M NaCl/10 mM EDTA was added, and the sample was incubated for another 90 min at 50°C. The sample was then diluted with an equal volume of 10 mM Pipes/0.4 M NaCl/1 mM EDTA and precipitated with ethanol. The precipitate was redissolved directly in 25 μ l of 50 mM Tris-HCl, pH 8.3/7.5 mM KCl/7.5 mM MgCl₂/1 mM dithiothreitol/4

Abbreviations: C3 and C4, third and fourth components of complement; bp, base pair(s); kb, kilobase(s); KW1, KW4, KW7, and KW10, recombinant λ phages λ MC3/KW1, λ MC3/KW4, λ MC3/KW7, and λ MC3/KW10.

* Present address: Dept. of Human Genetics, Univ. of Michigan Medical School, Ann Arbor, MI 48109.

† Present address: Max Planck Institut für Biochemie, D-8033 Martinsried, Federal Republic of Germany.

‡ To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

mM sodium pyrophosphate containing 0.8 mM each of dGTP, dTTP, and dCTP and 90 μ Ci (1 Ci = 3.7×10^{10} becquerels) of [α - 32 P]dATP (specific activity, 450 Ci/mmol). The sample was preincubated for 5 min at 45°C, and the reaction was started by the addition of 20 units of avian myoblastosis virus reverse transcriptase. After incubation for 90 min at 42°C, the reaction was stopped by addition of 20 mM EDTA/0.5 M NaCl/0.2% NaDodSO₄ (final concentrations). Unincorporated nucleotides were removed by column filtration through Bio-Gel A15m (Bio-Rad). Preparation of primer-extended cDNA by using the single-stranded 26-nucleotide *Hae* III–*Bam*HI subfragment of the 232-bp *Hind*III–*Bam*HI fragment that contains the T-A-T-A-A box (see Figs. 3 and 4) was performed exactly as described by Hagenbuechle *et al.* (28).

Materials. A/Jax mice were bred in the animal facility of our Institute. Restriction enzymes and T4 DNA ligase were from Bethesda Research Laboratories; polynucleotide kinase was from Boehringer Mannheim; nitrocellulose BA 85 was from Schleicher & Schuell; and [32 P]NTPs were from the Radiochemical Centre. Avian myoblastosis virus reverse transcriptase was provided by J. Beard (Office of Program Resources and Logistics, Viral Cancer Program, National Cancer Institute).

RESULTS

Isolation of Genomic Clones Carrying C3 Sequences. Recombinant phages containing C3 genomic DNA sequences were identified among a total of 1.2×10^6 plaques by hybridization with cloned C3 cDNA. Six C3 genomic DNA clones were isolated and their DNA was characterized by restriction enzyme analysis with *Bam*HI, *Hind*III, and *Eco*RI. The inserts of four of the six clones gave rise to different but partially overlapping cleavage patterns. This allowed an unambiguous assignment of the orientation for the four DNA insertions with respect to each other and of the order of restriction fragments within the inserts. In doubtful cases, the arrangement was determined by analyzing partial *Bam*HI digests of 5' end-labeled DNA fragments. The resulting map of restriction sites for the cloned phages λ MC3/KW1 (KW1), λ MC3/KW4 (KW4), λ MC3/KW7 (KW7), and λ MC3/KW10 (KW10) is shown in Fig. 1. Two of the six initial isolates were identical to KW10. Recombinant phages KW4 and KW7 carry abutting genomic DNA fragments of 14 and 18 kb, respectively, ending at a common *Bam*HI site as shown by the overlapping restriction patterns of their inserts

with that of KW1. The insert sequences of KW10 are entirely contained in those of KW7. The sum of all the cloned C3 genomic sequences is represented by the inserts of KW4 and KW7. Together they span a contiguous stretch of 32 kb of genomic DNA.

In order to determine which fraction of the C3 gene is contained in KW4 and KW7, Southern blot hybridizations were carried out with *Bam*HI restriction fragments of DNA from these phages immobilized on nitrocellulose. The filters were hybridized in parallel with seven different nick-translated C3 cDNA restriction fragments of known map positions with respect to the C3 mRNA. All of these probes, representing more than 90% of the mRNA, including its 3' end, were found to hybridize with *Bam*HI fragments of genomic DNA mapping between the 2.7-kb *Bam*HI fragment of KW4 and the 9.0-kb *Bam*HI fragment of KW7 in the same relative order as in the cDNA map (data not shown). These results suggested that the major part of the C3 gene is contained in KW4 and KW7 and that these two clones represent the 5' and 3' halves of the gene, respectively.

Localization of the 3' and 5' Ends of the Gene. Hybridization data of *Bam*HI/*Hind*III-digested DNA from KW7 with a 3'-terminal cDNA probe, suggested that the 3' end of the gene was located near a unique *Hind*III restriction site within the terminal 9.0-kb *Bam*HI fragment of KW7. In order to map the precise position of the 3' end of the gene, the sequence of ≈ 300 nucleotides encompassing this *Hind*III site was determined and compared with the known 3'-terminal sequence of C3 cDNA (22). The sequence of the last 182 nucleotides of the mRNA was found to be identical with the genomic DNA sequence (Fig. 2). The poly(A) attachment site was found 16 nucleotides downstream from the *Hind*III site. The genomic DNA sequence diverged from the cDNA sequence at two points: at this poly(A) attachment site, where the cDNA continues with a series of adenosines, and 182 nucleotides further upstream (arrow in Fig. 2). The dinucleotide A-G found at this position most likely marks the 3' border of the last intron of the gene.

By hybridizing 5'-terminal radiolabeled cDNA sequences to *Bam*HI/*Hind*III-digested DNA from recombinant phage KW4 in a Southern blot hybridization experiment, the 5' end of the gene was mapped. The cDNA probe was prepared by reverse transcription of the 5'-terminal sequences of C3 mRNA with a 5'-proximal cloned C3 cDNA fragment as a primer, so called primer-extended cDNA. The result suggested that the 5' end

FIG. 1. (a) Restriction map of the C3 gene. *Bam*HI restriction fragments are shown by open boxes (the 0.13-kb fragment marked by an asterisk is possibly located towards the 3' side of the neighboring 0.75-kb fragment). $\circ$, *Eco*RI sites; Δ , *Hind*III sites; ∇ , Positions of the T-A-T-A-A box (5' end of the gene) and of the poly(A) addition site (3' end of the gene). (b) Restriction map of genomic DNA inserts carried by four different C3 clones. Vertical bars indicate *Bam*HI restriction sites.

FIG. 2. Mapping of the 3' end of the C3 gene. Comparison of the nucleotide sequences of genomic DNA (upper row) and of the 3' end of C3 mRNA (22) (lower row); dashes represent homologous nucleotides. The DNA sequence of genomic DNA was determined in both directions from the *Hind*III site in the 9-kb *Bam*HI fragment of KW10 (see Fig. 1). Where mRNA sequences are written in the bottom row, they differ from genomic sequences. The first six adenosine residues of the poly(A) tail are shown. The arrow in the top row marks the limit of the last intron of the gene.

of the gene was located within a 232-bp *Hind*III–*Bam*HI subfragment of the 1,500-bp *Bam*HI fragment of KW4 (Fig. 3). The DNA sequence of this fragment was determined (Fig. 4). Downstream from the *Hind*III site by 76 bp, the hexanucleotide T-A-T-A-A-A was found, which is a typical sequence element of eucaryotic RNA polymerase promoter regions (29, 30). In order to map the cap site on the genomic DNA, a single-stranded end-labeled 26-bp *Hae* III–*Bam*HI fragment was used as a primer for reverse transcription of C3 mRNA. The size of extended cDNA was precisely measured by electrophoresis alongside the sequence ladder of the 232-bp *Hind*III–*Bam*HI fragment in a 6% polyacrylamide/urea gel. Thereby, the 5' end of the mRNA was mapped 28 ± 2 bp downstream from the center of the T-A-T-A-A-A box. The first ATG triplet was located 84 bp downstream from the T-A-T-A-A-A box, at the beginning of an open reading frame that extended over 81 nucleotides, part of which matched exactly the 5' sequences determined for the cDNA (see next section).

Nucleotide Sequences Encoding the C3 Signal Peptide. From the nucleotide sequence of primer-extended C3 cDNA, we have identified the sequence that encodes the first 10 amino acid residues from the NH₂ terminus of the C3 β polypeptide subunit: NH₂-Ile-Pro-Met-Tyr-Ser-Ile-Ile-Thr-Pro-AsN (22). Coding sequences for these amino acids were also found in the

genomic clone KW4; however, they were interrupted by an intron starting from the codon for methionine, the third residue in this amino acid sequence (Fig. 4). The ATC codon specifying the first Ile residue was located within the *Bam*HI site at the boundary of the 232-bp *Hind*III–*Bam*HI fragment that contains the T-A-T-A-A-A box described above (Fig. 3). Between the ATG translation initiation triplet and this Ile codon, we found another 23 triplets in the same reading frame, uninterrupted by translation termination codons. No other translation initiation signal was found between the T-A-T-A-A-A hexanucleotide and this Ile codon in any of the other reading frames. Therefore, the DNA sequence predicts that the initial translation product of the C3 mRNA should contain 24 additional amino acids beyond the NH₂-terminal Ile residue of the mature C3 β subunit. The amino acid sequence of this extension peptide, deduced from the genomic DNA, is shown in Fig. 4. By comparing the DNA sequences of genomic DNA and primer-extended cDNA covering this region, it could further be concluded that the first exon of the gene encodes the 24 amino acids of this NH₂-terminal extension peptide plus the next two residues (Ile and Pro) and the beginning of the third residue (Met) of the C3 β subunit. The remainder of the Met triplet and triplets for the next 62 amino acid residues of the C3 β subunit were found in the second exon. DNA sequences were further determined for se-

FIG. 3. Topography of the 5' end of the C3 gene. (a) Restriction map of the 5'-terminal region of the gene (clone KW4) for the enzymes *Eco*RI (↓), *Hind*III (↑), and *Bam*HI (vertical bars). (b) Fine structure of the 232-bp *Hind*III–*Bam*HI fragment, which contains the 5' end of the gene. Indicated are the T-A-T-A-A-A box, the first nucleotide (cap) of C3 mRNA, the translation initiation codon ATG, the 72-bp coding for the signal peptide, and the ATC triplet that specifies the NH₂-terminal Ile residue of the C3 β subunit.

FIG. 4. (Upper) Nucleotide sequence of the 232-bp *Hind*III–*Bam*HI fragment, which contains the 5' end of the gene and partial sequence of the 3'-neighboring 900-bp and 2,700-bp *Bam*HI fragments. Restriction sites mentioned in the text, the T-A-T-A-A hexanucleotide, the ATG translation initiation codon, and the first and last two nucleotides of the first intron are underlined. $\uparrow$, The exact exon/intron boundaries; $\downarrow$, the probable first nucleotide of the C3 mRNA. The NH₂-terminal amino acid residues, comprising the signal peptide and the first amino acids of the C3 β subunit, are written above the triplet codons from which they are deduced. (Lower) Block structure of the first, second, and the beginning of the third exon. The heavy line in the bottom row represents exons. The thin inclined line symbolizes intron sequences that are absent from the mRNA and the cDNA.

lected regions of the 2.7-kb *Bam*HI fragment and compared to known cDNA sequences (22). By this comparison it was possible to determine the exact position of the second and the start of the third exon in the gene (Fig. 4 Lower).

DISCUSSION

Two lines of evidence suggest that the recombinant phages KW4 and KW7 contain a complete copy of the mouse C3 gene including ≈ 1.2 kb of 5'-flanking and 7 kb of 3'-flanking sequences: (i) restriction fragments from the insertions carried in these phages hybridized with a variety of C3 cDNA fragments, which represent over 90% of the C3 mRNA sequences, and (ii) DNA sequence analysis showed that these genomic clones contain coding sequences for C3 mRNA that are identical to the corresponding cDNA sequences (22). The length of the mouse C3 gene from the cap site to the poly(A) attachment site is 24 kb—4 to 5 times longer than the C3 mRNA, which measures $5,100 \pm 200$ nucleotides (22).

The 5' end of the gene has been located by the position of the T-A-T-A-A hexanucleotide, by an ATG translation initiation codon, and by the map position of the 5' end of the mRNA. We conclude that the ATG triplet (Figs. 3 and 4) is used to initiate the translation of the C3 polypeptide because (i) it lies nearest to the only T-A-T-A-A element in the 232-bp *Hind*III–*Bam*HI fragment in which the 5' end of the C3 mRNA was mapped, and (ii), starting with this triplet, a polypeptide sequence can be synthesized in an open reading frame, which after 24 triplets enters into codons for Ile, Pro, and the beginning of a Met codon. These last three amino acids constitute the NH₂ terminus of the mouse C3 β chain according to our cDNA sequence data (22). The coding sequence is then interrupted by an intron and continues in the second exon for a further 188

nucleotides starting from the last nucleotide of the Met codon.

Secreted glycoproteins are frequently processed from intracellular precursor proteins that contain an NH₂-terminal extension peptide, a leader peptide, or signal peptide. These peptides serve to insert the NH₂ termini of the nascent protein chains into the membranes of the rough endoplasmic reticulum and to direct their transport across the membrane (vectorial transport). They are generally removed during translation by specific proteolytic cleavage (cotranslational processing; ref. 13). As C3 is a secreted glycoprotein, it was to be expected that its intracellular precursor should contain a single peptide. The genomic DNA sequence reported above predicts the existence of an NH₂-terminal extension peptide in the initial C3 translation product of the following sequence: Met-Gly-Pro-Ala-Ser-Gly-Ser-Gln-Leu-Leu-Val-Leu-Leu-Leu-Leu-Ala-Ser-Ser-Pro-Leu-Ala-Leu-Gly. This extension peptide possesses several characteristic properties of signal peptides (31): (i) its length falls into the usual range of 20–30 amino acid residues; (ii) it consists exclusively of uncharged amino acids; (iii) its last residue is glycine, a preferred residue for the proteolytic action of the signal peptidase; (iv) within its last six residues, a proline is found; and (v) its central portion consists of eight hydrophobic residues in an uninterrupted sequence. Therefore, we propose that the initial translation product of C3 mRNA contains an NH₂-terminal extension peptide with the above sequence, which functions as a signal peptide. It has so far not been detected at the protein level. If it were removed by cotranslational processing, it would have escaped detection in published experiments, which were designed to analyze the NH₂ terminus of the intracellular precursor polypeptide of C3 called pro-C3 (32).

From the comparison of genomic DNA and cDNA sequences, we conclude that the signal peptide and the following two amino acids of the C3 β subunit are encoded by a separate

first exon of the gene. Signal peptides of other glycoproteins (immunoglobulins, ovomucoid; refs. 33–35) have been observed to be encoded by separate exons. This suggests that the signal peptide is a distinct functional domain of the protein and possibly of independent genetic origin. Whether other functional domains of the C3 protein are encoded by separate exons should be determined by future sequence analysis.

In order to investigate whether the mouse genome contains one or several variants of the C3 gene, mouse liver DNA was subjected to Southern blot analysis in parallel with DNA from recombinant phages KW4 and KW7. A mixture of cloned C3 cDNA fragments that covered all of the C3 mRNA sequences that have been obtained as cloned cDNA until now, plus the 232-bp *Bam*HI–*Hind*III genomic DNA fragment containing the 5' end of the gene, were used as probes. All of the *Bam*HI restriction fragments, which were detected in the digest of genomic DNA, also were found in one of the two phages. Similar results also were obtained with the restriction enzymes *Eco*RI and *Hind*III (unpublished data). Therefore, the mouse genome carries only one type of C3 gene recognizable by this analysis. If more than one copy of the C3 gene were present, then all should give rise to the same *Bam*HI, *Eco*RI, and *Hind*III restriction patterns. No pseudogenes of the mouse C3 gene were detected. In this respect, the mouse C3 gene is similar to the human C3 gene (36) but quite different from the human and murine C4 genes. Although the C4 genes are most likely derived from the same ancestral gene as the genes specifying C3 and C5, they are duplicated inside the major histocompatibility complex of both species (37, 38). In contrast, the C3 gene is located outside that complex in both species (36, 39).

We thank Drs. G. McMaster, U. Schibler, B. Tack, K. Odink, and B. Hirt for discussions and support; Drs. J. Karn and S. Brenner for the phage vector; Dr. E. Buetti for help during the preparation of *in vitro* packaging mixes; V. Müller for excellent technical assistance; and S. Cherpillod for expert help with the preparation of the manuscript. K. Wiebauer and H. Domdey were recipients of postdoctoral fellowships from the European Molecular Biology Organization and Deutsche Forschungsgemeinschaft, respectively. This work was supported by a research grant from the Swiss National Science Foundation.

1. Müller-Eberhard, H. J. & Schreiber, R. D. (1980) *Adv. Immunol.* **29**, 1–53.
2. Porter, R. R. & Reid, B. M. (1979) *Adv. Protein Chem.* **33**, 1–71.
3. Colten, H. R. (1976) *Adv. Immunol.* **22**, 67–118.
4. Fey, G. & Colten, H. R. (1981) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **40**, 2099–2104.
5. Propp, R. P. & Alper, C. A. (1968) *Science* **162**, 672–673.
6. Kushner, I., Edgington, T. S., Trimble, C., Liem, H. H. & Müller-Eberhard, U. (1972) *J. Lab. Clin. Med.* **80**, 18–25.
7. Strunk, R. C., Tashjian, A. J. & Colten, H. R. (1975) *J. Immunol.* **114**, 331–335.
8. Hugli, T. E. & Müller-Eberhard, H. J. (1978) *Adv. Immunol.* **26**, 1–53.
9. Tack, B. F., Harrison, R. A., Janatova, J., Thomas, M. L. & Prahl, J. W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5764–5768.
10. Fearon, D. T. (1980) *J. Exp. Med.* **152**, 20–30.
11. Lambris, J. D., Dobson, N. J. & Ross, G. D. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1828–1832.
12. Milstein, C., Brownlee, G. G., Harrison, T. M. & Mathews, M. B. (1972) *Nature (London) New Biol.* **239**, 117–120.
13. Blobel, G. & Dobberstein, B. (1975) *J. Cell. Biol.* **67**, 852–862.
14. Groner, B. & Hynes, N. E. (1980) *J. Virol.* **33**, 1013–1025.
15. McMaster, G. K., Beard, P., Engers, H. D. & Hirt, B. (1981) *J. Virol.* **38**, 317–326.
16. Karn, J., Brenner, S., Barnett, L. & Cesareni, G. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5172–5176.
17. Hohn, B. & Murray, K. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3259–3263.
18. Hohn, B. (1979) *Methods Enzymol.* **68**, 299–309.
19. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
20. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
21. Odink, K. G., Fey, G., Wiebauer, K. & Diggelmann, H. (1981) *J. Biol. Chem.* **256**, 1453–1458.
22. Domdey, H., Wiebauer, K., Müller, V., Kazmaier, M. & Fey, G. (1979) *Proc. Natl. Acad. Sci. USA*, in press.
23. Davis, R. W., Botstein, D. & Roth, J. R. (1980) *Advanced Bacterial Genetics—A Manual for Genetic Engineering* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).
24. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
25. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387–2398.
26. Schibler, U., Pittet, A.-C., Young, R. A., Hagenbüchle, O., Tosi, M., Gellman, S. & Wellauer, P. K. (1982) *J. Mol. Biol.* **155**, 247–266.
27. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
28. Hagenbüchle, O., Bovey, R. & Young, R. A. (1980) *Cell* **21**, 179–187.
29. Goldberg, M. (1979) Dissertation (Stanford University, Stanford, CA).
30. Corden, J., Waslylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. & Chambon, P. (1980) *Science* **209**, 1406–1414.
31. Austen, B. M. (1979) *FEBS Lett.* **103**, 308–313.
32. Goldberger, G., Thomas, M. L., Tack, B. F., Williams, J., Colten, H. R. & Abraham, G. N. (1981) *J. Biol. Chem.* **256**, 12617–12619.
33. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1485–1489.
34. Stein, J. P., Catterall, J. F., Kristo, P., Means, A. R. & O'Malley, B. W. (1980) *Cell* **21**, 681–687.
35. Gilbert, W. (1981) *Science* **214**, 1305–1312.
36. Whitehead, A. S., Solomon, E., Chambers, S., Bodmer, W. F., Povey, S. & Fey, G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5021–5026.
37. Awdeh, Z. L. & Alper, C. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3576–3580.
38. Shreffler, D. C. (1976) *Transplant. Rev.* **32**, 140–167.
39. Penalva da Silva, F., Hoecker, G. F., Day, N. K., Vienne, K. & Rubinstein, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 963–965.