Molecular cloning of developmentally regulated, low-abundance mRNA sequences from embryonic muscle

(cDNA cloning/embryo-specific transcripts/cell differentiation)

CHARLES P. ORDAHL*, DIMITRIS KIOUSSIS†‡, SHIRLEY M. TILGHMAN†‡, CATHERINE E. OVITT*, AND JAMES FORNWALD*

*Department of Anatomy and †Fels Research Institute, Temple University Medical School, Philadelphia, Pennsylvania 19140

Communicated by Sidney Weinhouse, May 1, 1980

ABSTRACT To study the role of low-abundance, embryonic muscle-specific gene transcripts, we have developed a method to screen cDNA clones from embryonic muscle for such sequences. The protocol involves two stages: first, partial enrichment for cDNA clones carrying possible embryo-specific sequences by selecting clones of low-abundance sequences; and second, determination, by hybridization to RNA attached to diazobenzylxymethyl-paper, which sequences from this category are regulated in an embryonic muscle-specific manner during development. At least three different clones were obtained which hybridized to sequences present in early muscle development but absent, or present at relatively low levels, at late embryonic and adult muscle stages. Two of these clones were not muscle-specific because they hybridized to poly(A)⁺RNA from liver or brain or both. The third clone, 106A4, did not detectably hybridize to total poly(A)⁺RNA at any stage of brain or liver development tested. This sequence also was not detectable in poly(A)⁺RNA from embryonic muscle progenitor cells. Thus, the 106A4 sequence is a likely candidate for an embryonic muscle-specific sequence. We have demonstrated that the 106A4 sequence is a mRNA, although the specific identity and function of the translated product is unknown. The method used to identify embryonic muscle-specific cDNA clones should be generally applicable for obtaining clones for low abundance transcripts regulated in a tissue-specific or developmental stage-specific manner.

Recombinant DNA techniques have permitted the isolation and detailed characterization of the organization of a number of eukaryotic genes. In general these techniques have been most successful when the mRNA of interest is a predominant species in a population of mRNAs. However, mRNA sequences present in cells at low or even intermediate abundance levels are difficult or impossible to clone selectively by using these same approaches, unless they have an unusual structural feature permitting selective enrichment of them (1).

We are interested in studying a class of low-abundance transcripts that are under developmental regulation during muscle differentiation. As an *in vivo* model system we have used leg muscle development in the embryonic chicken because it is well characterized in its biochemical as well as its more classical embryological aspects (2–5). Recent work from many laboratories has shown that specific genes and gene sets are differentially expressed during muscle development (6–9). In terms of transcriptional regulation it is possible to discriminate at least two broad categories of sequences that are under developmental-regulation during embryonic muscle differentiation. In the first category are the muscle-specific genes which are transcribed at relatively low levels early in muscle development but which produce highly abundant transcripts at the terminal phase of myogenesis (7, 10–12).

A second, and less-well understood, class of genes are those which are transcribed at early stages but whose transcripts are undetectable at the later terminal myogenic stages (6, 8). The details of the regulation of these latter sequences remain unclear, largely because their presence at relatively low levels even at early developmental stages makes analysis difficult. However, because these latter sequences are both developmentally regulated and apparently embryo-specific they are good candidates for being involved in the critical early events of cell differentiation in which pluripotent progenitor cells “decide” which terminal phenotype to express.

In order to study these low-abundance sequences we have developed a protocol for identifying their cDNA clones derived from reverse transcription of early embryonic muscle total cell poly(A)⁺RNA. We obtained several clones that appear to belong to this category. The approach we have used should be generally useful for isolating cDNA clones of sequences present in cells at low frequency which are regulated in either a tissue-specific or developmental stage-specific manner.

METHODS AND MATERIALS

Preparation of RNA. Total cell RNA was extracted from embryonic and adult tissues as described by Rowe *et al.* (13) except that digestion of homogenates with proteinase K was omitted. After washing once with 3 M sodium acetate/5 mM EDTA to remove cellular DNA, the poly(A)⁺RNA fraction was prepared by chromatography on oligo(dT)-cellulose (14). Recovery from a single chromatography step was approximately 2–3% of input total cell RNA.

Preparation of Chimeric Plasmids. Double-strand cDNA (ds cDNA) for cloning was synthesized using 5 µg of day 10 embryo leg muscle total cell poly(A)⁺RNA according to the method of Wickens *et al.* (15) except for the following modifications: (i) the synthesis of cDNA first strand was conducted as described (16); (ii) for the synthesis of cDNA second strand, an additional 0.5 µmol of MgCl₂ was added to supplement the MgCl₂ removed as a coprecipitate with sodium pyrophosphate after boiling of the first strand reaction mixture.

ds cDNAs were introduced into the unique *Hind*III site of pBR322 (17) by using the dG-dC tailing and hybridizing protocol of Villa-Komaroff *et al.* (18). The ds cDNA was tailed with approximately 30 residues at the 3'-OH with terminal deoxynucleotidyl transferase as described (19). Because a low background of parental transformation is essential in these experi-

Abbreviations: bp, base pair; cDNA, complementary DNA synthesized by reverse transcription; DBM-paper, diazobenzylxymethyl-paper; ds cDNA, double-stranded cDNA; RNA paper, DBM-paper to which RNA has been covalently attached.

‡ Present address: Institute for Cancer Research, Fox Chase, Philadelphia, PA 19111.

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.

FIG. 1. Size analysis of cDNA inserts. Plasmid DNA prepared from chloramphenicol-treated cultures as described (26, 27) was digested with *Eco*RI and *Bam*HI and electrophoresed on a 1% agarose gel. Lanes: A, pBR322 DNA; B, clone 106C8 DNA; C, 107A6 DNA; D, 106A4 DNA. Net cDNA insert sizes were: 106C8, 541 bp; 107A6, 451 bp; and 106A4, 394 bp.

ments, we treated *Hind*III-linearized pBR322 with nuclease S1, extracted with chloroform, and excluded on Sephadex G-100. The treated pBR322 was then tailed with approximately 30 dG residues at its 3'-OH terminus, hybridized with the cDNA, and used to transform LE 392 by the procedure of Villa-Komaroff *et al.* (18). Approximately 95% of ampicillin-resistant colonies were sensitive to tetracycline, indicating the presence of recombinant plasmids. The overall transformation efficiency was $5\text{--}40 \times 10^3$ recombinants per μg of ds cDNA.

Hybridization to Paper-Bound RNA. Twenty-five micrograms of total cell poly(A)⁺RNA was electrophoresed on 1.5% agarose gels containing 10 mM methyl mercuric hydroxide as described (20). Chicken ribosomal RNA was used as size

markers (2160 and 4160 nucleotides, respectively) (21). Electrophoresed RNAs were blotted onto diazobenzylxymethyl-paper (DBM-paper; Schleicher & Schuell) according to their protocol which is a modification of that described by Alwine *et al.* (22); this product is called "RNA paper." After blotting, the RNA paper was "prehybridized" as described by Wahl *et al.* (23). Hybridization was conducted for 16 hr at 42°C in 3–10 ml of a modified Denhardt's solution (23) with $10\text{--}30 \times 10^6$ cpm of plasmid [³²P]DNA, labeled by nick-translation (24). The RNA paper was washed and exposed to x-ray film (23) for 4–100 hr.

RESULTS

Selection of Low-Abundance cDNA Clones. Embryonic muscle ds cDNA was prepared by using day 10 embryonic leg muscle total poly(A)⁺RNA as a template. After nuclease S1 treatment to remove single-strand regions and hairpin loops, the cDNA was tailed with dC and hybridized with the dG-tailed pBR322 vector, and the hybrid plasmid was used to transform pBR322. Transformants were screened for drug resistance and then replica plated onto nitrocellulose filters and prepared for colony hybridization.

Colony hybridization (25) was performed with labeled cDNA synthesized by using adult leg muscle or adult brain poly(A)⁺RNA as templates. After radiographic exposure, colonies failing to show hybridization (approximately 60% of the total) were selected as "low-abundance" clones and small (5 ml) cultures were grown from each to prepare a small amount of plasmid DNA for size analysis of the cDNA insert. Cloned plasmids with inserts >200 base pairs (bp) were then grown in larger cultures to prepare sufficient plasmid DNA for more detailed analysis. Of approximately 350 original recombinant clones, 40 were finally selected by the above criteria for further analysis by hybridization.

Hybridization Screening for Embryo-Specific Sequences. To determine which cDNA clones contained embryo-specific sequences, 10–20 μg of plasmid DNA was prepared and the insert plus about 400 bp of plasmid DNA was excised by digestion with *Eco*RI and *Bam*HI. Fig. 1 shows an electrophoretogram of three recombinant clones digested in this manner. For each, the presence of insert cDNA is indicated by the slower migration of the smallest fragment compared to that of pBR322 DNA digested in the same manner. After isolation of the insert fragments from agarose gels (28), they were labeled to a specific activity of 20–100 cpm/pg by nick-translation (24) in the presence of [³²P]dCTP. The labeled cDNA inserts were then hybridized to total poly(A)⁺RNA from various embryonic and adult tissues displayed on DBM-paper (22, 23). In order to provide a screen for phenotype specificity, we used RNA from three tissue systems: muscle, brain, and liver (Figs. 2 and 3). The embryonic as well as adult stages of brain (lanes F, G, and H) and liver (lanes I and J) were chosen to permit discrimination of possible embryo-specific sequences that might be ubiquitous from those that might be specific for early muscle differentiation.

The stages of muscle development used were chosen to span the developmental range from a pluripotent progenitor cell population (Figs. 2 and 3, lane A) to adult muscle (lane E). In general terms the muscle developmental stages can be characterized (2–5) as follows: stage 24 limb buds, contain a mesodermal core of progenitor cells pluripotent for muscle, cartilage, and possibly other phenotypes; day 10 leg muscle, predominantly myoblasts which proliferate but do not accumulate muscle terminal differentiation marker molecules; day 14 leg muscle, a mixture of myoblasts and fused, terminally differentiating multinucleate myotubes; day 18 leg muscle, almost

FIG. 2. Detection of cloned embryonic muscle cDNA sequences in developing muscle, brain, and liver. Twenty five micrograms of total cell poly(A)⁺RNA from each of the tissues indicated below was electrophoresed on agarose gels containing 10 mM methyl mercuric hydroxide, blotted into DBM-paper, and hybridized to ³²P-labeled cloned cDNA insert. The hybridized RNA paper was then exposed to x-ray film for 16–100 hr to determine the positions of hybridized RNA. (Left) Clone 106C8, 16 hr exposure; (Right) clone 107A6, 96 hr exposure. Lanes (RNA samples): A, stage 24 embryo (day 4½) limb bud; B, day 10 embryo (Ed 10) leg muscle; C, Ed 14 leg muscle; D, Ed 18 leg muscle; E, adult leg muscle; F, Ed 11 brain; G, Ed 18 brain; H, adult brain; I, Ed 18 liver; J, adult liver.

FIG. 3. RNA paper hybridization of clone 106A4. In both panels, lanes A–J are as in Fig. 2. (Right) Additional lanes K and L contain poly(A)⁺RNA from stage 20 and 24 whole embryos (2, 5); 100 hr radiographic exposure. (Left) 16 hr radiographic exposure.

entirely fused, multinucleate myotubes which are contractile and synthesize and accumulate muscle marker molecules in quantity.

Our primary goal in these experiments was to search for sequences that demonstrated embryo specificity in the context of muscle development. Each of the three cDNA clones shown in Figs. 2 and 3 demonstrated this specificity in that all hybridized to poly(A)⁺RNA sequences that were clearly present at the earliest muscle stage tested (day 10, lane B) but were absent, or present at very low abundance, in adult muscle (lane E). With the exception of clone 106A4, to be discussed in detail below, none of the sequences shown demonstrated strict embryo specificity: they hybridized to the adult as well as embryonic stages of liver or brain.

Clone 106C8 (Fig. 2 Left) hybridized strongly to a 3000-nucleotide RNA species in poly(A)⁺RNA from stage 24 limb buds (lane A), day 10 leg muscle (lane B), and day 14 leg muscle (lane C). This species was present at greatly decreased levels at late embryonic stages of leg muscle development (day 18, lane D) and in adult leg muscle (lane E). This 3000-nucleotide poly(A)⁺RNA species also was clearly present at all tested stages of brain (lanes F, G, and H) and liver (lanes I and J) development. Therefore, although the 3000-nucleotide poly(A)⁺RNA species demonstrates embryo specificity in muscle, it is present throughout the development of brain and liver. Although present in brain, its copy frequency is too low to have permitted its detection during the colony hybridization screening.

In adult leg muscle (lane E), hybridization of 106C8 also occurred to a novel species of poly(A)⁺RNA approximately 1520 nucleotides in length. Upon long-term exposure, a small amount of this second species could be detected both in day 18 leg muscle (lane D) and in adult liver (lane J). In addition, long-term exposure demonstrated up to three additional hybridizing species in adult leg muscle, one of which was larger than 3000 nucleotides. The meaning of this hybridization pattern is unclear. The possibility that the 3000-nucleotide species is simply a nuclear limited precursor to a smaller mRNA is discounted by the observation that *in vivo* it exists in a poly-some configuration (data not shown). It is possible, therefore, that the 3000- and 1520-nucleotide poly(A)⁺RNA species are transcribed from different genes which share regions of homology.

Clone 107A6 (Fig. 2 Right) hybridized to a 3560-nucleotide poly(A)⁺RNA present at very low abundance levels in early limb buds (lane A) and early leg muscle (lane B) but in extremely low to undetectable levels at all later muscle development stages (lanes C, D, and E). The conclusion that the 107A6

sequence is a very low-abundance sequence is based upon the longer exposure times necessary for detection (96 hr compared to 16 hr for 106C8). (This long exposure time is responsible for the increased background in this radiograph.) Despite the extremely low level of 107A6 in later stages of muscle development, it was detectable at low levels throughout brain development (lanes F, G, and H). We have been unable to detect its presence above background in liver (lanes I and J).

At present we can only speculate as to the function of sequences that are restricted to the embryonic stages of one tissue but are expressed throughout the development of another tissue. One obvious possibility is that such sequences code for products needed by many embryonic tissues but only a few adult tissues. A second possibility is that, during their embryonic status, cells may transcribe phenotypically inappropriate genes which they subsequently repress during terminal differentiation (30). By using cloned probes for developmentally regulated sequences it should be possible to begin to test such possibilities.

Cloned cDNA 106A4 May Be Embryonic Muscle-Specific. Cloned sequence 106A4 hybridized to a 1560-nucleotide poly(A)⁺RNA species only in early embryonic muscle (Fig. 3 Left, lanes B and C). This hybridization pattern is consistent with what one would predict for an embryonic muscle-specific sequence. There are three salient observations in support of such a designation.

First, 106A4 is probably muscle-specific because it did not hybridize detectably with poly(A)⁺RNA from either brain or liver (Fig. 3 Right) even with very long exposure times. The possibility remains that 106A4 is present in the nonadenylylated RNA fraction of the brain and liver, but a hybridization assay of this fraction is not feasible due to the dilution by the vast quantities of ribosomal RNA. Finally, because liver and brain obviously do not exhaust the variety of possible nonmuscle phenotypes, this uncertainty is unavoidable at present.

Second, total poly(A)⁺RNA from stage 24 limb buds (day 4½) also did not contain detectable quantities of the 106A4 sequence (Fig. 3 Right, lane A). In the chicken limb, phenotypic commitment of mesenchyme cells occurred at, or shortly after, stage 25 (approximately day 5 of incubation) (2). Prior to this time, the limb mesoderm appears to be plastic in terms of developmental potency for the various limb mesodermal phenotypes although phenotypic "proclivities" have already been established (31, 32). Therefore, the transcription of the 106A4 sequence appears to be either turned-on, or greatly potentiated, in muscle after phenotypic commitment of the pluripotent mesenchyme cells in the limb bud. This sequence is not specific to limb development because total cell poly(A)⁺RNA from whole embryos at stage 24 and at stage 20 contained it in appreciable quantities (lanes K and L). This result is presumably due to the presence of muscle tissue in other areas of the embryo which begin to differentiate earlier than the limb buds. These observations lead to the tentative conclusion that, in the limb, detectable transcription (and possibly first transcription) of the 106A4 sequence occurs after phenotypic commitment of muscle cells.

Third, transcription of the 106A4 sequence was decreased dramatically in fully fused muscle tissue. Fig. 3 Right shows that this sequence was clearly detectable in fused muscle tissue (lanes D and E) after long exposure times. Thus, the 106A4 sequence was not completely absent in the adult but rather it demonstrated embryo specificity only in a quantitative sense. It may also be relevant, in this regard, to point out that adult muscle retains a small population of embryo-like cells, known as satellite cells, which are capable of proliferating like myoblasts and subsequently fusing to form authentic myotubes (33). Thus, it is possible that the low levels of 106A4 sequence de-

tected in adult muscle may be contributed by this small subpopulation of mononucleate satellite cells.

Our prediction is that the 106A4 sequence is specific for mononucleate myoblasts which predominate at early muscle development stages. We would stress that our assignment of this sequence to myoblasts, or even muscle, at this point is only inferential because muscle tissue always contains a large number of nonmuscle cell types (4) and it is possible that this sequence is derived from one of these companion nonmuscle cell types. If this were the case, however, we would expect to detect this sequence also in RNA from liver or brain tissue because the companion nonmuscle cell types are expected to be virtually ubiquitous. Inevitably it will be necessary to identify unequivocally the cell that synthesizes the 106A4 sequence. The most satisfying proof would involve cytological *in situ* hybridization to determine microscopically which cells contain this sequence. The limits of detection in such experiments dictate that the sequence of interest be present at relatively high intracellular copy frequency. Unfortunately, the 106A4 sequence is present at a relatively low copy frequency per cell as estimated from the RNA paper hybridization. We have tried to quantitate the cellular concentration of 106A4 by using the cloned fragment for solution hybridization or as an affinity hybridization medium (34, 35). We find, as have others (34, 35), that precise quantitation of very low-abundance sequences is difficult. Our preliminary results place the copy frequency at between 15 and 60 copies per muscle cell nucleus at embryonic day 9.

FIG. 4. Determination that the transcript complementary to cDNA clone 106A4 is mRNA. Postmitochondrial supernatant from day 12 embryo leg muscle was fractionated on sucrose gradients containing Mg^{2+} or EDTA to separate polysomal and postpolysomal ribonucleoprotein by the technique of Ouellette *et al.* (29). After oligo(dT)-chromatography to isolate poly(A)⁺RNA, equivalent proportions (25%) recovered from each fraction (not exceeding 10 μ g) were electrophoresed, blotted onto DBM-paper, and hybridized to ^{32}P -labeled 106A4 insert as described in Fig. 3. Postpolysomal and polysomal poly(A)⁺RNA from Mg^{2+} -containing gradients are shown in lanes A and B, respectively. Identical fractions from EDTA-containing gradients are shown in lanes C and D.

106A4 Sequence Is a mRNA. Because total poly(A)⁺RNA was used for the RNA paper hybridization shown in Figs. 2 and 3 it is possible that the sequences detected were not translated mRNAs but either nuclear RNAs or mRNAs that were packaged in embryonic muscle for periods of time in ribonucleoprotein particles that are inactive in protein synthesis (36–38). We therefore wished to establish whether 106A4 is associated with polysomes in a manner consistent with its being actively translated. A postmitochondrial supernatant from day 12 leg muscle was loaded onto preparative sucrose gradients containing either 7.5 mM $MgCl_2$ or 20 mM EDTA, centrifuged, and fractionated into postpolysomal and polysomal fractions as described by Ouellette *et al.* (29). The poly(A)⁺RNA prepared from each fraction was displayed on RNA paper for hybridization.

Fig. 4 shows that, in Mg^{2+} -containing gradients (which preserve polysome structure), the 106A4 sequence was roughly equally distributed between the postpolysomal and polysome fractions, indicating that approximately half the 106A4 sequence is associated with polysomes. In the comparable fractions from the EDTA-containing gradient, however, all of the 106A4 sequence was distributed in the postpolysome region and none was detected in the polysome region. This indicates that the 106A4 sequence does not cosediment with polysomes fortuitously but is associated with an authentic polysome structure capable of disruption with EDTA. This result is consistent with the 106A4 sequence being an actively translated message sequence in embryonic muscle. However, our attempts to isolate sufficient quantities of the RNA complementary to 106A4 for translation *in vitro* have not yet been successful.

DISCUSSION

We have shown that it is practicable to develop screening protocols for the isolation of recombinant clones of classes of low-abundance mRNAs which are regulated in definable terms. Here, we were interested in examining representatives from a class of low-abundance sequences which are present at early myogenic stages but not at later ones. Therefore, our screening protocol utilized hybridization to RNA from different developmental stages of muscle as well as from other tissues. It is not difficult to envisage similar screening protocols which vary other features such as physiological state or hormonal induction. Thus, in its broad application, the protocol we have used should be applicable to the study of representative low-abundance mRNAs in various experimental systems.

The detailed analysis of regulated, low-abundance mRNAs may prove to be of critical importance to our understanding of the role that differential gene expression plays in differentiation. Low-abundance mRNAs, which constitute the most complex class of messages, must code for the majority of diverse proteins in the cell. Moreover, Galau *et al.* (39) indicated that a significant fraction of the histospecifically regulated proteins in liver are coded for by low-abundance mRNAs. The organization and regulation of the genes coding for these mRNAs may therefore be more generally representative than those of highly abundant mRNAs.

A second point of interest is the verification that a population of low-abundance transcripts is regulated in muscle development in "reverse phase" compared to typical muscle-specific gene transcripts. That is, these sequences are present at the earliest myogenic stages and absent, or very decreased in abundance, at the terminal stage of muscle differentiation. Although the existence of such a sequence class in muscle has been demonstrated (6) and implicated from the work of others (8), a detailed analysis of those sequences has been refractory to standard biochemical approaches.

Of the three sequences from this category presented here, only one behaved as if it were specific to the muscle cell type, and this is not yet unequivocally proven. The other two are found in at least one other tissue type and neither appears to be developmentally regulated in these nonmuscle tissues. It will be interesting to determine the roles that the proteins coded for by these sequences play in embryonic muscle differentiation.

We gratefully acknowledge the expert technical assistance of Joyce Brown in the dissection of tissues and embryos and the continuing encouragement and criticism of Arnold I. Caplan. This work was supported by National Institutes of Health Grant 1 R01/GM 25400 and a Basil O'Connor Starter Research Grant from The March of Dimes-National Foundation. C.P.O. is a recipient of National Institutes of Health Research Career Development Award 1 K04 HD 00290. S.M.T. is a recipient of a Junior Faculty Award from the American Cancer Society.

1. Noyes, B. E., Mevarech, M., Stein, R. & Agarwal, K. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1770-1774.
2. Herrmann, H., Heywood, S. M. & Marchok, A. C. (1970) *Curr. Top. Dev. Biol.* **5**, 181-234.
3. Hauschka, S. D. (1968) in *The Stability of the Differentiated State*, ed. Ursprung, H. (Springer, New York), Vol. 1, pp. 37-57.
4. Marchok, A. C. & Herrmann, H. (1967) *Dev. Biol.* **15**, 129-155.
5. Searles, R. (1973) in *Developmental Regulation*, ed. Coward, S. J. (Academic, New York), p. 220.
6. Ordahl, C. P. & Caplan, A. I. (1976) *Dev. Biol.* **54**, 61-72.
7. Paterson, B. M. & Bishop, J. O. (1977) *Cell* **12**, 751-765.
8. Devlin, R. B. & Emerson, C. P. (1979) *Dev. Biol.* **69**, 202-216.
9. Leibovitch, M. P., Leibovitch, S. A., Harel, J. & Kruh, J. (1979) *Eur. J. Biochem.* **97**, 321-326.
10. Strohman, R. C., Moss, P. S., Micou-Eastwood, J., Spector, D., Przybyla, A. & Paterson, B. (1977) *Cell* **10**, 265-273.
11. John, H. A., Patrino-Georgoulas, M. & Jones, K. W. (1977) *Cell* **12**, 501-508.
12. Caravatti, M., Perriard, J. C. & Eppenberger, H. M. (1979) *J. Biol. Chem.* **254**, 1388-1394.
13. Rowe, D. W., Moen, R. C., Davidson, J. M., Byers, P. H., Bornstein, P. & Palmiter, R. D. (1978) *Biochemistry* **17**, 1581-1589.
14. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408-1412.
15. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem.* **253**, 2483-2495.
16. Ordahl, C. P. & Caplan, A. I. (1978) *J. Biol. Chem.* **253**, 7683-7691.
17. Sutcliffe, J. F. (1978) *Nucleic Acids Res.* **5**, 2721-2728.
18. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3727-3731.
19. Roychoudhury, R., Jay, E. & Wu, R. (1976) *Nucleic Acids Res.* **3**, 101-116.
20. Bailey, J. M. & Davidson, N. (1976) *Anal. Biochem.* **70**, 75-85.
21. Leoning, U. E. (1969) *J. Mol. Biol.* **38**, 355-365.
22. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5350-5354.
23. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3683-3687.
24. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1184-1188.
25. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961-3965.
26. Tilghman, S. M., Kioussis, D., Gorin, M. B., Garcia-Ruiz, J. P. & Ingram, R. S. (1979) *J. Biol. Chem.* **254**, 7393-7399.
27. Meagher, R. B., Tait, R. C., Betlach, M. & Boyer, H. B. (1977) *Cell* **10**, 521-536.
28. McDonell, M. W., Simon, M. N. & Studier, F. W. (1977) *J. Mol. Biol.* **110**, 119-146.
29. Ouellette, A. J., Kumar, A. & Malt, R. A. (1976) *Biochim. Biophys. Acta* **425**, 384-395.
30. Caplan, A. I. & Ordahl, C. P. (1978) *Science* **201**, 120-130.
31. Bonner, P. H. & Hauschka, S. D. (1974) *Dev. Biol.* **37**, 317-328.
32. Ahrens, P. B., Solursh, M., Reiter, R. S. & Singley, C. T. (1979) *Dev. Biol.* **69**, 436-450.
33. Bischoff, R. (1979) *J. Cell Biol.* **83**, 387a (abstr.).
34. Kafatos, F., Jones, C. W. & Efstratiadis (1979) *Nucleic Acids Res.* **7**, 1541-1552.
35. Meyuhas, O. & Perry, R. P. (1980) *Gene* **10**, 113-129.
36. Yaffe, D. & Dym, H. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 543-547.
37. Heywood, S. M., Kennedy, D. S. & Bester, A. J. (1975) *FEBS Lett.* **53**, 69-72.
38. Buckingham, M. E., Cohen, A. & Gros, F. (1976) *J. Mol. Biol.* **103**, 611-626.
39. Galau, G. A., Klein, W. H., Britten, R. J. & Davidson, E. H. (1977) *Arch. Biochem. Biophys.* **179**, 584-599.