

Differential processing of RNA transcribed from the single-copy *Drosophila* myosin heavy chain gene produces four mRNAs that encode two polypeptides

(*Drosophila melanogaster*/alternate splicing/alternate polyadenylation/muscle)

CHARLES E. ROZEK* AND NORMAN DAVIDSON†

*Department of Biology, Case Western Reserve University, Cleveland, OH 44106; and †Department of Chemistry, California Institute of Technology, Pasadena, CA 91125

Contributed by Norman Davidson, December 16, 1985

ABSTRACT We report the sequence of genomic DNA at the 3' end of the single-copy *Drosophila* myosin heavy chain (MHC) gene and the structure and sequence at the 3' end of four MHC mRNAs. Two mRNAs, 7.2 kilobases (kb) and 8.0 kb in length, are expressed in all stages of development in which detectable levels of muscle-specific mRNAs accumulate. These mRNAs differ by alternate choice of two poly(A) sites within the same exon. Sequence information predicts that these two mRNAs can encode one MHC polypeptide. Two additional MHC mRNAs, 8.0 kb and 8.6 kb in length, are expressed only in late pupal and adult stages of development. These two stage-specific MHC mRNAs use the same poly(A) sites as the MHC mRNAs described above but have a different splicing pattern and thus include an additional exon. Sequence information predicts that these two stage-specific MHC mRNAs encode a second MHC polypeptide with a different COOH terminus.

Conversion of a primary transcript of a eukaryotic gene into an mRNA involves several steps, including capping, removal of introns, methylation, and polyadenylation. For some structural genes, variations in these processing steps occur in a tissue- or developmental-specific pattern and result in the formation of several different mRNAs from a single transcriptional unit (1). These variations include alternate sites for initiation of transcription (2), alternate splicing patterns (3), and alternate use of poly(A) sites (4). Detailed analyses of several structural genes and their mRNAs have revealed that these alternate processing pathways sometimes result in the formation of mRNAs that encode different protein isoforms (1, 5).

We and others have shown that the *Drosophila* myosin heavy chain (MHC) gene is single copy (6, 7). We have also shown that three size classes of mRNA are transcribed from this gene and that the relative accumulation levels of the three classes vary during development. In this communication, we report that the three size classes of mRNA are composed of four distinct mRNAs that result from the four possible combinations of two alternate splicing patterns and two choices of poly(A) sites. The sequence of cDNA clones predicts that the alternate splicing pattern results in the formation of MHC mRNAs that encode two MHC polypeptides that differ in their COOH-terminal amino acid sequences.

EXPERIMENTAL PROCEDURES

Materials. Restriction enzymes, polynucleotide kinase, *Escherichia coli* DNA polymerase I, and the large fragment

of *E. coli* DNA polymerase I were obtained from Boehringer Mannheim. Avian myeloblastosis virus AMV reverse transcriptase was obtained from Bethesda Research Laboratories. [α - 32 P]NTPs (3000 Ci/mmol; 1 Ci = 37 GBq) and [γ - 32 P]ATP were obtained from ICN.

Construction and Isolation of MHC cDNA Clones. cDNA clones were constructed using the methods of Okayama and Berg (8, 9). Three RNA samples were used for the construction of cDNA libraries: poly(A)⁺ RNA from 15-hr embryos, poly(A)⁺ RNA from 75-hr pupae, and high molecular weight RNA from 15-hr embryos. The high molecular weight 15-hr embryonic RNA was isolated from low-gelling temperature agarose gels containing methyl mercuric hydroxide as described (7). *E. coli* K-12 strain HB101 was used for bacterial transformation (10). Clones containing the desired inserts were selected by hybridization with nick-translated hybridization probes.

MHC cDNA clones were also isolated from a pupal cDNA library constructed by Falkenthal *et al.* (11) using methods described above.

DNA Sequence Analysis. 32 P-labeled DNA fragments for sequencing were labeled either by filling in 3' termini with [32 P]dNTPs with the large fragment of *E. coli* DNA polymerase I or by end-labeling with [32 P]ATP and T4 polynucleotide kinase (12). Both strands of DNA were sequenced using the methods of Maxam and Gilbert (12).

Transcriptional Mapping with S1 Nuclease. Transcriptional mapping was essentially as described by Favaloro *et al.* (13). Hybridizations were in one of two buffers (A or B). Hybridization buffer A (high stringency) contained 80% formamide, 40 mM Pipes (pH 6.4), 400 mM NaCl, and 1 mM EDTA. Hybridization buffer B (low stringency) contained 40 mM Tris-HCl (pH 7.8), 300 mM NaCl, and 2 mM EDTA. Samples were heated to 100°C for 10 min and incubated either at 52°C for 3 hr (buffer A) or 42°C for 16 hr (buffer B). Following incubation, the samples were digested with S1 nuclease and analyzed as described (13).

RNA Blotting. Poly(A)⁺ RNAs from 15-hr embryos and 75-hr pupae were electrophoresed in agarose gels containing formaldehyde, transferred to nitrocellulose, and hybridized with 32 P-labeled probes as described (7).

RESULTS

Overview. For clarity, we begin with a concise statement of the final results. The evidence leading to these interpretations is presented in the following sections.

We reported previously that *Drosophila* has a single MHC gene (7). The transcriptional unit is 19 kilobases (kb) long and

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.

Abbreviations: MHC, myosin, heavy chain; MLC-ALK, myosin light chain dissociated by alkali; EL, embryo-larval; PA, pupal-adult; kb, kilobase(s).

contains at least nine introns. There are three size classes of MHC mRNAs expressed from the gene, with molecular lengths predicted by gel mobilities of 7.2 kb, 8.0 kb, and 8.6 kb. The transcription of the 7.2-kb and 8.6-kb MHC mRNAs appeared to be developmentally regulated. Furthermore, structural studies indicated that the three size classes of mRNAs differed primarily by their patterns of alternate splicing at their 3' ends.

Fig. 1A shows a restriction map of the MHC genomic DNA and the MHC transcriptional unit. We determined previously that the exons involved in alternate splicing are transcribed from a region of the transcriptional unit between the *Sal* I site and the 3' end of the 1.55-kb *Eco*RI fragment (Fig. 1B). S1 nuclease transcriptional mapping and R-loop experiments determined the approximate positions of three MHC exons (C, B, and A) that are involved in the formation of several different MHC mRNAs (Fig. 1C).

The conclusions of the present investigation are summarized in Fig. 2. There are four MHC mRNAs. One pair is expressed in all stages of *Drosophila* development during which muscle mRNAs accumulate (late embryo-larval, EL; late pupal-adult, PA). For these two mRNAs of length 7.2 kb and 8.0 kb, exon C is spliced directly to exon A. The length difference is due to use of the two poly(A) sites, poly(A)-1 and poly(A)-2, within exon A, that are separated by 0.55 kb. These RNAs are named CA-1 and CA-2. The second pair of MHC mRNAs, of length 8.0 kb and 8.6 kb, accumulate only during the PA stage of development. For this pair, exon C is spliced to exon B, which is spliced to exon A. The length difference of these two mRNAs is again due to the use of the two poly(A) sites within exon A. These mRNAs are named CBA-1 and CBA-2. Thus, in our present study, we show that the 8.0-kb size class consists of two different mRNAs, CA-1 and CBA-1, the latter of which is expressed only in the PA stage.

Our sequence determinations predict that there is an in-frame stop codon beginning 82 nucleotides downstream of the exon C-exon A junction, so that exon A of mRNA's CA-1 and CA-2 can code for a 27 amino acid COOH terminus. There is a stop codon just 3 nucleotides after the exon C-exon B junction, so exon B of mRNA's CBA-1 and CBA-2 codes for 1 amino acid and thus these mRNAs can encode a MHC polypeptide with a different COOH terminus.

Isolation of cDNA Clones. Separate cDNA libraries were constructed from RNA isolated from the two stages in development when MHC mRNAs are maximally expressed, late embryonic (15 hr after egg laying) and late pupal (75 hr after pupariation). Transformants were initially screened using as a hybridization probe the 6.0-kb *Eco*RI restriction fragment from the MHC transcriptional unit (Fig. 1A). This restriction fragment had been shown previously to contain sequences that are homologous with all size classes of MHC mRNAs. We thus identified 18 late embryonic cDNA clones

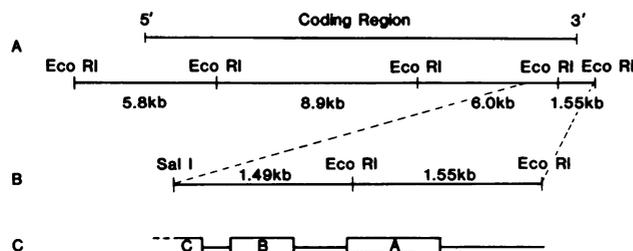


FIG. 1. *Drosophila* MHC gene and the organization of exons at the 3' end of the transcriptional unit. (A) *Eco*RI restriction map of the *Drosophila* MHC gene and the limits of the transcriptional unit. (B) Restriction map of 3' end of the transcriptional unit. (C) Organization of exons at the 3' end of the MHC transcriptional unit.

among 4000 transformants and 48 late pupal cDNA clones among 2000 transformants. Additionally, we isolated two cDNA clones among 10,000 transformants in a late pupal cDNA library constructed by Falkenthal *et al.* (11).

Genomic DNA Sequence. Sequence data extending for 2439 base pairs downstream from the *Sal* I site (within exon C, see Figs. 1 and 2) have been obtained. This sequence, which extends *ca.* 300 base pairs beyond the identified poly(A)-2 site of exon A, is presented in Fig. 3 with our interpretation, based on data presented below, as to the exon-intron junctions. By analysis of the sequences and consideration of our earlier mapping data (7) we identified several restriction endonuclease fragments that were useful as more discriminating hybridization probes to characterize the various MHC mRNAs and their cDNAs.

mRNA CA-1. cDNA clones have been isolated that hybridized with the exon C probe (*Sal* I-*Bgl* II, see Figs. 1 and 2) but not with the exon B probe (*Bgl* II-*Pst* I, Figs. 1 and 2) or with the 3' exon C probe (the 1.55-kb *Eco*RI fragment, Figs. 1 and 2). Such clones have been isolated and sequenced from embryonic and pupal cDNA libraries. Their DNA sequence indicated that they have the structure shown in Fig. 2. The clones terminated with a sequence CAAAATA(*n*) [*n* > 50] corresponding to the poly(A)-1 site of Fig. 3. (None of our data precludes the possibility that one to three of the adenosines following the ribosylthymine at the marked site in the cDNA is genomic rather than added by polyadenylation.) Note that the expected consensus AATAAA sequence is found 13 nucleotides upstream from poly(A)-1.

The structure of the cDNA clone (MHC mRNA CA-1) is consistent with the predicted structure of the 7.2-kb MHC mRNA. However, in our previous studies, we concluded that this RNA was present only in the EL stages not in PA stages. A more careful RNA gel blot using poly(A) RNA and exon-specific probes shows that this mRNA is expressed in the PA stage as well but at a lower intensity than for the 8.0-kb and 8.6-kb mRNAs (Fig. 4A). Furthermore, the 7.2-kb mRNA is detected by hybridization with the 5' exon A-specific probe (*Hpa* II-*Eco*RI, Fig. 4B) but not with the 3' exon A probe (Fig. 4C) or the exon B probe (Fig. 4D). Previous S1 nuclease analysis (7) by hybridization of the end-labeled 1.55-kb *Eco*RI fragment (see Figs. 1-3) in 80% formamide at rather high stringency failed to demonstrate a protected fragment that would represent an EL or PA RNA ending at site poly(A)-1. Examination of our sequence revealed that the segment of exon A of length 117 nucleotides extending downstream from the *Eco*RI site to site poly(A)-1 is only 13% G+C (see Fig. 3). Such an RNA-DNA hybrid would not have been stable in 80% formamide. When an S1 nuclease experiment was performed under less stringent conditions (see *Experimental Procedures*) either with embryonic or with pupal RNA, the expected fragment of length 117 nucleotides (from the *Eco*RI site) was observed (data not shown).

mRNA CA-2. Several cDNA clones were isolated from EL and PA libraries that hybridized to the exon C probe and the 3' exon A probe but not to the exon B probe. Sequence analysis showed that these clones contained the C-to-A splice and the poly(A)-2 site 747 nucleotides downstream from the C-A splice junction. S1 nuclease analysis under stringent conditions confirmed that an mRNA ending in the poly(A)-2 site was present in 15-hr embryonic and late pupal RNAs. Note that a consensus AATAAA site is present 13 nucleotides upstream from the poly(A)-2 site.

The predicted molecular length of this RNA is thus 8.0 ± 0.1 kb relative to the 7.2-kb mRNA. RNA gel blots with the 3' exon A-specific probe show that the 8.0-kb, but not the 7.2-kb embryonic RNA, hybridizes with the 3' exon A-specific probe (Fig. 4C) as well as with the exon C (Fig. 4A) and 5' exon A (Fig. 4B), but not with the exon B probe (Fig.

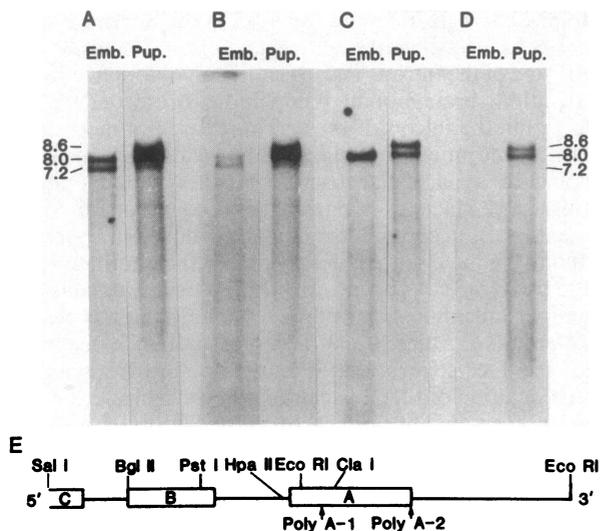


FIG. 4. Hybridization of RNA blots of embryo and pupal poly(A)⁺ RNAs with exon-specific probes. Poly(A)⁺ RNA from embryos (Emb.) (1 μ g) and pupae (Pup.) (3 μ g) was subjected to electrophoresis on 1% agarose gels containing 2.2 M formaldehyde. The RNA was transferred to nitrocellulose and hybridized with the following nick-translated restriction fragments shown in E: *Sal* I-*Bgl* II, exon C probe (A); *Hpa* II-*Eco*RI, 5' exon A probe (B); *Cla* I-*Eco*RI, 3' exon A probe (C); *Bgl* II-*Pst* I, exon B probe (D).

poly(A)-1 termination sites are used in the formation of MHC mRNAs in all stages of development (data not shown); (iii) previously we identified an R-loop structure formed only with PA RNA consistent with the exon arrangement of MHC mRNA CBA-2 (7). Thus, the structure assigned is the only one consistent with these observations.

DISCUSSION

Two Developmentally Regulated Protein Isoforms Due to Alternate Splicing. The data presented above show that there are two splicing patterns at the 3' end of the MHC gene. There are two mRNAs (CA-1 and CA-2) in which exon C is directly spliced to exon A (Fig. 2). This splicing pattern occurs for the late EL and late PA stages of development. The alternative splicing pattern, C to B to A, occurs only for the PA stages of development. For both splicing patterns, two choices of poly(A) sites within exon A are used with approximately equal probability and apparently are not under developmental regulation.

We note that in our previous communication (7) we reported that low-resolution mapping studies gave no evidence for MHC mRNA heterogeneity upstream of exon C. However, this region of length *ca.* 6 kb has not been subjected to the detailed analysis reported here for the 3' ends of the mRNAs. Thus, we cannot rule out the possibility that

additional alternate processing may occur in the 5' region of the MHC transcriptional unit.

The translated amino acid sequences deduced from the cDNA sequences are shown in Fig. 5. Beginning at the *Sal* I site within exon C, there is an open reading frame of length 204 nucleotides, corresponding to 68 amino acids. Splicing of exon A to C results in an open reading frame that can encode an additional 27 amino acids and then a termination codon. This is the predicted amino acid sequence of the COOH terminus of the protein encoded by mRNAs CA-1 and CA-2, which is expressed in EL and PA stages of development. For the exon C-to-exon B splice, there is a single amino acid encoded within exon B, then a termination codon. Thus, the PA-specific protein that can be translated from MHC mRNAs CBA-1 and CBA-2 is predicted to have a shorter and different COOH terminus than does that of the CA-1 and CA-2 translation product.

Bernstein and coworkers (quoted in ref. 14) had determined previously some nucleotide sequence at the 3' end of the *Drosophila* MHC gene. From their sequence they identified an open reading frame. This sequence is in close agreement with that corresponding to the 3' end of the exon C sequence reported by us. McLachlan and Karn (15) have pointed out that there is considerable amino acid homology between the predicted COOH-terminal sequence of the *Drosophila* and the *Caenorhabditis elegans* proteins when compared to the amino acid sequence of a rabbit protein. However, one interesting feature is that the most 3' COOH-terminal amino acids have highly diverged.

Characterization of cDNA clones from several organisms has revealed there are tissue-specific MHC mRNAs that are transcribed from subsets of multigene families (14, 16). DNA sequence analysis has revealed that within a given organism the cDNA clones are highly homologous for large segments of the coding region. In contrast to the high degree of homology in large segments of the coding region, there is a contiguous region in which complete divergence among MHC cDNA sequences has been observed. This gene-specific region comprises the last few amino acids at the COOH-terminal end of the coding segment and the entire 3' untranslated region of the mRNAs (17). Our evidence predicts that, in *Drosophila*, alternate splicing of MHC mRNAs from the single-copy MHC gene can result in the formation of two MHC polypeptides with divergent COOH termini. Thus, the predicted primary sequences of the *Drosophila* MHC polypeptides have similar structural features when compared to predicted MHC polypeptides from other organisms.

The association of MHC polypeptides into dimeric molecules and filaments appears to be a function of the COOH-terminal region. Recent evidence has implicated about 50 residues encompassing the COOH terminus as contributing to the dimerization of MHC polypeptides (18). Several studies have also shown that antibodies specific for the COOH-terminal regions of avian and *Acanthamoeba* MHC peptides can prevent association of MHC monomers to

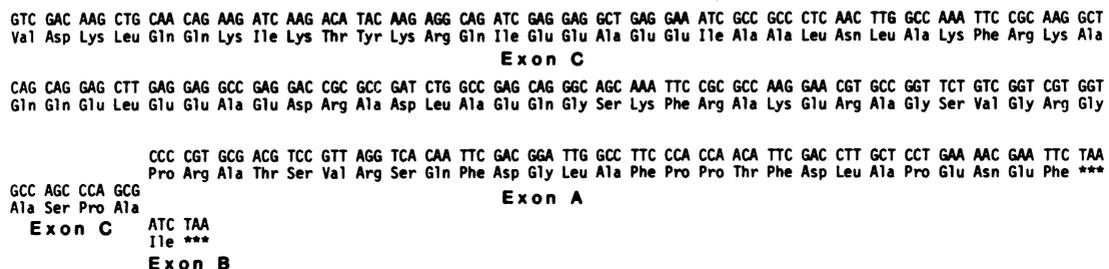


FIG. 5. Deduced amino acid sequence of the MHC polypeptide. Shown is the DNA sequence of exon C beginning at the *Sal* I restriction site (Fig. 2) and the deduced amino acid sequence as well as the proposed reading frames generated by the splicing of exon C to A and exon C to B. Termination codons (TTA) are indicated by ***.

dimers or these antibodies can dissociate existing MHC dimers (19, 20). Thus, if *Drosophila* MHC polypeptides are assembled into homodimeric molecules such as MHC polypeptides from *C. elegans* (21), the unique COOH termini on the *Drosophila* MHC polypeptides could play a key role in recognition for the formation of homodimeric MHC molecules. We also note that the predicted amino acid sequence of the COOH-terminal peptide encoded by exon A contains four proline residues that may contribute to a distinctive tertiary structure and, thus, could play a role in recognition and aggregation.

In summary, the most important conclusion of the present study is that *Drosophila* uses a mechanism of alternate splicing to produce two different MHC isoforms from a single gene. One isoform is expressed at all stages when MHC mRNAs accumulate; the other is specific for the pupal and adult stages of development. Determination of whether the expression of the two MHC isoforms occurs in a tissue-specific manner in pupae and adults awaits further investigation.

Splice Sites and the Mechanisms of Developmentally Regulated Alternate Splicing. There are four sites involved in splice-junction recognition that could conceivably play a role in the developmental regulation of the alternate splicing patterns. These are the splice donor sequences at the 3' ends of exons C and B and the splice acceptor sequences at the 5' ends of exons B and A. Examination of Fig. 3 shows that the two splice donor sequences conform to the consensus sequences at exon-intron junctions (22). The intron-exon junction of exon A, which is not developmentally regulated, is also a consensus sequence. However, the intron-exon junction of exon B, which is used only in PA stages, is nonconsensus.

As shown in Fig. 3, the 3' splice-junction sequence preceding exon B contains the consensus CAG recognition sequence; however, this is preceded by a purine-rich stretch of nucleotides instead of the consensus pyrimidines. The *Drosophila* MHC gene contains six purines in the 10 nucleotides preceding the CAG splice-junction sequences. A list of known splice-junction sequences (22) has been examined to ascertain if other genes have been described that have purine-rich stretches of nucleotides preceding the splice-junction sequence C/TAG. The search was limited to the 10 nucleotides preceding the recognition sequence C/TAG. Among 103 3' acceptor sequences there were seven cases in which >50% of the nucleotides were purines. To our knowledge, none of the exons following these purine-rich sequences is alternatively spliced.

Recently, Falkenthal *et al.* (23) have shown that developmental regulation of alternate splicing occurs for the single-copy *Drosophila* myosin alkali light chain (MLC-ALK) gene, thus producing two isoforms that differ in their COOH-terminal sequences. The mechanisms that regulate the alternate splicing of the MHC and MLC-ALK mRNAs appear to be quite different. For the MLC-ALK case, one mRNA that is present in EL and PA stages contains an exon that has a nonconsensus splice acceptor with a CAG preceded by a purine-rich sequence. A second MLC-ALK mRNA is found only in PA stages that are missing this exon. Thus, although both exons of the MLC-ALK and MHC mRNAs are preceded by similar nonconsensus sequences, they are alternately spliced in different stages of development. We surmise that the molecular mechanisms for alternate splicing of these mRNAs may be quite different.

With regard to the molecular mechanism of developmentally regulated splicing, we presume that there may be a special set of *trans*-acting molecules present in some late pupal and adult muscle tissues that can carry out the nonconsensus C-to-B splice. These molecules are presumably either small nuclear RNAs or proteins (24, 25). Recently, Turner (26) has proposed some interesting ideas on how a

scanning mechanism could function in the alternate splicing of mRNAs. He suggests that certain criteria must be met at the 3' side of the intron. These criteria involve splice-junction recognition, branch-point recognition, or secondary structure within the intron. If these criteria are not met, then the next 3' splice junction is utilized. The localization of alternate splicing to a particular tissue or cell is not an absolute requirement since as long as the recognition factors are present the 3' acceptor sites could be chosen in a stochastic fashion. We have not yet determined if the alternate splicing of the MHC mRNAs occurs in different tissues or cells during the pupal and adult stages of development, but it is clear that an alternative choice of a 3' acceptor sequence occurs. Our evidence does suggest that if unique factors are involved in the recognition of the 3' acceptor site preceding exon B that the synthesis or appearance of these factors is also under developmental control. The recent report of Forbes *et al.* (27) regarding the developmentally regulated transcription of U1 small nuclear RNAs in *Xenopus* makes this an attractive hypothesis. Finally, we believe that if such molecules are expressed in PA muscle tissue, it would be unusual if they did not have more than one primary transcript as their substrate.

C.E.R. is indebted to Drs. Susan Celniker and Morris Burke for many helpful suggestions during the execution of experiments and preparation of the manuscript and to Margaret Chidlaw for help in preparation of the manuscript. This research has been supported by research grants to C.E.R. and N.D. from the National Institutes of Health as well as a generous equipment and supply allowance from the Biology Department of Case Western Reserve University to C.E.R.

- Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. (1982) *Nature (London)* **298**, 240-244.
- Benyajati, C., Spoerel, N., Haymerle, H. & Ashburner, M. (1983) *Cell* **33**, 125-133.
- Nevins, J. R., Blanchard, J. M. & Darnell, J. E. (1980) *J. Mol. Biol.* **144**, 377-386.
- Setzer, D. R., McGrogan, M., Nunberg, J. H. & Schimke, R. T. (1980) *Cell* **22**, 361-370.
- Rogers, J., Early, P., Carter, C., Calame, K., Bond, M., Hood, L. & Wall, R. (1980) *Cell* **20**, 303-312.
- Bernstein, S. I., Mogami, K., Donady, J. & Emerson, C. P. (1983) *Nature (London)* **302**, 393-397.
- Rozek, C. E. & Davidson, N. (1983) *Cell* **32**, 23-34.
- Okayama, H. & Berg, P. (1982) *Mol. Cell Biol.* **2**, 161-170.
- Okayama, H. & Berg, P. (1983) *Mol. Cell Biol.* **3**, 280-289.
- Cohen, S. N., Chang, A. C. Y. & Hsu, L. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 2110-2114.
- Falkenthal, S., Parker, V. P., Mattox, W. M. & Davidson, N. (1984) *Mol. Cell Biol.* **4**, 956-965.
- Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.
- Favaloro, J., Triesman, R. & Kamen, R. (1980) *Methods Enzymol.* **65**, 718-749.
- Mahdavi, V., Periasamy, M. & Nadal-Ginard, B. (1982) *Nature (London)* **297**, 659-664.
- McLachlan, A. D. & Karn, J. (1982) *Nature (London)* **299**, 226-231.
- Sinha, A. M., Umeda, P. K., Kavinsky, C. J., Rajamanickam, C., Hsu, H.-J., Jakovic, S. & Rabinowitz, M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5847-5851.
- Nadal-Ginard, B., Medford, R. M., Nguyen, H. T., Periasamy, M., Wydro, R. M., Hornig, D., Gubits, R., Garfinkel, L. I., Weiczorek, D., Bekesi, E. & Mahdavi, V. (1982) in *Muscle Development: Molecular and Cellular Control*, eds. Pearson, M. L. & Epstein, H. F. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 143-168.
- Lu, R. C., Nyitray, L., Balint, M. & Gergely, J. (1983) *Biophys. J.* **41**, 228a (abstr.).
- Kiehart, D. P., Kaiser, D. A. & Pollard, T. D. (1984) *J. Cell Biol.* **99**, 1015-1023.
- Winkelman, D. A., Lowey, S. & Press, J. L. (1983) *Cell* **34**, 295-306.
- Schatz, F., Garcea, R. L. & Epstein, H. F. (1978) *Cell* **15**, 405-411.
- Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459-472.
- Falkenthal, S., Parker, V. P. & Davidson, N. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 449-453.
- Keller, E. B. & Noon, W. A. (1985) *Nucleic Acids Res.* **13**, 4971-4981.
- Lerner, M. R., Boyle, J. A., Mount, S. M., Walin, S. L. & Steitz, J. A. (1980) *Nature (London)* **283**, 220-224.
- Turner, P. (1985) *Nature (London)* **316**, 105-106.
- Forbes, D. J., Kirschner, M. W., Caput, D., Dahlberg, J. E. & Lund, E. (1984) *Cell* **38**, 681-689.