

Muscle coding sequences and their regulation during myogenesis : cloning of muscle actin cDNA probes

par A. MINTY, M. CARAVATTI, B. ROBERT, Arlette COHEN, P. DAUBAS, A. WEY-DERT, F. GROS, Margaret BUCKINGHAM

*Département de Biologie moléculaire, Institut Pasteur
25, rue du Dr. Roux, 75724 Paris Cedex.*

Summary. For a number of years our group has been mainly interested in the regulation of muscle gene expression during myogenesis. Using primary cultures and cell lines we have tried to find out whether the coding sequences for muscle proteins are already present in an unexpressed form or if there is a transcriptional switch at the onset of differentiation. Metabolic studies on pulse-labelled RNA, together with translation and molecular hybridization experiments have given a certain number of indications. More recently the development of genetic engineering techniques has made it possible to answer these questions directly with probes which are complementary to specific muscle coding sequences. We have identified a plasmid which contains a coding sequence for muscle actin. Other recombinant plasmids are being characterized. Such plasmids, used as probes, will permit us to study the organization and expression of the genes coding for the contractile proteins in muscle cells.

The terminal stages of muscle cell differentiation can be studied in tissue culture either in primary cultures from foetal muscle or in permanent muscle cell lines. Under normal conditions in mammalian cultures, the mononucleated myoblasts multiply until confluence is reached when they begin to fuse together to form multinucleated myotubes increasing in size to appear as spontaneously-contracting branching muscle fibres at the end of the period in culture. The synthesis and accumulation of characteristic muscle proteins such as the acetylcholine receptor, muscle enzymes and structural proteins of the contractile apparatus are initiated at the time the first myotubes appear (Buckingham, 1977). The events are shown in schematic form in figure 1.

Our laboratory is interested in the regulation of muscle gene expression, and two aspects of terminal myogenesis are of particular interest in this context. One concerns the degree of regulatory coordination necessary to produce the different muscle proteins at the onset of differentiation, and in the correct amounts. This is particularly critical in the case of the structural proteins of the sarcomere where stoichiometric requirements are strict. The second point concerns the multigenic families which exist for most of the contractile proteins. In the last few years, it has become evident from studies on the primary sequences of actins from different tissues, for example, that minor differences in the aminoacid sequence exist, thus skeletal

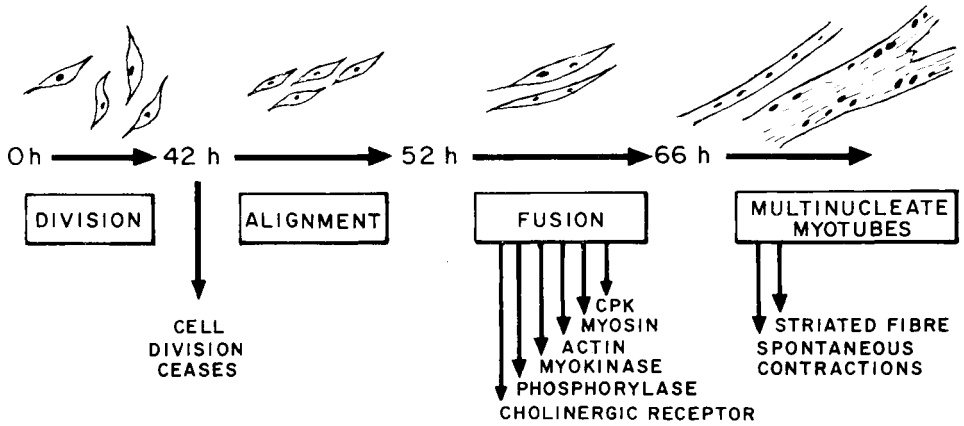


FIG. 1. — Schematic representation of terminal myogenesis in a primary culture of calf muscle cells.

muscle actin differs from the non muscle actins by 25 residues out of 375 and the non muscle β and γ -actins also differ from each other, but only by 3-4 residues (Vandekerkhove and Weber, 1979). There is, thus, a closely related family of actin genes and the same is true for other types of contractile proteins, although the sequence divergence tends to be rather greater. The heavy and light chains of myosin, in addition to their presence in different isozymic forms in different tissues, also appear to be developmentally regulated, so that, for example, the myosin light chain LC_1 has an embryonic isozyme (Whalen *et al.*, 1978). The structure and organization of these genes is of interest both from an evolutionary and a regulatory point of view. A little information is available about the actin gene family. In *Dictyostelium* there appear to be multiple copies of the actin genes, dispersed singly or in clusters in the DNA. Unlike most other examples of structural protein genes, these genes do not contain introns (Kindle and Firtel, 1978; McKeown *et al.*, 1978). More recently, actin genes have been studied in *Drosophila* (Fyrberg *et al.*, 1980), the sea urchin (Merlino *et al.*, 1980) and the chick (Cleveland *et al.*, 1980); in these organisms about 5-7 actin genes have been found, a number which could correspond to the expected number of different actin primary sequences. The genes do contain introns. At present no data is available on other contractile coding sequences and their possible juxtapositions in the genome.

A number of indirect approaches have been employed by us and by other laboratories in an attempt to study muscle gene expression during terminal myogenesis (Buckingham, 1977). Experiments using actinomycin D suggested that mRNA coding for muscle protein accumulates just prior to cell fusion (e.g. Yaffé and Dym, 1973). Pulse labeling and isotopic chase experiments also suggested that a class of mRNAs accumulated as cytoplasmic ribonucleoprotein particles just before fusion, entering the polysomes at the onset of differentiation. Furthermore these mRNAs were present as more rapidly turning-over sequences in ribonucleoprotein particles in the dividing myoblasts (Buckingham *et al.*, 1976). The suggestion, therefore, is that there are changes in the stability and cytoplasmic localization of certain classes of sequences just prior to myotube formation. By this approach it was not possible to identify unambiguously

the nature of these sequences. Another way which also provides information about classes of sequences rather than individual messenger RNAs, is that of molecular hybridization. This is a powerful technique for comparing the total populations of coding sequences in different cells (Minty and Birnie, 1981). Most eukaryotic cells contain 10,000-30,000 different coding sequences ranging from one or two copies to several thousand copies per cell. By homologous and heterologous hybridization between populations of poly A⁺ RNA and their complementary DNAs it is possible to obtain information about the presence or absence of different sequences and their relative abundance in various cell types. Experiments carried out in our laboratory on a muscle cell line indicated that while the majority of sequences are present in both myoblasts and myotubes, some new sequences do appear and their appearance seems to be regulated at a transcriptional level. Other sequences which are less abundant in the myoblast become more abundant in the myotube (Affara *et al.*, 1980). The approach which has permitted the examination of muscle coding sequences is that of *in vitro* translation of muscle cell RNA and comparison of the peptides synthesized *in vitro* with those synthesized by cultures *in vivo*. It has been shown in a number of cell systems that translatable muscle messenger RNAs accumulate in the cytoplasm at the time when their *in vivo* synthesis is increasing (Yablonka and Yaffé, 1977 ; Devlin and Emerson, 1979 ; Daubas *et al.*, 1980). Cytoplasmic modulation of messenger utilization appears also to play a role in fine regulation since a number of imbalances between the levels of *in vivo* and *in vitro* synthesis are observed (Daubas *et al.*, 1981). This approach does not of course provide any information about possible non-translatable forms of mRNA, and also involves assumptions about the efficiency of translation of different messengers.

In order to obtain direct information about the genomic organization and expression of muscle coding sequences it is necessary to purify complementary DNA sequences which can then be used as hybridization probes. This approach has become accessible because of the recent development of recombinant DNA technology which permits the isolation of homogeneous complementary DNA molecules after insertion into a plasmid and bacterial cloning (Glover, 1977). We have employed this technique to isolate muscle coding sequences from heterogeneous fractions of mouse muscle mRNA. The protocol is summarized in figure 2. Poly A⁺ RNA was separated on low salt/sucrose gradients to obtain partially enriched fractions for the myosin light chains, the actins, and the myosin heavy chain mRNAs. Although the proteins are very abundant in muscle (20-60 p. 100 total protein), the level of protein synthesis and of the corresponding mRNA is very much lower (1-5 p. 100) which makes purification of the mRNAs difficult (c/f globin mRNA 50-90 p. 100). Once the mixture of double-stranded cDNAs has been integrated into plasmids (fig. 2) and bacterial clones selected with the expected antibiotic resistance, the major problem is identification of the clones. A preliminary screening was carried out *in situ* by the technique of Grunstein and Hogness (1975), using ³²P-labelled cDNA from the original RNA fraction and from muscle and non-muscle RNA to obtain as much information as possible. Clones were then selected and grown in bulk culture, and the plasmid DNA was purified. This DNA was used in hybridization/translation assays with muscle mRNA, either in the hybrid-arrested translation technique (HART) (Paterson *et al.*, 1977) or the DBM filter technique (Alwine *et al.*, 1979) in order to identify the mRNA species which hybridizes to the recombinant

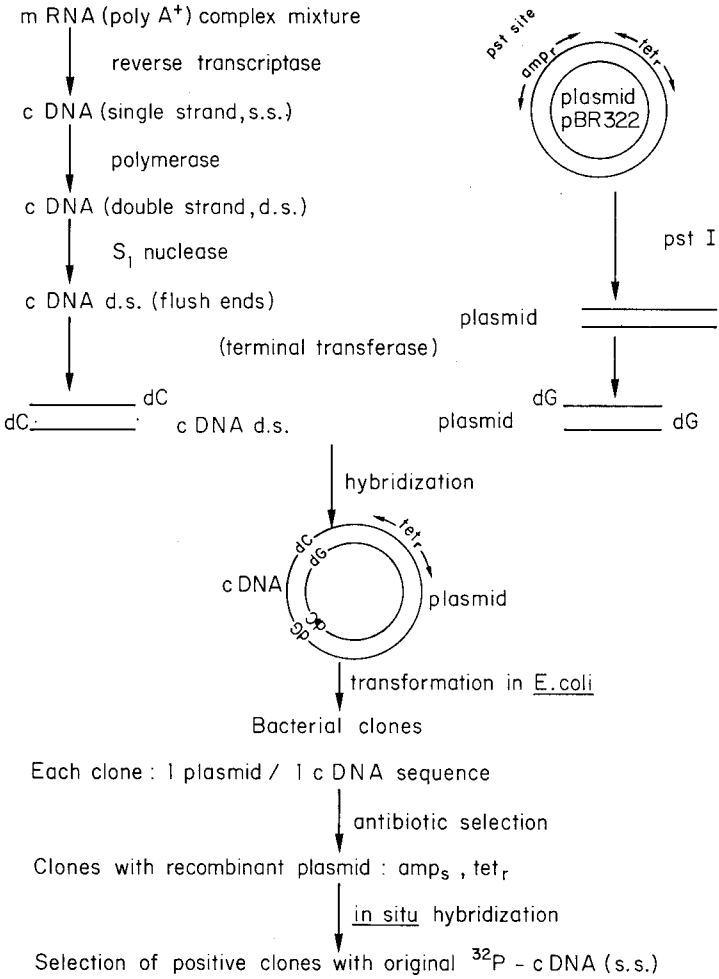


FIG. 2. — Schematic representation of cloning procedure. mRNA : messenger RNA ; cDNA : complementary DNA ; tet_r : resistance to tetracycline ; amp_r : resistance to ampicillin ; amp_s : sensitivity to ampicillin ; pst I : restriction enzyme pst I.

muscle cDNA in the plasmid. The results presented here are those obtained for an actin clone which we have characterized (Minty *et al.*, 1981). The DBM filter technique indicated that this clone cross-hybridizes with muscle and non-muscle actin mRNA. After hybrid-arrested translation and analysis of the products on two-dimensional gels, it appears to affect the translation of α - and γ -actin mRNAs more than that of β -actin mRNA. In order to determine which sequence is incorporated into the plasmid, the inserted cDNA segment was isolated and its DNA sequence analyzed by the method of Maxam and Gilbert (1977). The result is shown in figure 3 where it is clear that the clone contains a sequence corresponding to α -actin. Analysis of this sequence by restriction mapping and partial DNA sequencing indicates that it contains

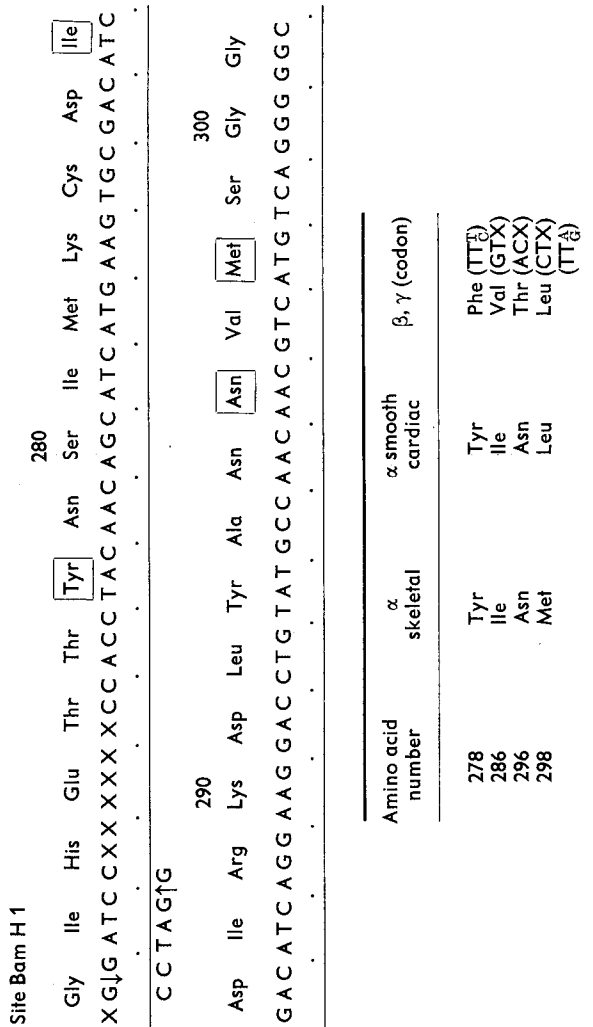


FIG. 3. — Partial nucleotide sequence of actin plasmid, showing the codon assignments characterizing the different isozymes of actin (taken from Vandekerckhove and Weber, 1979).

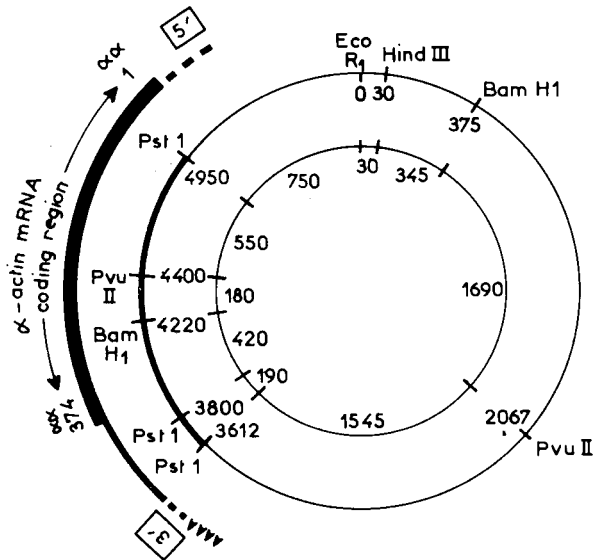


FIG. 4. — Schematic representation of actin plasmid (α A 91). The orientation of the actin coding sequence is shown in relation to the recombinant DNA sequence (—). Restriction sites and their nucleotide distance from the origin of replication or from each other are also shown.

1 350 nucleotides, including part of the non-coding 3' terminal of the mRNA and lacking a short sequence at the 5' coding end of the messenger (fig. 4).

We have begun to use this plasmid as a probe to look at actin gene organization and expression. By the Northern blot technique (Alwine *et al.*, 1979), we have shown that muscle and non-muscle mRNAs have sizes of 1 650 and 2 150 nucleotides, respectively. Furthermore, the 200 nucleotide Pst-Pst fragment (fig. 5), representing the non-coding sequence of α actin mRNA, is specific for this form of actin. The results on the cross-hybridization of different actins are of interest in terms of the evolutionary divergence of the genetic code. The apparent difference between β - and γ -actins is surprising in view of the very close correspondence in amino acid sequence between β - and γ -actin. This might suggest that the α -muscle form has evolved from the non muscle γ -actin.

Preliminary results, using this plasmid to look at mouse genomic DNA suggest that there are a limited number (10-20) of actin sequences in the mammalian genome. Other muscle coding sequences (e.g. myosin LC₁ and LC₃) have also been cloned in our laboratory and experiments are in progress to isolate the corresponding genes and to analyze their expression during myogenesis.

6^e Réunion du groupe Développement I.N.R.A.,
Clermont-Ferrand/Theix, 22-23 mai 1980.

Acknowledgements. — The laboratory is supported by grants from the Délégation Générale à la Recherche Scientifique et Technique, the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, the Commis-

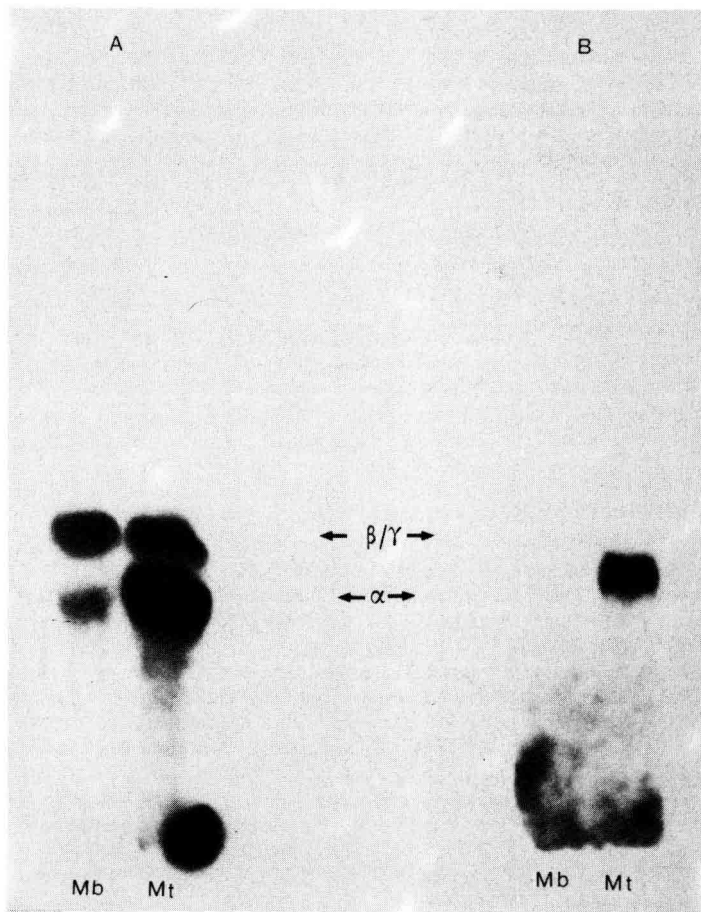


FIG. 5. — A Northern blot (Alwine *et al.*, 1979), showing myoblast and myotube poly A⁺ RNA which has been separated by gel electrophoresis, transferred from the gel to DBM paper and hybridized either with the plasmid α -A 91 (left-hand side) or the α -specific fragment (right-hand side).

sariat à l'Énergie Atomique, the Ligue Nationale Française contre le Cancer, the Fondation pour la Recherche Médicale Française and the Muscular Dystrophy Associations of America.

A. J. M. has been the recipient of fellowships from EMBO and the Muscular Dystrophy Associations of America, M. C. from the Swiss National Research Foundation, and P. D. from the Ligue Nationale Française contre le Cancer.

Résumé. Depuis plusieurs années, l'intérêt de notre groupe s'est porté sur la régulation de l'expression du phénotype musculaire pendant la myogenèse. Dans les cultures primaires et les lignées cellulaires nous avons essayé d'établir si les séquences qui codent pour les protéines musculaires sont déjà présentes sous une forme non exprimée ou s'il a un « switch » transcriptionnel au moment de la différenciation. Les études métaboliques sur les ARN marqués et les expériences de traduction et d'hybridation moléculaire ont donné un certain

nombre d'indications. Le développement de la technologie du génie génétique nous permet maintenant d'aborder la question plus directement avec l'obtention des sondes homogènes complémentaires aux séquences musculaires. Nous avons identifié un plasmide qui renferme une séquence d'ADN codant pour l'actine musculaire. D'autres plasmides recombinants sont en voie d'identification. Ces plasmides, utilisés comme sondes spécifiques, vont nous permettre d'étudier l'organisation et l'expression des gènes codant pour les protéines contractiles dans les cellules musculaires.

References

- AFFARA N., JACQUET M., ROBERT B., BUCKINGHAM M. E., GROS F., 1980. Changes in gene expression during myogenic differentiation : regulation of mRNA sequences expressed during myotube formation. *J. mol. Biol.*, **140**, 441-458.
- ALWINE J. C., KEMP D. J., PARKER B. A., REISER J., RENART J., STARK G. R., WAHL G. M., 1979. Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzoyloxymethyl paper. In WU R., *Methods in enzymology*, vol. **68**, 220-242, Acad. Press.
- BUCKINGHAM M. E., 1977. Muscle protein synthesis and its control during the differentiation of skeletal muscle cells *in vitro*. In PAUL J., *International Review of biochemistry. Biochemistry of cell differentiation II*, vol. **15**, 269-332. Publ. MTP Press.
- BUCKINGHAM M. E., COHEN A., GROS F., 1976. Cytoplasmic distribution of pulse labelled poly A containing RNA, particularly 26 S RNA, during myoblast growth and differentiation, *J. mol. Biol.*, **103**, 611-626.
- CLEVELAND D. W., LOPATA M. A., MacDONALD R. J., COWAN N. J., RUTLER, W. J., KIRSCHNER, M. W., 1980. Number and evolutionary conservation of α and β tubulin and cytoplasmic β and γ actin genes using specific cloned cDNA probes. *Cell*, **20**, 95-105.
- DAUBAS P., CAPUT D., BUCKINGHAM M., GROS F., 1981. A comparison between the synthesis of contractile proteins and the accumulation of their translatable mRNAs during calf myoblast differentiation. *Develop. Biol.* (in press).
- DEVLIN R. B., EMERSON C. P., 1979. Co-ordinate accumulation of contractile protein mRNAs during myoblast differentiation. *Develop. Biol.*, **69**, 202-216.
- FYRBERG E. A., KINDLE K. L., DAVIDSON N., SODJA A., 1980. The actin genes of drosophila : a dispersed multigene family. *Cell*, **19**, 365-378.
- GLOVER D. M., 1977. Gene cloning : A new approach to understanding relationships between DNA sequences. In PAUL J., *International Review of Biochemistry. Biochemistry of cell differentiation II*, vol. **15**, 269-332, Publ. MTP Press.
- GRUNSTEIN M., HOGNESS, D. S., 1975. Colony hybridization : a method for the isolation of cloned DNAs that contain a specific gene. *Proc. nat. Acad. Sci., USA*, **72**, 3961-3966.
- KINDLE K. L., FIRTEL R. A., 1978. Identification and analysis of *Dictyostelium actin genes*, a family of moderately repeated genes. *Cell*, **15**, 763-778.
- MAXAM A., GILBERT W., 1977. A new method for sequencing DNA. *Proc. nat. Acad. Sci., USA*, **74**, 560-564.
- McKEOWN M., TAYLOR W. C., KINDLE K. L., FIRTEL R. A., BENDER N., DAVIDSON N., 1978. Multiple actin genes in *Dictyostelium*. *Cell*, **15**, 789-800.
- MERLINO G. T., WATER R. D., CHAMBERLAIN J. P., JACKSON D. A., EL-GWELY M. R., KLEINSMITH L. J., 1980. Cloning of sea urchin actin gene sequences for use in studying the regulation of actin gene transcription. *Proc. nat. Acad. Sci., USA*, **77**, 765-769.
- MINTY A. J., BIRNIE G. D., 1981. Messenger RNA populations in eucaryotic cells — evidence from recent nucleic acid hybridization experiments bearing on the extent and control of differential gene expression. In M. BUCKINGHAM, *Cellular regulation*, Vol. **III**, *Development and differentiation* C. R. C. Press (in press).
- MINTY A. J., CARAVATTI M., ROBERT B., COHEN A., DAUBAS P., WEYDERT A., GROS F., BUCKINGHAM M., 1980. Construction and characterization of a recombinant plasmid molecule containing a complementary DNA transcript of mouse α -actin mRNA. *J. biol. Chem.* (in press).

- PATERSON B. M., ROBERTS B. E., KUFF E. L., 1977. Structural gene identification and mapping by DNA-mRNA hybrid-arrested cell-free translation. *Proc. nat. Acad. Sci., USA*, **74**, 4370-4374.
- VANDEKERKHOVE J., WEBER K., 1979. The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine fast skeletal muscle and rabbit slow skeletal muscle. *Differentiation*, **14**, 123-133.
- WHALEN R. G., BUTLER-BROWNE G. S., GROS F., 1978. Identification of a novel form of myosin light chain present in embryonic muscle tissue and cultured muscle cells. *J. mol. Biol.*, **126**, 415-431.
- YABLONKA Z., YAFFÉ, D., 1977. Synthesis of myosin light chain and accumulation of mRNA coding for light chain like polypeptides in differentiating muscle cultures. *Differentiation*, **8**, 133-143.
- YAFFÉ D., DYM H., 1973. Gene expression during differentiation of contractile muscle fibers, *Cold Spring Harbor Symp. Quant. Biol.*, **37**, 543-547.
-