Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 6807–6810, October 1985
Biochemistry

Cloning and characterization of a nonmuscle myosin heavy chain cDNA

(motility/*Dictyostelium discoideum*/expression vector)

ARTURO DELOZANNE*, MINDY LEWIS†, JAMES A. SPUDICH*, AND LESLIE A. LEINWAND†

Departments of †Microbiology and Immunology and of Genetics, Albert Einstein College of Medicine, New York, NY 10461; and *Department of Cell Biology, Stanford University School of Medicine, Stanford, CA 94305

Communicated by Frank Lilly, June 24, 1985

ABSTRACT **Despite many biochemical and structural similarities between muscle and nonmuscle myosins, their genes appear to have completely diverged, since muscle myosin molecular clones will not hybridize to RNA from nonmuscle sources. Here we report the isolation and characterization of a partial myosin heavy chain (MHC) cDNA clone from the slime mold *Dictyostelium discoideum*. We have isolated this clone from a λgt11 expression cDNA library by antibody screening. In contrast to the highly conserved sarcomeric muscle MHC multigene families in other organisms, there appears to be only one gene encoding MHC in the *Dictyostelium* genome. The cloned portion of this gene does not hybridize to the genomic DNAs of other eukaryotic organisms. Analysis of the predicted amino acid sequence of the partial *Dictyostelium* MHC clone shows that while there is no sequence homology to known striated muscle MHCs, the structure- and coiled-coil-forming capacities have been conserved.**

Myosins are ubiquitous in eukaryotes, where they convert chemical energy into mechanical force through the hydrolysis of ATP. This generation of force manifests itself in muscle cells by contraction and in nonmuscle cells (e.g., macrophages and fibroblasts) by many fundamental cellular processes such as phagocytosis and cell division. The myosin molecule consists of two heavy chains (about 200 kDa each) and two pairs of light chains (15–20 kDa each). The carboxyl-terminal portion of the heavy chain forms an α-helical tail, and the globular amino-terminal portion contains the ATPase activity and the binding sites for the myosin light chains and actin. An intriguing aspect of myosin structure and function is that there are many biochemical similarities between muscle and nonmuscle myosins, and yet they appear to have completely diverged at the sequence level. Myosin heavy chain (MHC) cDNA probes from skeletal or cardiac muscle will hybridize to each other but will not cross-hybridize with RNA from smooth muscle or nonmuscle tissue (1).

The availability of cloned DNA sequences corresponding to sarcomeric MHCs has been invaluable in elucidating features of their primary structure and regulated expression. Sarcomeric MHCs are encoded by highly conserved multigene families of at least 10 members (2–5). Rat MHC genes will cross-hybridize with DNA from such diverse organisms as humans, goldfish, and sea urchins (6). Protein and nucleic acid studies have shown that within striated muscle, there are multiple tissue- and developmental-specific MHC forms that are very closely related, but distinct (e.g., see ref. 7). It is presumed that these various forms are functionally significant. Molecular genetic analysis of different MHC forms has begun to identify those areas of the molecule that are type specific and those that are more highly conserved. Similar approaches to the study of nonmuscle MHC have not been possible due to the inability to isolate such molecular clones. Here we present the isolation of a partial MHC cDNA clone from the slime mold *Dictyostelium discoideum*. We have isolated this clone from a λgt11 expression cDNA library by screening with a polyclonal anti-*Dictyostelium* myosin antibody. Preliminary characterization of the *Dictyostelium* MHC gene points to fundamental differences at the gene level between muscle and nonmuscle motility.

MATERIALS AND METHODS

Antibody Screening of *D. discoideum* λgt11 Expression cDNA Library. The λgt11 cDNA library containing 10^5 recombinants (generous gift of S. Cohen and H. Lodish, Massachusetts Institute of Technology) was prepared from *Dictyostelium* AX3 cells at 8 hr of development. Four library equivalents were plated onto *Escherichia coli* strain Y1090 at 3×10^4 plaque-forming units per plate, grown for 4 hr at 42°C, induced with 0.01 mM isopropyl β-D-thiogalactoside for 2 hr, and screened with a polyclonal rabbit anti-*Dictyostelium* myosin antibody that was diluted 1:100. Screening and visualization of plaques was carried out as described (8).

***E. coli* Fusion Protein Detection in Lysogens by Blot Analysis.** Lysogens of the different positive clones were made in *E. coli* strain Y1089. After induction, lysates were prepared as described (8) and electrophoresed along with controls on two 4% stacking/7.5% running polyacrylamide gels. Proteins from one gel were electrophoretically transferred to nitrocellulose and the filter was processed with anti-*Dictyostelium* myosin antibody, using the horseradish peroxidase (HRP) technique for visualization of antigen–antibody complexes (9). Proteins from the other gel were visualized by staining with Coomassie brilliant blue.

RNA and DNA Blot Analysis. *Dictyostelium* AX3 cellular RNA was extracted by the guanidine·HCl/hot phenol procedure (10). RNA was run on a denaturing 1% agarose gel and transferred to nitrocellulose as previously described (11). Filters were prehybridized and hybridized at 65°C in 0.75 M NaCl/0.075 M sodium citrate/0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone with denatured salmon sperm DNA at 100 μg/ml. Plasmid DNA probes were radioactively labeled by nick-translation (12) and included in hybridizations at 2×10^5 cpm/ml. Filters were washed in 0.30 M NaCl/0.03 M sodium citrate/0.2% $NaDodSO_4$ at 65°C. *Dictyostelium* MHC plasmid DNA probes were obtained from inserts of λgt11 cDNA clones by excision with *Eco*RI and subcloning in pBR322.

Genomic DNA from various organisms was cut with *Eco*RI and electrophoresed on 0.7% agarose gels. After transfer to nitrocellulose (13), filters were prehybridized and hybridized as described above.

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; kb, kilobase(s).

DNA Sequence Analysis. Sequencing of the λgt11 cDNA inserts subcloned in pBR322 and phage M13 was performed by base-specific chemical cleavage (14) and dideoxy chain termination methods (15), respectively.

RESULTS

The λgt11 expression system has been used to isolate a number of eukaryotic molecular clones (8). This vector expresses inserted "foreign" DNA as a fusion protein with β-galactosidase. A cDNA library can be screened with an antibody to isolate molecular clones for which no nucleic acid probe is available. In 1×10^5 recombinant *Dictyostelium* λgt11 cDNA clones screened with a polyclonal anti-*Dictyostelium* myosin antibody, 40 positives were detected. Ten clones were subjected to secondary screening. Examples of a primary and secondary screening are shown in Fig. 1 *A* and *B*. To demonstrate that the antibody used in the screening selected only the expected fusion proteins, blot analysis was carried out on proteins isolated from *E. coli* cells lysogenic for the various λgt11 clones. As can be seen in Fig. 1 *C* and *D*, only the recombinant phage lysogens synthesize proteins that react with anti-*Dictyostelium* myosin antibody. The fusion proteins are larger than native β-galactosidase (114 kDa) and the increase in size is proportional to the size of the inserts. The specificity of the antibody is shown in Fig. 1 *E* and *F*, where the antibody recognizes only MHC from a total *Dictyostelium* cell lysate.

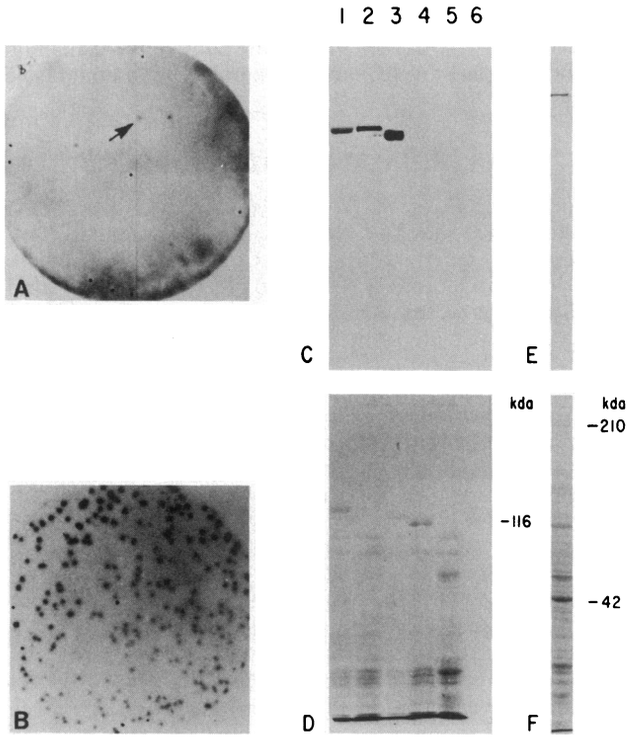

FIG. 1. Antibody detection of *Dictyostelium* myosin-β-galactosidase fusion proteins in λgt11 clones. (*A* and *B*) Primary and secondary antibody screening of a *Dictyostelium* λgt11 cDNA library with a rabbit anti-*Dictyostelium* myosin. Screening and visualization of plaques was carried out as described (8). (*A*) Primary screening of 3×10^4 plaques; (*B*) secondary screening of the plaque indicated by the arrow in *A*. (*C* and *D*) Protein blot analysis of λgt11 clones. Lysogens of the different positive clones were made in *E. coli* strain Y1089. After induction, lysates were prepared as described (8) and electrophoresed along with controls on two 4% stacking/7.5% running polyacrylamide gels. (*C*) Proteins from one gel were electrophoretically transferred to nitrocellulose and the filter was processed with anti-*Dictyostelium* myosin antibody, using the horseradish peroxidase technique for visualization of antigen–antibody complexes (16). (*D*) The duplicate gel for *C* was stained with Coomassie blue. The horizontal line at 116 kDa indicates the size of β-galactosidase. Lane 1, *Dictyostelium* MHC-21 *E. coli* Y1089; lane 2, *Dictyostelium* MHC-15 *E. coli* Y1089; lane 3, *Dictyostelium* MHC-1 *E. coli* Y1089; lane 4, *E. coli* Y1089 (λgt11); lane 5, *E. coli* Y1089; lane 6, β-galactosidase. (*E* and *F*) Protein blot analysis of a *Dictyostelium* total cell lysate electrophoresed on a 10% polyacrylamide gel and processed as for *C*. (*E*) Horseradish peroxidase immunostain. (*F*) Coomassie blue-stained gel. Myosin and actin are indicated by the horizontal lines at 210 and 42 kDa, respectively.

Inserts from four λgt11 clones were excised with *Eco*RI and found to be small, ranging from 100 to 310 base pairs. They were subcloned in the *Eco*RI site of pBR322 and M13mp8 for further characterization. Hybridization of the inserts to size-fractionated *Dictyostelium* RNA identified a single RNA species of 7.1 kilobases (kb) (Fig. 2, lane 1), which is close to the predicted length of the unusually long MHC mRNA. The protein is 200 kDa and would require at least 6 kb of a mRNA to encode it. This size is characteristic of MHC mRNA since the average size of eukaryotic mRNA is about 2.0 kb (17). An actin probe hybridized to this same preparation identified the two expected actin transcripts of 1.35 and 1.25 kb (18), confirming the integrity of the RNA (Fig. 2, lane 2).

To further identify this cloned sequence as myosin, DNA sequencing was carried out to examine features of the polypeptide encoded by the cDNA clones. In the absence of nonmuscle myosin protein sequence, other structural features can be examined. Two of the structural features that are characteristic of myosin are an α-helical propensity and a periodicity of uncharged residues. The four cDNA clones represent overlapping portions of a single sequence. One end of the sequence is A-rich, suggesting that it may be the 3′ end (Fig. 3). Only one long open reading frame can be identified in the sequence of this region. As expected, this reading frame is in frame with the translation of β-galactosidase (data not shown). The largest clone contains 240 nucleotides of protein-encoding sequence in addition to 70 nucleotides of the 3′ untranslated region (Fig. 3). There is no homology to previously described muscle MHC sequences (data not shown) (19). An examination of the recently published

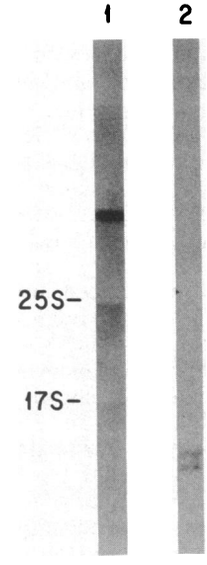

FIG. 2. Blot analysis of *Dictyostelium* cDNA clones. Lane 1, hybridization of DNA from clone pDdMHC-15 with total *Dictyostelium* RNA; lane 2, hybridization of the same RNA with an actin probe, pCERF DRp14, provided by R. Firtel (University of California, San Diego).

```
                                                          30
CTC GTT GTC GAA AAA CTC GAA ACA GAT TAC
 L   V   V   E   K   L   E   T   D   Y
                                                          60
AAG AGA GCC AAG AAA GAA GCT GCT GAT GAA
 K   R   A   K   K   E   A   A   D   E
                                                          90
CAA CAA CAA CGT CTT ACT GTT GAA AAC GAT
 Q   Q   Q   R   L   T   V   E   N   D
                                                         120
CTC CGT AAA CAC CTC AGT GAA ATC TCA TTA
 L   R   K   H   L   S   E   I   S   L
                                                         150
CTC AAA GAT GCC ATT GAT AAG TTA CAA CGT
 L   K   D   A   I   D   K   L   Q   R
                                                         180
GAT CAC GAT AAG ACC AAA CGT GAA TTG GAA
 D   H   D   K   T   K   R   E   L   E
                                                         210
ACA GAA ACT GCC AGC AAA ATC GAA ATG CAA
 T   E   T   A   S   K   I   E   M   Q
                                                         240
AGA AAG ATG GCC GAT TTC TTT GGT GGT TTC
 R   K   M   A   D   F   F   G   G   F
                                                         270
AAA GCT TAA ATC AAT TTG ATT TCT TCT TAA
 K   A   .
                                                         300
TTC AAT GCT AAA AAA AAA AAT ATT ATA AAA

AAA AAA ACA G
```

FIG. 3. Nucleotide sequence and predicted amino acid sequence (in single-letter code) of the combined *Dictyostelium* MHC clones.

peptide sequence of the tail of *Acanthamoeba* myosin II also failed to show any homology (data not shown) (16).

Computer-assisted analysis of the cDNA sequence was used to make predictions about the structure of the encoded polypeptide and its similarity to muscle MHC. Analysis of the sequence by the method of Chou and Fasman (20) shows that it has considerable potential to form an α-helix throughout most of its length (Fig. 4). The average P_α value is 1.14. The average value for the rabbit muscle myosin subfragment 2 is 1.18 (21). Any value above 1.03 is considered to have a high probability for α-helix formation (20). The last several amino acids of the predicted *Dictyostelium* sequence do not have α-helical propensity. The tailpiece of nematode body wall MHC also has such a feature (22). The structure predicted by such analyses for *Dictyostelium* MHC is thus consistent with the known conformation of the tail region of muscle myosin (23, 24).

Another parameter of the *Dictyostelium* sequence that was

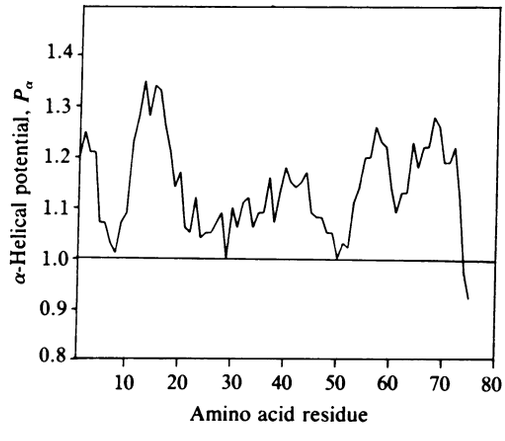

FIG. 4. Propensity for formation of α-helix by the *Dictyostelium* MHC, as analyzed by the method of Chou and Fasman (20).

Table 1. Periodicity of uncharged residues in the seven-residue repeat unit of various MHCs

MHC source	% uncharged residues in each position of the repeat unit						
	1st	2nd	3rd	4th	5th	6th	7th
Dictyostelium	*75*	41	33	*72*	54	54	36
Nematode muscle	*86*	37	40	*86*	52	52	58
Rabbit muscle	*95*	36	38	*87*	47	49	48
Rat muscle	*85*	34	39	*87*	50	52	50

analyzed was the periodicity of uncharged residues (Table 1). The tail regions of nematode and rabbit MHCs have been shown to be α-helical coiled-coils (22, 24). This coiled-coil conformation is presumably stabilized by hydrophobic interactions between neighboring nonpolar amino acid side chains. Examination of muscle myosin from various sources has demonstrated the existence of a seven-residue repeat unit in which a high percentage of first and fourth positions are occupied by hydrophobic amino acids (19). A comparison of the distribution of uncharged amino acids of the *Dictyostelium* and muscle MHC protein sequences (Table 1) shows that the nonmuscle myosin clearly has the characteristic myosin periodicity.

To determine how many MHC genes are present in the *Dictyostelium* genome, each subclone was hybridized to *Dictyostelium* DNA digested with several endonucleases. Fig. 5*A* shows that the largest cDNA clone hybridizes to a single *Eco*RI, *Hin*dIII, or *Kpn* I genomic fragment and hybridizes to two *Bcl* I genomic fragments. There is a single *Bcl* I site in the insert. It has been possible to construct a restriction map of the genomic region encompassing the MHC cDNA clone. It appears from this analysis that there is a single MHC gene in *Dictyostelium*. The degree of evolutionary conservation of the *Dictyostelium* MHC gene was examined by hybridizing the *Dictyostelium* cDNA insert to *Eco*RI digests of DNA isolated from a wide variety of eukaryotic organisms. While a single *Eco*RI fragment from *Dictyostelium* showed an intense hybridization signal, no cross-hybridization was observed with DNA from any of the other species examined (Fig. 5*B*). In contrast, a 600-base-pair fragment of the 3′ end of a human sarcomeric MHC cDNA clone (unpublished results) cross-hybridized extensively with DNA from other organisms from mouse to goldfish, but not *Dictyostelium* (Fig. 5*C*).

DISCUSSION

We have reported here the molecular cloning of the cDNA for a nonmuscle MHC. These clones are thought to encode MHC for the following reasons: (*i*) They encode a protein that is recognized by an antibody highly specific for *Dictyostelium* MHC. (*ii*) The cDNA clones hybridize to a single RNA species of 7.1 kb. (*iii*) Computer analyses predict a strong α-helical character and, more importantly, a sequence that can form a coiled-coil.

The gene number and sequence conservation of the *Dictyostelium* MHC gene appear to be different from those of previously described muscle MHC genes. Most sarcomeric MHC genes are encoded by multigene families (2–5) with the exception of *Drosophila*, which has one gene encoding three different MHCs (25). The possibility of only one MHC gene in *Dictyostelium* is very interesting since the organism contains 17 actin genes, which cross-hybridize to the DNA from other eukaryotic organisms (26). We have not excluded that there are other MHC genes in *Dictyostelium* with different carboxyl-terminal portions, but this would also be different from the sarcomeric MHC genes, whose tail portions are highly conserved (19).

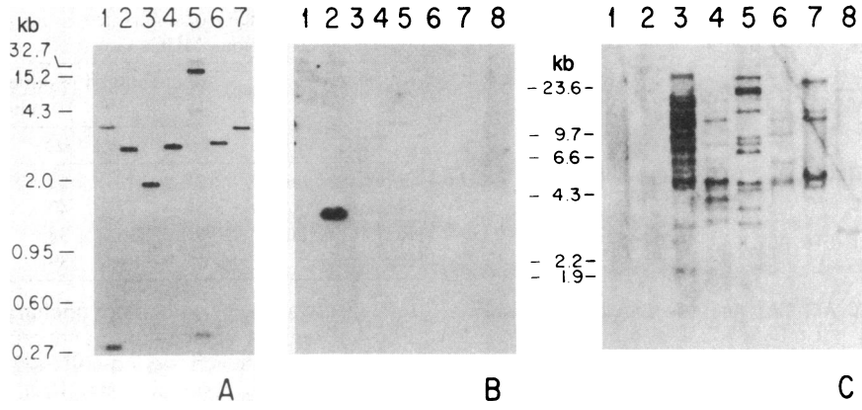

FIG. 5. Hybridization of MHC cDNA clones to genomic DNA. Radioactively labeled probes were gel-purified inserts from pDdMHC-15 for *A* and *B* and pSMHCA for *C*. pSMHCA is a 3'-proximal fragment of an adult human skeletal muscle cDNA clone (unpublished results). (*A*) *Dictyostelium* genomic DNA cut with the following enzymes: lane 1, *Eco*RI + *Hin*dIII; lane 2, *Bcl* I + *Hin*dIII; lane 3, *Kpn* I + *Hin*dIII; lane 4, *Hin*dIII; lane 5, *Bcl* I; lane 6, *Kpn* I; and lane 7, *Eco*RI. The digests were probed with the insert from pDdMHC-15. (*B*) Genomic DNAs from eight different organisms cut with *Eco*RI and probed with the insert from pDdMHC-15. Identity of genomic DNAs for *B* and *C*: lane 1, sea urchin (*Lytechinus pictus*); lane 2, *Dictyostelium*; lane 3, goldfish (*Carassius auratus*); lane 4, amphibian (*Xenopus*); lane 5, chicken; lane 6, mouse; lane 7, rat; and lane 8, human. (*C*) The same genomic DNAs probed with the insert from pSMHCA.

The conservation of the coiled-coil structure between muscle and nonmuscle MHC in the absence of sequence homology raises an interesting question. Is there any sequence-specific function to the carboxyl-terminal portion of the MHC or is it purely the coiled-coil structure that is important? If the latter were correct the sequence of the MHC tail from various species might be expected to have diverged significantly. There is, however, a high degree of carboxyl-terminal sequence conservation among rabbit, chicken, and rat striated muscle MHCs (19). It seems likely that the nonhomology between the muscle MHC genes and the *Dictyostelium* sequence reflects functional diversity of myosin-mediated motility. For example, the carboxyl-terminal portion of the MHC in nonmuscle cells may play an important role in control of assembly–disassembly of myosin filaments (27), whereas in muscle the filaments are stable structures.

It should now be possible to define the regions of the myosin molecule that are important for its function, by introducing alterations into the cloned MHC DNA, expressing it, and assaying for its various functions. In this manner, we hope to map the functionally important regions of both muscle and nonmuscle myosins and to better define the molecular aspects of motility.

Note Added in Proof. We have recently isolated a 13-kb genomic segment containing the *Dictyostelium* MHC gene. Genomic Southern blots using this larger segment as a probe confirm the existence of a single *Dictyostelium* MHC gene.

We acknowledge the excellent secretarial assistance of D. Jackson and L. Vitagliano. This research was supported by the National Institutes of Health (Grants GM 30387 and GM 33289 to J.A.S. and GM 29090 to L.A.L.) and the American Heart Association. L.A.L. is an American Heart Association Established Investigator. A.D. is supported by National Institutes of Health Grant GM 33289 and M.L. by National Institutes of Health Training Grant 5T 32GM07128.

1. Medford, R., Wydro, R., Nguyen, H. & Nadal-Ginard, B. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5749–5753.
2. Wydro, R., Nguyen, H., Gubits, R. & Nadal-Ginard, B. (1983) *J. Biol. Chem.* **258**, 670–678.
3. Robbins, J., Freyer, G., Chisholm, D. & Gilliam, T. C. (1981) *J. Biol. Chem.* **256**, 549–556.
4. Leinwand, L. A., Saez, L., McNally, E. & Nadal-Ginard, B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3716–3720.
5. Sinha, A. M., Umeda, P., Kavinsky, C., Rajamanickam, C., Hsu, H., Jakovcic, S. & Rabinowitz, M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5847–5851.
6. Nguyen, H., Gubits, R., Wydro, R. & Nadal-Ginard, B. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5230–5234.
7. Whalen, R. G., Schwartz, K., Bouveret, P. & Gros, F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5197–5200.
8. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1194–1198.
9. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 194–195.
11. Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M. & Darnell, J. E., Jr. (1981) *Cell* **23**, 731–739.
12. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
13. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
14. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Coté, G., Robinson, E., Appella, E. & Korn, E. (1984) *J. Biol. Chem.* **259**, 12781–12787.
17. Lewin, B. (1975) *Cell* **4**, 77–93.
18. Kindle, K. L. & Firtel, R. A. (1978) *Cell* **15**, 763–778.
19. Kavinsky, C., Umeda, P., Sinha, A., Elzinga, M., Tong, S., Zak, R., Jakovcic, S. & Rabinowitz, M. (1983) *J. Biol. Chem.* **258**, 5196–5205.
20. Chou, P. & Fasman, G. (1978) *Adv. Enzymol.* **47**, 45–148.
21. Capony, J. P. & Elzinga, M. (1981) *Biophys. J.* **33**, 148 (abstr.).
22. McLachlan, A. D. & Karn, J. (1982) *Nature (London)* **229**, 226–231.
23. Lowey, S. & Cohen, C. (1962) *J. Mol. Biol.* **4**, 293–308.
24. Elzinga, M. & Truss, B. (1980) in *Methods in Peptide and Protein Sequence Analysis*, ed. Birr, Chr. (Elsevier, Amsterdam), pp. 213–224.
25. Rozek, C. E. & Davidson, N. (1983) *Cell* **32**, 23–34.
26. Firtel, R. A. (1981) *Cell* **24**, 6–7.
27. Kuczmarski, E. R. & Spudich, J. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7292–7296.