1085-9195/99/30/413–435/$15.75

A Potentially Exhaustive Screening Strategy Reveals Two Novel Divergent Myosins in *Dictyostelium*

Eva C. Schwarz, Heidrun Geissler, and Thierry Soldati*

Department of Molecular Cell Research, Max-Planck-Institute for Medical Research, Jahnstrasse 29, 69120 Heidelberg, Germany

ABSTRACT

In recent years, the myosin superfamily has kept expanding at an explosive rate, but the understanding of their complex functions has been lagging. Therefore, *Dictyostelium discoideum*, a genetically and biochemically tractable eukaryotic amoeba, appears as a powerful model organism to investigate the involvement of the actomyosin cytoskeleton in a variety of cellular tasks. Because of the relatively high degree of functional redundancy, such studies would be greatly facilitated by the prior knowledge of the whole myosin repertoire in this organism. Here, we present a strategy based on PCR amplification using degenerate primers and followed by negative hybridization screening which led to the potentially exhaustive identification of members of the myosin family in *D. discoideum*. Two novel myosins were identified and their genetic loci mapped by hybridization to an ordered YAC library. Preliminary inspection of *myoK* and *myoM* sequences revealed that, despite carrying most of the hallmarks of myosin motors, both molecules harbor features surprisingly divergent from most known myosins.

*Author to whom all correspondence and reprint requests should be addressed.

Index Entries: Cytoskeleton; myosin; PCR; *Dictyostelium discoideum.*

INTRODUCTION

Myosins have been described in a wide variety of organisms, from yeast to man and from amoeba to plants. Indeed, expression of actin and myosins appears to be a hallmark of all eukaryotes. This rapidly expanding protein superfamily, the founder of which is the filament-forming myosin II, is divided in at least 15 classes *(1–3)*. Even a eukaryote of relatively low complexity such as *Saccharomyces cerevisiae* expresses five myosins from three different classes *(4–7)*.

Despite the presence of more than a hundred myosin sequences in the databases, relatively little is known about their cellular functions. Nevertheless, evidence is slowly accumulating about their involvement in a variety of cellular processes such as motility, cytokinesis, phagocytosis, endocytosis, polarized secretion and exocytosis, organelle movement, and mRNA transport *(1)*. In addition, recent data indicate that myosins play active roles in signal transduction pathways *(8)*.

All the myosins identified so far share a common tripartite modular structure comprised of the head or motor domain, the neck domain, and the tail domain. Directed movement along actin filaments is performed by the motor domain carrying both binding sites for ATP and actin. Despite overall structure and sequence conservation in this region, the head domains of an increasing number of myosins show either variation in otherwise extremely conserved motifs (e.g., of the ATP-binding site) or presence of insertions in structural domains usually conserved in length (surface loop 1 and loop 2).

Dictyostelium discoideum is a genetically and biochemically tractable model organism. Its small haploid genome of 3.4×10^7 bp contains only few and usually small introns and gene targeting by homologous recombination is efficient *(9)*. Axenic lab strains give easy access to milligram quantities of proteins. Recent data led to a revisiting of the place of *D. discoideum* in evolution. Protein-based phylogenetic analyses suggest that it diverged after the yeasts and before the metazoan radiation *(10)*. Whereas the single cell, immotile *S. cerevisiae* expresses five myosins it is estimated that the genome of a mammalian cell might contain up to three dozen of such motors.

Owing to its intermediate position in evolution, *D. discoideum* also reflects its intermediate complexity in the makeup of its cytoskeletal composition and might total just over a dozen myosins (Table 1). Previous studies using a combination of approaches led to the identification of 10 different myosins including the classical (conventional) myosin II *(mhcA)*, at least six members of the class I, called *myoA* to *myoF*, and two myosins of high molecular weight, *myoI*, and *myoJ*. Only three unconventional myosins, MyoB, MyoC, and MyoD have been directly purified and studied biochemically *(11)*. Gene disruption experiments indicated that myosins of the class I fulfill various crucial tasks within the cell *(12–17)* but also revealed a high degree of functional redundancy.

In order to dissect the individual as well as the overlapping functions of this family of motors in any model organism, it appears essential to first acquire knowledge of the complete repertoire of myosins. Also, identification in *D. discoideum* of myosins that are close cousins of those recently discovered in extremely complex cellular systems such as the hair cells of the mammalian cochlea, will tremendously ease their molecular functional dissection. Such an example is represented by MyoI, a myosin of class VII recently involved in phagocytosis *(18,19)*. Mutations in mammalian myosin VII are associated with deafness and deaf/blindness syndromes in mammals *(20)*, the blindness being potentially correlated with impaired phagocytosis in retinal pigmented epithelial cells *(1,21)*.

Data presented in previous reports support the presumption that the *D. discoideum* repertoire of myosins is not yet complete. Conserved head regions have been used as probes to hybridise an ordered array of yeast artificial chromosomes (YACs) covering the genome of *D. discoideum* and revealed thirteen potential myosin loci *(22)*, but no myosin sequence has yet been retrieved from the YACs bearing *myoG*, *myoK*, and *myoL*. Southern blot analysis of *D. discoideum* genomic DNA with a probe derived from the tail of *myoJ* detected a second unidentified locus *(23)*. The aim of the present investigation was to complete this catalog by following a thorough and exhaustive screening strategy. We report here of the cloning of one myosin corresponding to the previously mapped *myoK* locus as well as the identification and cloning of a novel myosin *myoM*, expanding the repertoire to a potential total of 14 members.

Table 1

The Myosin Family of *Dictyostelium discoideum*

Myosin	Size	Genomic iocus	Reference	Accession number
Class I				
myoA	113 kD	chromosome 3 3870 kb	*(47)*	em:S73909 p22467
myoB	124 kD	chromosome 5 1870 kb	*(37,48)*	em:m26037 p34092
myoC	135 kD	chromosome 2 5610 kb	*(49)*	em:l35323 p42522
myoD	124 kD	chromosome 2 860 kb	*(11)*	em:l16509 p34109
myoE	113 kD	chromosome 5 1150 kb	*(50)*	em:l06805 q03479
myoF	113 kD (?)	chromosome 5 2070 kb	*(51)*	em:l35319
myoK	94 kD	chromosome 5 3170 kb	this report	AF090534
Class II				
mhcA	240 kD	chromosome 4 5280 kb	*(52,53)*	em:m14628 p08799
Class VII or X				
myoI	260 kD (?)	chromosome 5 2330 kb	*(19,22)*	em:l35321
Class XI or V				
myoJ	258 kD	chromosome 2 3810 kb	*(23,54)*	em:u42409
Not classified yet				
myoH	? fragment	chromosome 5 5000 kb	*(22)*	em:l35320
myoM	195 kD	chromosome 6 1850 kb	this report	AF090533

The table presents all the myosin sequences from *D. discoideum* known so far. The apparent molecular weights given derive from the publications cited (including myoF and myoI, where the question marks indicate that full length sequences are not available yet, but sizes were approximated from the length of the mRNAs, as judged from Northern blotting; no such information is available for myoH). Location to one of the six chromosomes is indicated and approximate distance from the end is given in kb. More information corresponding to the YAC physical mapping studies of A. Kuspa and W. F. Loomis is available on the following web site (http://www-biology.ucsd.edu/loci_maps.html). The last two columns indicate a primary reference and the accession numbers (for cDNA and protein sequences, when available), respectively.

MATERIALS AND METHODS

Cell Culture

Cells of the *Dictyostelium discoideum* strain AX2 (kindly provided by G. Gerisch, MPI for Biochemistry, Martinsried) were cultivated in shaken cultures at 23°C and 190 rpm. The cells were grown in HL-5c nutrient medium at a density comprised between 1×10^5 cells/mL and 5×10^6 cells/mL *(24)*.

Preparation of Genomic DNA and Total RNA from Dictyostelium discoideum

D. discoideum cells were harvested at a density of $3–5 \times 10^6$ cells/mL. Genomic DNA was purified according to a standard procedure *(25)*. Total RNA was purified by acid guanidinium thiocyanate-phenol-chloroform extraction as described by Chomczynski and Sacchi *(26)*. In preliminary experiments, following this protocol strictly, we observed that after separation on agarose gels, the high molecular weight RNAs including the 26S rRNa band were somehow missing. This problem was completely abolished by introducing the following modifications. Prior to extraction, the cells were washed once in 1× TE, pH 8.0. The phenol was saturated not simply with water, but with 2 *M* sodium acetate, pH 5.2. Accordingly, this same 2 *M* sodium acetate solution was used in the extraction mixture. Buffering at this precise pH appears crucial to optimise this purification for *D. discoideum*, but we also found it critical for the extraction of total RNA from *Escherichia coli* (data not shown). In addition, *D. discoideum* has an extremely high content of carbohydrates especially glycogen *(27)*. Therefore, a final wash of the RNA pellet with 4M LiCl was introduced to dissolve glycogen, but not the bulk of RNA, as described for glycogen rich liver cells *(28)*.

Reverse Transcription of Total RNA

The reverse transcription of 1 μg total RNA was performed in a 20 μL reaction. RNase Inhibitor and M-MLV Reverse Transcriptase Rnase H Minus were used as described by the manufacturer (Promega). The reverse transcription reactions were carried out not with a poly-dT primer, but with a myosin-specific primer chosen appropriately according to the subsequent PCR reactions (*see* Table 2 and Fig.1). Four microliters of reverse transcription reactions were used as template for PCR.

Table 2

Characteristics of the Different Primers Used for PCR

Name	Orientation	Conserved motif (Protein)	Sequence (DNA, with codon triplets delineated by spaces)	Degeneration/ no. of Inosines	Annealing temperature
Myo1	sense	GESGAGKT	5' GGI GA$^{G}/_{A}$ $^{T}/_{A}$ $^{G}/_{C}$ I GGI GCI GGI AA $^{G}/_{A}$ AC 3'	16/5 I	70°C
Myo2	sense	EAFGNAKT	5' GA$^{G}/_{A}$GCI TT$^{T}/_{C}$ GGI AA$^{T}/_{C}$ GCI AA$^{G}/_{A}$ AC 3'	6/3 I	66°C
Myo3	anti-sense	EAFGNAKT	5' GT $^{T}/_{C}$ TTI GC$^{G}/_{A}$ TTI CC$^{G}/_{A}$ AAI GC$^{T}/_{C}$ TC 3'	32/3 I	66°C
Myo4	sense	ERNYHIFY	5' GA$^{G}/_{A}$ $^{A}/_{C}$GI AA$^{T}/_{C}$ TA$^{T}/_{C}$ CA$^{T}/_{C}$ ATI TT$^{T}/_{C}$ T 3'	64/2 I	54°C
Myo5	anti-sense	ERNYHIFY	5' $^{G}/_{A}$TA $^{G}/_{A}$AA IAT $^{G}/_{A}$TG $^{G}/_{A}$TA $^{G}/_{A}$TT IC$^{T}/_{G}$ $^{T}/_{C}$TC 3'	128/2 I	54°C
Myo6	anti-sense	LDIYGFE	5' $^{T}/_{C}$TC $^{G}/_{A}$AA ICC $^{G}/_{A}$TA ^{T}/G/$_{A}$AT $^{G}/_{A}$TC 3'	48/1 I	48°C
Myo7	anti-sense	FEQFCINY	5' $^{G}/_{A}$TA $^{G}/_{A}$TT IAT $^{G}/_{A}$CA $^{G}/_{A}$AA $^{T}/_{C}$TG $^{T}/_{C}$TC $^{G}/_{A}$A 3'	128/1 I	57°C

This table summarizes some of the features of the primers used for the screening strategy. The primer nomenclature corresponds to Fig. 1. Sense and antisense orientation, the corresponding conserved protein motifs from which they are derived, and the primer sequences are presented. Slashes between bases indicate the two to three nucleotide degeneration introduced during the synthesis. In place of four bases degeneration, we introduced inosines. The total degree of degeneration (total number of different oligonucleotides within one population) and the total number of inosines are also given. The last column gives an approximate annealing temperature for the primer population.

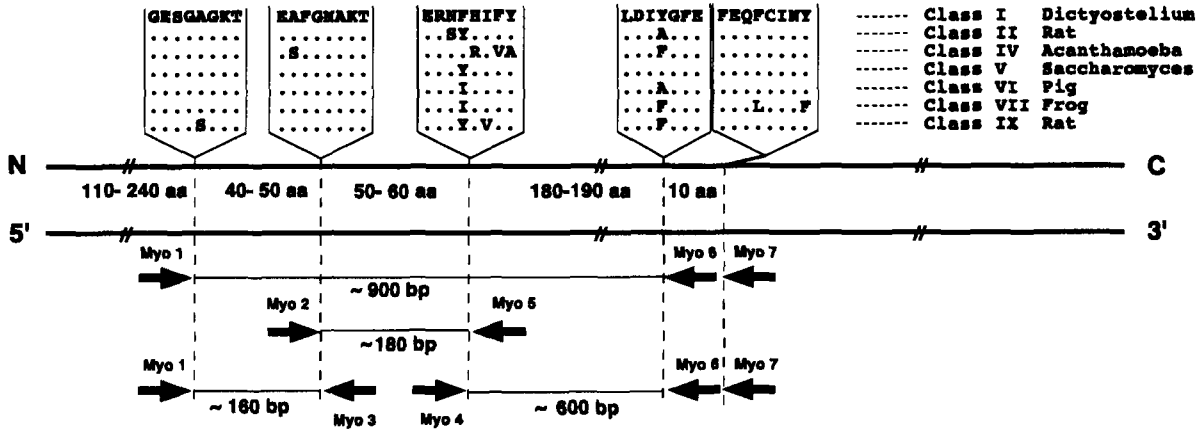


Fig. 1. Design of the PCR strategy within the myosin head domain. The head domain is very conserved throughout the myosin classes and in different organisms. This figure is a linear representation of myosin head domains at the protein and nucleic acid levels (thick lines from **N** to **C** and from **5'** to **3'**, respectively). Accordingly, approximate lengths are given in amino acids **(aa)** or base pairs **(bp)**. Sequences of the conserved protein motifs used to derive guessmer-oligonucleotides **(arrows)** are shown in one letter code, and the myosin classes and organisms of origin are indicated on the right. Dots represent residue identity. The nomenclature for the primers is the same as in Table 2.

Polymerase Chain Reaction (PCR)

In a PCR reaction many parameters can be tuned. The design of primers as well as the crucial optimization of conditions such as the pH, the $MgCl_2$ concentration, and the type of template used are described in the Results section. Except when mentioned otherwise, PCR was carried out with a denaturation time of 30 s at 94°C, a primer annealing time of 30 seconds using one of four different temperatures (40, 45, 50, 55°C), a primer extension of 30 s at 72°C. A standard profile included 30 cycles, with an initial denaturation at 94°C for 5 minutes and a final 7 min elongation time at 72°C. At last, the samples were cooled to 4°C until further use. Twenty microliter reactions were set up with 200 μ*M* dNTPs and 1 μ*M* of each primer and 0.5 U AmpliTaq DNA polymerase (Perkin Elmer). Various amounts of different templates were used, 400 ng of genomic DNA, 4 μL of reverse transcription reactions or 1 μL of a stock solution of a λgt11 phage library (*D. discoideum* cDNA, 4×10^{10} pfu/mL, Clontech). All PCR reactions were carried out in a GeneAMP PCR System 2400 instrument (Perkin Elmer).

Molecular Cloning of the PCR Products

The PCR products were separated from primers and nucleotides by using a spin column (QIAquick PCR purification Kit, Qiagen). Without further size selection, the whole sample was cloned into the pCR-Script vector according to the manufacturer (pCR-Script Amp SK(+) Cloning kit, Stratagene). For this purpose, the ends of the amplified DNA had to be polished by Pfu polymerase and cloned into the blunt, but not dephosphorylated SrfI site. SrfI is present during the ligation reaction and cuts self-ligated pCR-Script vectors. One tenth of the ligation was electroporated into different *Escherichia coli* strains: sure strain *(29)*, XL1-Blue strain *(30)*, or XL1-Blue MRF′ strain *(31)*, using an Electro Cell Manipulator 600 instrument (BTX Company).

Analysis of the Recombinant Clones by PCR and Colony Hybridization

In order to ascertain the presence of an insert, white clones were randomly picked, transfered in 50 μL H_2O and 2 μL of this suspension used as template for PCR with T3 and T7 primers. Positive clones were sequenced, analyzed, and fragments identified as myo-

sins were used as probes for the subsequent rounds of hybridization screening.

DNA probes were labeled nonradioactively with Digoxigenin-11-dUTP (DIG-dUTP) by PCR as described by the manufacturer (PCR DIG labeling Mix, Boehringer Mannheim), using the same myosin-specific primers as for the creation of the libraries. After ligation, transfection, and plating, colonies were transferred onto Nylon membranes (82 mm diameter, Boehringer Mannheim) and crosslinked by UV (UV crosslinker, Amersham Life Science). Hybridization and detection were carried out according to the manufacturer (for more details, see instructions in "The DIG System User's Guide for Filter Hybridization," Boehringer Mannheim). Briefly, the prehybridization solution was in Boehringer's standard buffer. For hybridization, each of the DIG-labeled probe was diluted in prehybridization solution at about 15 ng/mL. Prehybridization and hybridization were carried out at 68°C in an oven/shaker (Amersham Life Science).

Sequencing and Sequence Analysis

DNA sequencing was performed with the DNA Sequencing kit Version 2.0 (Amersham Life Science) which is based on the dideoxy chain termination method of Sanger *(32)* and the incorporation of radioactively labeled (α-^{35}S)-dATP. Oligonucleotide primers for PCR and sequencing were synthesized in the facilities of our institute. Sequences were analysed with BLAST and FASTA, and aligned with the clustalW program from the HUSAR package (DKFZ, Heidelberg). The phylogenetic tree was constructed by "bootstrapping" the alignment 1000-fold and plotted with the TreeView freeware.

RESULTS

Primer Design

Our aim was to design a PCR-based strategy that would unbiasedly amplify all the myosin sequences present in the genome of *D. discoideum*. Initially, all the myosin sequences present in the databases were aligned and the five most conserved motifs within the head domain were chosen for deriving degenerate primers (or guessmers) in sense and antisense orientations (Fig. 1). We also chose these regions as they are scattered throughout the head and increase

the number of possible primer pair combinations. As not only the motifs but also the intervening distances seemed well conserved, the position of the primers gave a relative approximation of the expected length of PCR products. Note that, in order to additionally lower the risk of bias, the codon usage of *D. discoideum*, despite it being extremely skewed *(33)* was not taken into account. In addition we empirically tried to maintain a balance between the overall degeneracy and the number of inosines (Table 2).

PCR Conditions

Different parameters of a PCR reaction can be optimized, mainly in order to fine-tune its stringency. We tried to define conditions that would amplify all and every myosin, with minimal contamination of nonmyosin sequences. In other words, we tried to find the right balance between specificity and the complexity of the resulting pattern of bands. We varied the annealing temperature, the source of the template, the $MgCl_2$ concentration, and the pH for every of the nine possible primer pair combinations (Myo1/Myo3, Myo1/Myo5, Myo1/Myo6, Myo1/Myo7, Myo2/Myo5, Myo2/Myo6, Myo2/Myo7, Myo4/Myo6, Myo4/Myo7, *see* Fig. 1). Table 3 summarizes the different conditions tested.

Figure 2 illustrates how we monitored the influence of pH and $MgCl_2$ concentration. In this case, PCR was carried out with reverse transcribed RNA as template, an annealing temperature of 50°C and the primer pair combination Myo1/Myo5. In these conditions, the main product has a length of about 340 bp, close to the size predicted by the alignment (Fig. 1). Variations of pH and $MgCl_2$ concentration strongly influence the diversity of the amplification products. Both a higher pH (up to 9.5) and a higher $MgCl_2$ concentration (up to 3.5 m*M*) led to a greater complexity of the resulting band pattern (Fig. 2). Raising the $MgCl_2$ concentration up to 5 m*M* did not further increase the complexity (Fig. 2, lane 10).

We also tested different templates for the PCR reaction. The use of genomic DNA ensures that every myosin is equally represented, independently of potential cell cycle- or developmental stage-dependent regulation, contrary to cDNAs which obligatorily reflect the expression pattern of a given cell type. On the other hand, long introns could be a disadvantage for efficent PCR. Figure 3 shows a comparison of different primer pairs used with either genomic or reverse transcribed RNA as templates. The primer pair Myo1/Myo6

Table 3

Variation of Parameters for the PCR Reactions

Parameter	Range tested
Annealing temperature	40°C , 45°C, 50°C, and 55°C
Template	genomic DNA, reverse transcribed RNA, phage cDNA library (λgt11)
Magnesium concentration	1.5 m*M*, 2.5 m*M*, and 3.5 m*M* $MgCl_2$
pH	8.3, 9.0, 9.5, and 10.0

This table lists the different parameters that were optimized in order to influence the specificity and complexity of PCR products patterns. Most of the parameters were tested in all possible combinations for each of the nine primer pairs.

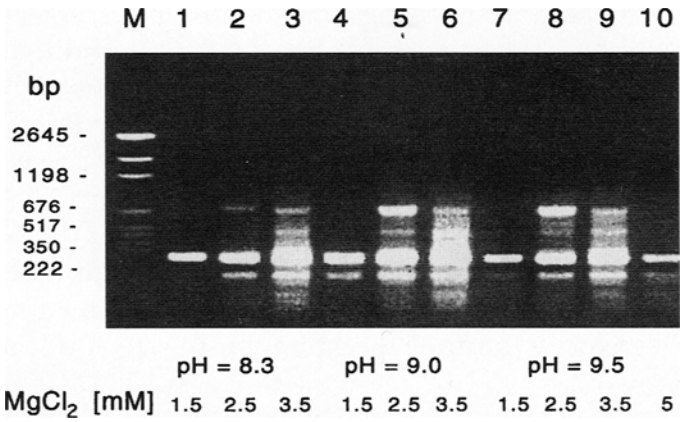


Fig. 2. The complexity of the band pattern strongly depends on the reaction conditions. In the example shown, one primer pair, Myo1 and Myo5 was tested with reverse transcribed RNA as template. The influence of three different **pH** and **$MgCl_2$** concentrations varying between 1.5 and 5.0 m*M* were monitored. The annealing temperature was 50°C for all conditions.

produces the apparently most diverse band pattern (Fig. 3, lanes 5 and 6) from both templates, and also gives rise to the most pronounced difference between both templates. It is unlikely that this diversity is primarily because of the presence of introns, as none of the products we cloned and sequenced contained any introns *(see* next two sections*)*. This higher heterogeneity indeed illustrates the superiority of genomic DNA in the creation of an exhaustive library of myosin fragments. The conditions for each primer pair had to be optimized, because, as observed for primer pair Myo1/Myo7 with

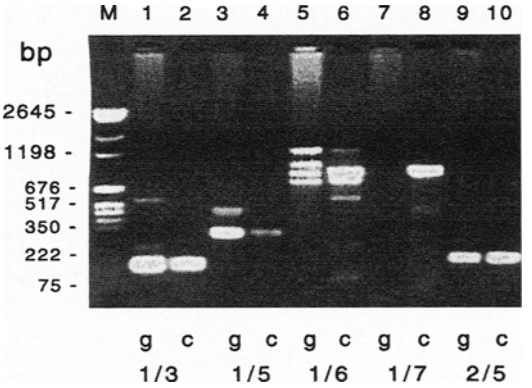


Fig. 3. The use of different templates—genomic or reverse transcribed RNA—generates different band patterns. PCR was carried out with a pH of 9.0, 3.5 m*M* $MgCl_2$ and an annealing temperature of 50°C, except for lanes 2, 4, and 8 for which 55°C was used. The templates used were either genomic DNA **(g)** or reverse transcribed RNA **(c)**. We monitored the influence of the type of template with different primer pair combinations [Myo1/Myo3, **(1/3)**; Myo1/Myo5, **(1/5)**; Myo1/Myo6, **(1/6)**; Myo1/Myo7, **(1/7)**; Myo2/Myo5, **(2/5)**]. The primer nomenclature is the same as in Table 2.

genomic DNA as template (Fig. 3, lane 7), suboptimal condition sometime result in no amplification at all. The latter problem was solved by decreasing the annealing temperature to 40°C (not shown).

Identification of the PCR Products from Different Myosin Libraries

In order to test our system and generate myosin probes for the hybridization screening strategy *(see below)* a PCR reaction giving rise to a relatively low complexity was chosen to create the first myosin library. Low complexity ensured a high specificity for myosins fragments. Therefore, PCR was carried out with the primer pair combination Myo1/Myo5, reverse transcribed RNA as template, at pH 9, with 3.5 m*M* $MgCl_2$ and an annealing temperature of 55°C. After ligation and transformation, white colonies were investigated by PCR and sequencing. Sequence analysis led to the identification of three different myosins, *mhcA*, *myoC*, and *myoI*. No nonmyosin sequence was found in this screen which emphasizes the high specificity and low complexity of this first library.

In order to increase the complexity stepwise, the second myosin library was created with the same primer pair, a pH of 9.5, 3.5 m*M* $MgCl_2$, but with an annealing temperature of 50°C and with genomic

DNA as template. As before, white colonies were investigated by PCR and sequencing. In this second screening, 52 inserts were analyzed and led to the identification of 20 × *mhcA*, 7 × *myoB*, 10 × *myoC*, 10 × *myoD*, 3 × *myoJ*, and only two artefactual amplification products. Note that none of these first two rounds of screening were conducted in an exhaustive manner and therefore, no solid conclusion can be derived from the type and number of sequences isolated. It may be that the myosins amplified by RT-PCR are among the most abundant myosin mRNAs in vegetative cells.

Identification of Two Novel Myosins by Negative Hybridization Screening

In order to maximally avoid repetitive fishing of already identified myosins, we then applied a negative hybridization screening strategy. For this purpose we generated DIG-dUTP-labeled probes from myosins identified in the first two libraries. Labeled probes from the first screen *(mhcA)* as well as from the second one *(myoB, myoC,* and *myoD)* were pooled with a final concentration of 15 ng/mL of hybridization solution, each. The efficient procedure based on hybridization allowed us to screen hundreds of novel clones in a very short time. Thus, for the third round of screening, the PCR reaction with greatest complexity was ligated (primer pair Myo1/Myo6, annealing temperature of 50°C, at pH of 9.0 with 3.5 m*M* $MgCl_2$ and genomic DNA as template; Fig. 3, lane 5). After colony lifting and hybridization, every colony that did not give a signal was picked and analyzed by sequencing. Final identification was carried out by computer analysis with the clustalW package as well as BLAST and FASTA search engines. We investigated 140 sequences and identified 3 × *myoA*, 5 × *myoH* and two additional, so far unknown myosin sequences. Probes derived from both fragments were sent to Drs. W. Loomis and A. Kuspa (UC, San Diego, CA) for physical mapping on an ordered YAC library which covers 98% of the *D. discoideum* genome. The first probe mapped to a locus on chromosome V, (between the genes *acgA* and *actK;* Table 1) which had been previously identified as a potential myosin locus then named *myoK (22),* but from which no sequence could be retrieved. The other fragment mapped to a new myosin locus on chromosome VI, between the genes *aldB* and *cabA2,* and was named *myoM* according to the alphabetical nomenclature (Table 1). Furthermore, these probes were

also used in the first step leading to isolation of full-length sequences for both *myoK* and *myoM* (Accession number AF090534, Schwarz, Neuhaus, Kistler, and Soldati; and Accession number AF090533, Geissler, Ullmann, and Soldati, respectively, manuscripts in preparation).

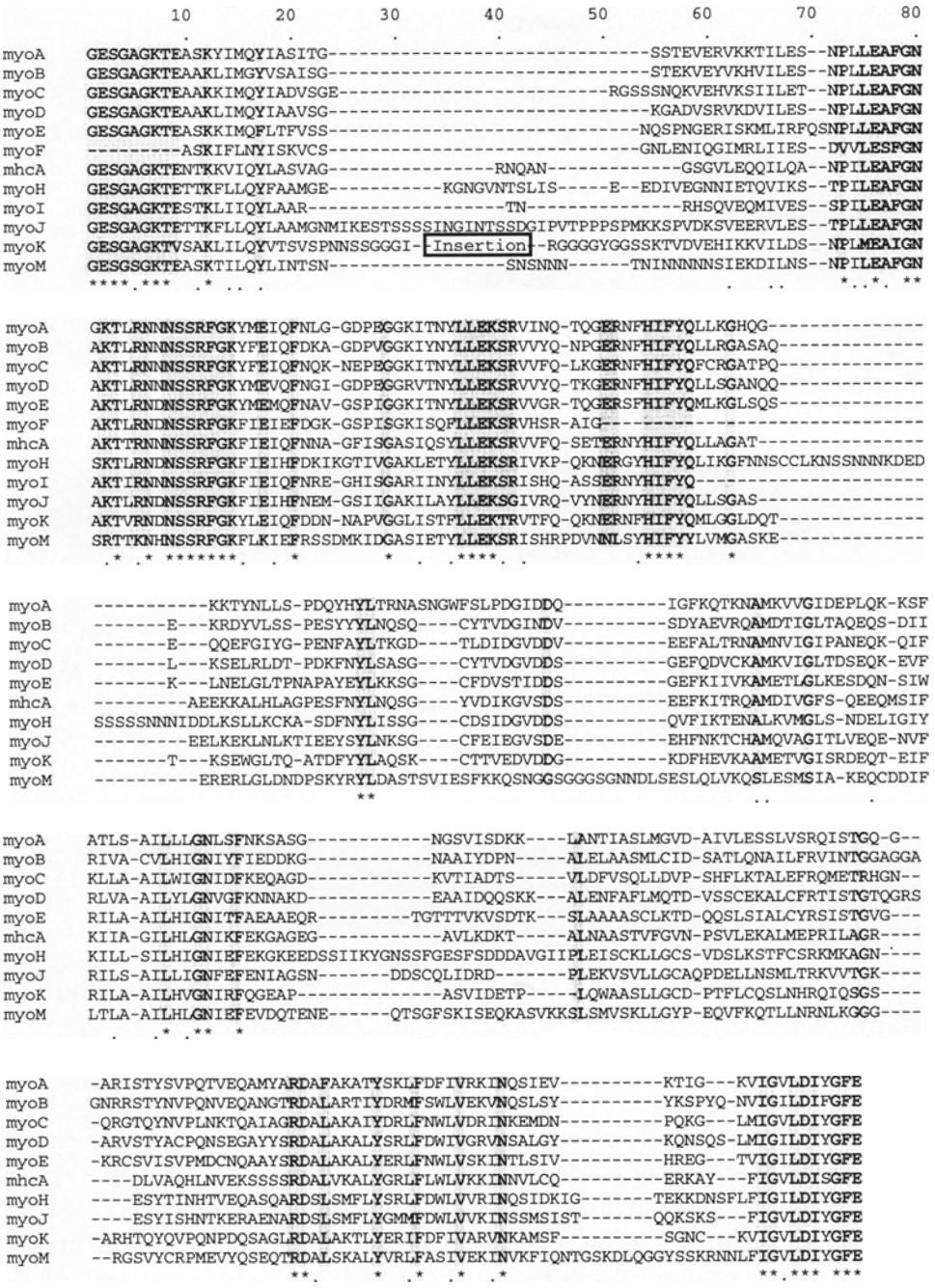
The complexity of this third library was high and therefore accompanied by lower specificity, resulting in the isolation of about hundred nonmyosin fragments. The known dynein heavy chain as well as a sequence very likely coding for a novel RNA helicase were amplified, apparently because their P-loop motifs (primer Myo1) relatively closely resemble the one of a myosin. Other sequences were either not clearly identifiable or were PCR artefacts such as amplifications with only one primer.

Preliminary Sequence Analysis of myoK and myoM Fragments

Both MyoK and MyoM exhibit very divergent features. MyoK carries a big insertion of about 143 amino acids right behind the GESGAGKT P-loop motif (Fig. 4), which is so far unique in the myosin world. MyoK is nevertheless clearly identified as a member of the class I, whereas MyoM seems to be the founder of a novel class of myosins (Fig. 5, and accompanying review). Both MyoK and MyoM carry a potential phosphorylation site, a Thr and a Ser, respectively, at a very precise position of the head domain called the TEDS site *(34)*. It was shown that, except for the myosins XIV found in Apicomplexan protozoan, over 95% of the myosins carry an acidic or a phosphorylatable residue at this position *(34)*. The residue is thus an Asp (D) or a Glu (E) except in myosin VI and in a subclass of myosin I where it is a Ser (S) or Thr (T), phosphorylation of which was shown to be essential for function *(35–38)*. Kinases able to phosphorylate these latter sites, the Myosin I Heavy Chain Kinases

---

Fig. 4. Alignment of the translated myosin fragments identified. The fragments are all derived from the study presented here, with the exception of *myoE* and *myoF*, which were described by others. The fragments span the region between primers Myo1 and Myo6, corresponding to the conserved motifs GESGAGKT and LDIYGFE, with the exception of *myoI*, which was identified in our first screening and spans only between the primer Myo1 and Myo5 (ERNYHIFY). On the last line of each 80 residues block, the stars indicate amino acid identity, whereas dots indicate conservative changes. Residues shaded in gray represent positions conserved



with maximally one deviation. MyoK has an additional 143 residues in an insertion **(boxed)** situated in the surface loop 1: "GGSGGGNGGIPQYDG-GGGPPSRGGGPPPTRGRGGPPPPIPQNRGAPPVSNGGAPPPVARGP VAPPPTRGAPPTRGGGPANRGGRGGGPPPVSTS".

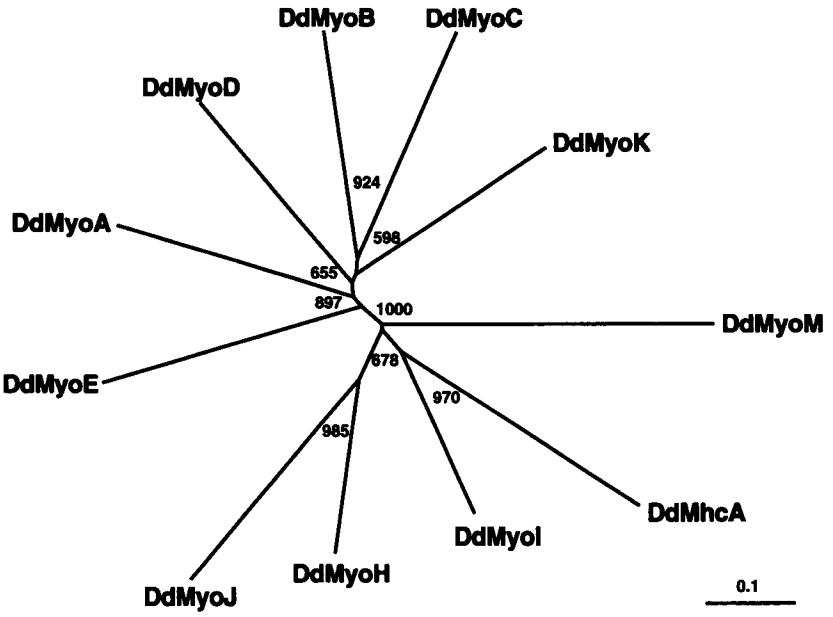


Fig. 5. Phylogenetic analysis of the *D. discoideum* myosin family. The computation was carried out with all the sequences available for *D. discoideum* myosins, with the exception of *myoF*, as the fragment known is too short to result in meaningful classification. The tree was computed using full length sequences (when available) by the ClustalW package and bootstrapped 1000 times; the internal edge labels indicate how often a given branching was found. The tree was drawn by the TreeView freeware. The scale bar represents 10% divergence. The accession numbers are given in Table 1.

(MIHCK), belong to the PAK/Ste20 family and were identified in mammals, *S. cerevisiae*, *A. Castellanii*, and *D. discoideum (38–41)*.

DISCUSSION

The aim of this work was to extend and complete the catalog of myosins in the eukaryotic model organism *D. discoideum*. Using degenerate primers derived from highly conserved motifs within the head domain, we were able to amplify eight of the ten myosins identified by previous studies. In addition, we identified two novel myosins. This was rendered possible by the rational design of the strategy. First, we optimized a palette of parameters so as to increase the complexity of the products of the PCR reactions, including the design of primers that were not biased according to the *D. discoideum*

codon usage. Second, we did not perform a size selection on the PCR reaction products. Third, we screened three different libraries of myosin fragments constructed with a stepwise increase of complexity, allowing us to first generate specific myosin probes that were then used in a final negative hybridization screening. During the course of the successive screenings, it became evident that there are no single reaction conditions that could realistically lead to amplification of the whole myosin family. This conclusion is supported by the identification of two novel exotic myosins, which show on the one hand relatively high divergence in the otherwise conserved regions used to design the primers *(myoM)*, and a unique insertion in a surface loop that results in a PCR amplification product of 1300 bp instead of about 900 bp (*myoK, see* Fig. 1). In the light of this success, it is still unclear why *myoE* and *myoF* resisted our efforts. A possible explanation was published after we had finished our screening. The amplification of adenine and thymidine (AT)-rich DNA in PCR reactions performed at the standard elongation temperature of 72°C can be very difficult, as reported for *Plasmodium falciparum* by Xin-zhuan Su, et al. *(42)*. Although the AT content of *D. discoideum* DNA (about 75% *(9)* is slightly less than that of *P. falciparum*, we have then found that decreasing the elongation temperature to 60 or 65°C leads to efficient amplification of products which failed at 72°C (data not shown).

Have we achieved an exhaustive screening? As already mentioned, a strategy employing physical mapping on an ordered YAC library *(22)* by sequential low stringency hybridization with two probes derived from 5' fragments of *mhcA* and *myoA* revealed three potential myosin loci from which no sequence could be isolated by subsequent PCR reactions. We now report that *myoK* is a true myosin locus, but we failed to identify other myosin fragments corresponding to the *myoG* and *myoL* loci. *MyoK* was missed in the previous screening probably because the PCR product should have been surprisingly 400 bp longer than expected. It is not clear whether the conditions used for PCR did not allow for amplification or whether it was mis-identified as nonspecific. In addition, we now show that at least one myosin locus, *myoM*, went undetected in the physical mapping, potentially because *myoM* is one of the most divergent myosin sequence identified to date. Our strategy "picked it" because despite overall high divergence, it deviates only minimally in the sequences used for priming. Therefore, the final census

will likely only be reached after the ongoing efforts at the level of an integrated Genome Project are completed.

Other groups already have reported on the use of PCR approaches to identify novel members of the myosin superfamily in different organisms. For example, Mooseker and collaborators used PCR to obtain a "fingerprint" of the diversity of unconventional myosins expressed in a given cell type of mammalian origin *(43)*. Also, Titus and collaborators developed a strategy using a combination of PCR on cDNA and genomic DNA to extend the repertoire of known myosins in the nematode *Caenorhabditis elegans (44)*. The overall strategy presented here was designed to allow for a nearly exhaustive survey of the myosin superfamily in any given organism. Because our primers do not take the codon usage of *D. discoideum* into account, we could readily test our PCR conditions on *Toxoplasma gondii* genomic DNA. *T. gondii* is an obligate intracellular parasite of the Apicomplexa family, which include *P. falciparum*. These protozoan actively penetrate their host cells by making use of their own acto-myosin cytoskeleton *(45,46)*, but little is known about the molecular details of this process. Results from a first round of screening revealed the presence of a novel myosin in this organism, which is a potential candidate to be involved in the gliding motility of these parasites (Hettmann, Herm, Geiter, Frank, Schwarz, Soldati, and Soldati; manuscript submitted).

In conclusion, our data on the potentially exhaustive screening of the myosin superfamily in *D. discoideum* and the preliminary success in *T. gondii* argue for both the high potential and broad applicability of this screening strategy.

Note

Following a similar strategy, the group of Dr. K. Sutoh (Tokyo University, Japan) has independently identified myosin sequences corresponding to *myoK* and *myoM*. Full-length cDNA and genomic sequences have also subsequently been cloned. It is worth noting that no sequence was isolated corresponding to the proposed myosin loci myoG and myoL.

ACKNOWLEDGMENTS

The authors owe a special thank you to Drs. W. Loomis and A. Kuspa (UC, San Diego, CA) for their generous offer to map both myoK and myoM loci. We also thank the other members of the labo-

ratory and of the group of Dietmar Manstein for helpful comments and support during the course of this study. The work of E.C. S. was supported by Grant No. So 376/1-1 from the Deutsche Forschungsgemeinschaft. H. G. was a recipient of a Doctoral fellowship from the Max-Planck-Gesellschaft.

REFERENCES

1. Mermall, V., Post, P. L., and Mooseker, M. S. (1998) Unconventional myosins in cell movement, membrane traffic, and signal transduction. *Science* **279,** 527–533.
2. Probst, F. J., et al. (1998) Correction of deafness in *shaker* -2 mice by an unconventional myosin in BAC transgene. *Science* **280,** 1444–1447.
3. Wang, A., et al. (1998) Association of unconventional myosin MYO15 mutations with human nonsyndromic deafness DFNB3. *Science* **280,** 1447–1451.
4. Sweeney, F. P., Pocklington, M. J., and Orr, E. (1991) The yeast type II myosin heavy chain: analysis of its predicted polypeptide sequence. *J. Muscle Res. Cell Motil.* **12,** 61–68.
5. Goodson, H. V., Anderson, B. L., Warrick, H. M., Pon, L. A., and Spudich, J. A. (1996) Synthetic lethality screen identifies a novel yeast myosin I gene (MYO5): myosin I proteins are required for polarization of the actin cytoskeleton. *J. Cell Biol.* **133,** 1277–1291.
6. Govindan, B., Bowser, R., and Novick, P. (1995) The role of Myo2, a yeast class V myosin, in vesicular transport *J. Cell Biol.* **128,** 1055–1068.
7. Haarer, B. K., Petzold, A., Lillie, S. H., and Brown, S. S. (1995) Identification of MYO4, a second class V myosin gene in yeast. *J. Cell Sci.* **107,** 1055–1064.
8. Bähler, M. (1996) Myosins on the move to signal transduction. *Current Opin. Cell Biol.* **8,** 18–22.
9. Loomis, W. F. and Kuspa, A. (1997) The genome of *Dictyostelium discoideum*, in: Dictyostelium - *A Model System for Cell and Developmental Biology.* (Maeda, Y., Inouye, K. and Takeuchi, I., eds.), Universal Academy Press, Tokyo, Japan, pp. 15–30.
10. Smith, D. W. and Loomis, W. F. (1997) DictyDB - A genomic database for *Dictyostelium discoideum.*, in: Dictyostelium - *A Model System for Cell and Developmental Biology.* (Maeda, Y., Inouye, K. and Takeuchi, I., eds.), Universal Academy Press, Tokyo, Japan, pp. 471–477.
11. Lee, S. F. and Cote, G. P. (1993) Isolation and characterization of three *Dictyostelium* myosin-I isozymes. *J. Biol. Chem.* **268,** 20,923–20,929.

12. Jung, G. and Hammer III, J. A. (1990) Generation and characterization of *Dictyostelium* cells deficient in a myosin I heavy chain isoform. *J. Cell Biol.* **110,** 1955–1964.
13. Wessels, D., Murray, J., Jung, G., Hammer III, J. A., and Soll, D. R. (1991) Myosin IB null mutants of *Dictyostelium* exhibit abnormalities in motility. *Cell Motil. Cytoskel.* **20,** 301–315.
14. Peterson, M. D., Novak, K. D., Reedy, M. C., Ruman, J. I., and Titus, M. A. (1995) Molecular genetic analysis of myoC, a *Dictyostelium* myosin I. *J. Cell Sci.* **108,** 1093–1103.
15. Novak, K. D., Peterson, M. D., Reedy, M. C., and Titus, M. A. (1995) *Dictyostelium* myosin I double mutants exhibit conditional defects in pinocytosis. *J. Cell Biol.* **131,** 1205–1221.
16. Temesvari, L. A., Bush, J. M., Peterson, M. D., Novak, K. D., Titus, M. A., and Cardelli, J. A. (1996) Examination of the endosomal and lysosomal pathways in *Dictyostelium discoideum* myosin I mutants. *J. Cell Sci.* **109,** 663–673.
17. Jung, G., Wu, X. F., and Hammer III, J. A. (1996) *Dictyostelium* mutants lacking multiple classic myosin I isoforms reveal combinations of shared and distinct functions. *J. Cell Biol.* **133,** 305–323.
18. Titus, M. A. (1997) An unconventional myosin required for phagocytosis. *Mol. Biol. Cell* **8,** 352a.
19. Uyeda, T. Q. and Titus, M. A. (1997) The myosins of *Dictyostelium*, in: Dictyostelium - *A Model System for Cell and Developmental Biology.* (Maeda, Y., Inouye, K. and Takeuchi, I., eds.), Universal Academy Press, Tokyo, Japan, pp. 43–64.
20. el-Amraoui, A., Sahly, I., Picaud, S., Sahel, J., Abitbol, M., and Petit, C. (1996) Human Usher 1B/mouse shaker-1: the retinal phenotype discrepancy explained by the presence/absence of myosin VIIA in the photoreceptor cells. *Hum. Mol. Gen.* **5,** 1171–1178.
21. Liu, X., Vansant, G., Udovichenko, I. P., Wolfrum, U., and Williams, D. S. (1997) Myosin VIIa, the product of the Usher 1B syndrome gene, is concentrated in the connecting cilia of photoreceptor cells. *Cell Motil. Cytosk.* **37,** 240–252.
22. Titus, M. A., Kuspa, A., and Loomis, W. F. (1994) Discovery of myosin genes by physical mapping in *Dictyostelium. Proc. Natl. Acad. Sci. USA* **91,** 9446–9450.
23. Hammer III, J. A. and Jung, G. (1996) The sequence of the *Dictyostelium* myo J heavy chain gene predicts a novel, dimeric, unconventional myosin with a heavy chain molecular mass of 258 kDa. *J. Biol. Chem.* **271,** 7120–7127.
24. Sussman, M. (1987) Cultivation and synchronous morphogenesis of *Dictyostelium* under controlled experimental conditions. *Meth. Cell Biol.* **28,** 9–29.

25. Manstein, D. J., Titus, M. A., De Lozanne, A. and Spudich, J. A. (1989) Gene replacement in *Dictyostelium*: generation of myosin null mutants. *EMBO J.* **8,** 923–932.
26. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162,** 156–159.
27. Nellen, W., Datta, S., Reymond, C., Sivertsen, A., Mann, S., Crowley, T. and Firtel, R. A. (1987) Molecular biology in *Dictyostelium*: tools and applications. *Meth. Cell Biol.* **28,** 67–100.
28. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1997) Preparation and analysis of RNA, in: *Current Protocols in Molecular Biology,* **1:**4.2.2.
29. Greener, A. (1990) E. coli SURE™ Strain: Clone "unclonable" DNA. *Strategies* **3,** 5–6.
30. Bullock, W. O., Fernandez, J. M., and Short, J. M. (1987) XL1-blue: A High Efficiency Plasmid Transforming Reca ES. *Biotechniques* **5,** 376–379.
31. Jerpseth, B., Greener, A., Short, J. M., Viola, J., and Kretz, P. L. (1992) XL1-blue MRF' *E. coli* Cells: McrA$^-$, McrCB$^-$, Mrr$^-$, HsdR$^-$ derivative of XL1-Blue cells. *Strategies* **5,** 81–83.
32. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
33. Warrick, H. M. (1987) Codon frequency in *Dictyostelium discoideum. Meth. Cell Biol.* **28,** 497–500.
34. Bement, W. M. and Mooseker, M. S. (1995) TEDS rule: a molecular rationale for differential regulation of myosins by phosphorylation of the heavy chain head. *Cell Motil. Cytosk.* **31,** 87–92.
35. Albanesi, J. P., Hammer III, J. A., and Korn, E. D. (1983) The interaction of F-actin with phosphorylated and unphosphorylated myosins IA and IB from *Acanthamoeba castellanii. J. Biol. Chem.* **258,** 10,176–10,181.
36. Albanesi, J. P., Fujisaki, H., Hammer III, J. A., Korn, E. D., Jones, R., and Sheetz, M. P. (1985) Monomeric *Acanthamoeba* myosins I support movement in vitro. *J. Biol. Chem.* **260,** 8649–8652.
37. Cote, G. P., Albanesi, J. P., Ueno, T., Hammer III, J. A., and Korn, E. D. (1985) Purification from *Dictyostelium discoideum* of a low-molecular weight myosin that resembles myosin I from *Acanthamoeba castellanii. J. Biol. Chem.* **260,** 4543–4546.
38. Lee, S. F., Egelhoff, T. T., Mahasneh, A., and Cote, G. P. (1996) Cloning and characterization of a *Dictyostelium* myosin I heavy chain kinase activated by Cdc42 and Rac. *J. Biol. Chem.* **271,** 27,044–27,048.
39. Brzeska, H., Lynch, T. J., Martin, B., and Korn, E. D. (1989) The localization and sequence of the phosphorylation sites of *Acanthamoeba*

myosins I. An improved method for locating the phosphorylated amino acid. *J. Biol. Chem.* **264,** 19,340–19,348.
40. Brzeska, H., Lynch, T. J., Martin, B., Corigliano-Murphy, A., and Korn, E. D. (1990) Substrate specificity of *Acanthamoeba* myosin I heavy chain kinase as determined with synthetic peptides. *J. Biol. Chem.* **265,** 16,138–16,144.
41. Wu, C., Lee, S. F., Furmaniak-Kazmierczak, E., Cote, G. P., Thomas, D. Y., and Leberer, E. (1996) Activation of myosin-I by members of the Ste20p protein kinase family. *J. Biol. Chem.* **271,** 31,787–31,790.
42. Su, X. Z., Wu, Y., Sifri, C. D., and Wellems, T. E. (1996) Reduced extension temperatures required for PCR amplification of extremely A+T-rich DNA. *Nucl. Acids Res.* **24,** 1574–1575.
43. Bement, W. M., Hasson, T., Wirth, J. A., Cheney, R. E., and Mooseker, M. S. (1994) Identification and overlapping expression of multiple unconventional myosin genes in vertebrate cell types. *Proc. Natl. Acad. Sci. USA* **91,** 6549–6553.
44. Baker, J. P. and Titus, M. A. (1997) A family of unconventional myosins from the nematode *Caenorhabditis elegans*. *J. Mol. Biol.* **272,** 523–535.
45. Dobrowolski, J. M. and Sibley, L. D. (1996) *Toxoplasma* invasion of mammalian cells is powered by the actin cytoskeleton of the parasite. *Cell* **84,** 933–939.
46. Pinder, J. C., Fowler, R. E., Dluzewski, A. R., Bannister, L. H., Lavin, F. M., Mitchell, G. H., Wilson, R. J. M., and Gratzer, W. B. (1998) Actomyosin motor in the merozoite of the malaria parasite, *Plasmodium falciparum*: implication for red cell invasion. *J. Cell Sci.* **111,** 1831–1839.
47. Titus, M. A., Wessels, D., Spudich, J. A., and Soll, D. (1993) The unconventional myosin encoded by the myoA gene plays a role in *Dictyostelium* motility. *Mol. Biol. Cell* **4,** 233–246.
48. Jung, G., Saxe III, C. L., Kimmel, A. R., and Hammer III, J. A. (1989) *Dictyostelium discoideum* contains a gene encoding a myosin I heavy chain. *Proc. Natl. Acad. Sci. USA* **86,** 6186–6190.
49. Titus, M. A., Warrick, H. M., and Spudich, J. A. (1989) Multiple actin based motor genes in *Dictyostelium*. *Cell Regulation* **1,** 55–63.
50. Urrutia, R. A., Jung, G., and Hammer III, J. A. (1993) The *Dictyostelium* myosin-IE heavy chain gene encodes a truncated isoform that lacks sequences corresponding to the actin binding site in the tail. *Biochim. Biophys. Acta* **1173,** 225–229.
51. Titus, M. A., Novak, K. D., Hanes, G. P., and Urioste, A. S. (1995) Molecular genetic analysis of myoF, a new *Dictyostelium* myosin I gene. *Biophys. J.* **68,** S152–S157.
52. Clarke, M. and Spudich, J. (1974) Biochemical and structural studies of actomyosin-like proteins from non-muscle cells. Isolation and

characterization of myosin from amoebae of *Dictyostelium discoideum*. *J. Mol. Biol.* **86,** 209–222.

53. Warrick, H. M., De Lozanne, A., Leinwand, L. A., and Spudich, J. A. (1986) Conserved protein domains in a myosin heavy chain gene from *Dictyostelium discoideum*. *Proc. Natl. Acad. Sci. USA* **83,** 9433–9437.
54. Peterson, M. D., Urioste, A. S., and Titus, M. A. (1996) *Dictyostelium discoideum* myoJ: A member of a broadly defined myosin V class or a class XI unconventional myosin? *J. Muscle Res. Cell Motil.* **17,** 411–424.