# Identification of act2, an essential gene in the fission yeast Schizosaccharomyces pombe that encodes a protein related to actin 

James P. Lees-Miller, Gilbert Henry, and David M. Helfman*<br>Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

Communicated by Barbara McClintock, October 2, 1991 (received for review July 30, 1991)


#### Abstract

Actins are a family of highly conserved proteins that are ubiquitously found among eukaryotic organisms. All actins that have previously been identified, including those of animals, plants, fungi, and protozoa, are 374-376 amino acids long and exhibit at least $70 \%$ amino acid sequence identity when compared with one another. We have cloned a gene from the fission yeast Schizosaccharomyces pombe that encodes a distantly related member of the actin protein family, herein referred to as act2. In contrast to all other actins, the derived amino acid sequence reveals that act 2 is $\mathbf{4 2 7}$ residues long and exhibits only $35-40 \%$ identity to actins, including act1 from Sch. pombe. Comparison to the known x-ray crystallographic structure of rabbit skeletal muscle actin indicates that the ATP and divalent metal ion binding sites are largely conserved in act2, while regions involved in actin-actin and actin-myosin interactions are relatively divergent. Disruption of the act 2 gene demonstrated that this gene encodes a function essential for germination of haploid spores. These findings indicate that while act2 and act1 are related proteins, they appear to have distinct functions. In addition, they demonstrate that the actin protein family is more diverse than was previously thought.


The shape and motility of eukaryotic cells are largely governed by three filamentous systems-actin filaments, intermediate filaments, and microtubules. Each of these systems possesses a major core protein-namely, actin, intermediate filament protein, and tubulin, respectively. In a given organism or cell type, considerable isoform diversity of both core proteins and associated proteins is often present. Actin filaments consist of a double-helical array of actin monomers polymerized in a polar manner. A number of proteins are associated with actin that function in nucleating, capping, stabilizing, severing, bundling, and mechanical movement of actin filaments. Through the interactions with these proteins, actin filaments participate in several essential cellular functions, including cell motility, cytokinesis, maintenance of cell structure, and organelle movement (reviewed in ref. 1). Many organisms have multiple isoforms of actin that often exhibit developmentally regulated and cell-type-specific expression (reviewed in refs. 2 and 3). Actin isoforms can also exhibit differential localization within a single cell (4). Actins that have been characterized from a broad range of phyla including vertebrates, invertebrates, fungi, plants, and protozoa have between $70 \%$ and $90 \%$ amino acid sequence identity and are 374-376 amino acids in length. Budding yeast (Saccharomyces cerevisiae) and fission yeast (Schizosaccharomyces pombe) each possess a single gene (actl) encoding an actin with an amino acid sequence that is $\approx 90 \%$ identical to the sequence of vertebrate actins $(5,6)$. The actl gene of budding yeast is known to be essential for viability (7). Several proteins that interact with actin have been identified in budding yeast, including profilin, tropomyosin, and myosin

[^0]

FIG. 1. Restriction map of the 5.8-kilobase (kb) HindIII-HindIII fragment carrying act2 is shown with the 5 '-end to the left (Top). The localization of the five exons (Middle) and the region deleted for gene disruption (Bottom) are illustrated. To make the disruption construct, an internal $2.5-\mathrm{kb}$ HindIII-BamHI fragment and the $3^{\prime}$-end 1.1-kb BamHI-HindIII fragment were each inserted into a pUC19 vector. The vector containing the internal fragment was cut with HindIII, a BamHI linker was added destroying the HindIII site, and the construct was digested with BamHI and Bgl II. The resulting $1.9-\mathrm{kb}$ fragment was ligated into the BamHI site of the vector containing the $3^{\prime}$-end fragment so as to leave an intact internal BamHI site into which the ura4 ${ }^{+}$gene was inserted. These manipulations resulted in the replacement of a $1.4-\mathrm{kb}$ HindIII-BamHI fragment containing $90 \%$ of the act 2 gene coding sequence with a 1.8-kb fragment containing the ura4 $4^{+}$gene (13). A $h^{+N} / h^{+N}$ ura4-D18/ura4-D18 ade6-210/ade6-216 leu1-32/leu1-32 diploid strain (SP826) was transformed to Ura ${ }^{+}$with a $4.8-\mathrm{kb}$ EcoRI-HindIII fragment excised from this plasmid. Southern hybridization to DNA from the transformants indicated that the appropriate homologous recombination had occurred in 10 of 16 cases. A $h^{+N} / h^{90}$ strain capable of high-frequency sporulation was selected from one of the original act2::ura4 ${ }^{+}$recombinants by iodine staining of replica plates and was used for tetrad analysis.
(reviewed in ref. 8). Actin in yeast is localized in both patches and cables, the former being associated with regions of cell growth including the bud neck of budding yeast and the cytokinetic septum of fission yeast (8-10). In this study we have characterized a yeast gene (act2) encoding an actin-like protein that, at present, is the most distantly related member of the known actin gene family. ${ }^{\dagger}$

## MATERIALS AND METHODS

Cloning of act2. A Sch. pombe cDNA library in Lambda ZAP (Stratagene) was kindly provided by Dallan Young (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). Oligodeoxynucleotides YTM. 15 (5'-AAAGGCAACCATG-GACAAAAT- $3^{\prime}$ ) and YTM. 13 ( $5^{\prime}$-ACAAAAGTTGATCA CAAGTT- $3^{\prime}$ ) were synthesized with the intention of cloning a yeast tropomyosin cDNA by amplification with the polymerase chain reaction (11). DNA amplification mixtures ( 50 $\mu \mathrm{l}$ ) consisted of $1 \mu \mathrm{~g}$ of cDNA library, 50 pmol of each primer, buffer, dNTPs, and Taq polymerase as recommended by Perkin-Elmer/Cetus and were incubated for 35 cycles of 1.5 min at $94^{\circ} \mathrm{C}, 3 \mathrm{~min}$ at $45^{\circ} \mathrm{C}$, and 2 min at $72^{\circ} \mathrm{C}$. A $660-$ nucleotide fragment was obtained that, when cloned into pUC19 and sequenced with Sequenase (United States Bio-

[^1]

Fig. 2. Nucleotide sequence and encoded amino acid sequence of the act2 gene. The gene possesses five exons that encode a total of 427 amino acids. The four introns are written in lowercase type, and the branch point and splicing signals are underlined.
chemical), was found to encode an amino acid sequence similar to that of actin. When the sequence of the full-length act 2 cDNA (obtained through plaque hybridization from the above library) was later determined, both oligonucleotides were found to have primed out of frame so that no similarity between actin and tropomyosin should be inferred. The act2 gene was selected through colony hybridization from a Sch. pombe genomic library in the pWH5 vector (12).

## RESULTS AND DISCUSSION

We first identified a cDNA encoding part of an actin-like protein sequence through selective amplification of a $S c h$. pombe cDNA library with the polymerase chain reaction (11). The partial cDNA was cloned into the pUC19 vector and used to detect a full-length cDNA from the same library. This cDNA was then used to screen a Sch. pombe genomic library in the plasmid vector pWH5 (12) (Fig. 1). The comparison of


```
MAD-KQT-VVVDNGSGVVKAGFSGGEDARPRAUFRSIIGRPNNVSMLIG
MAESESPA IVIDNGSGMCKAGIAGDDAPRAAFPSIVGRPKMPGIMVG
```



```
MADDED IQP IVGDNGTGMVKAGFAGDDAPRAVFPS IVGRPRLHTGVMMG
MADED IQP IVCDNGTGMVKAGF AGDDAPRANFPSIVGRPRHTGVMVG-
MEE-EIAALVIDNGSGMCKAGFAGDDADRAVFPSIVGRPGHOEVVMVG-
```




```
*)
RERCPEZLPKP-AF IGKBF-PG#HELTFNNSIMMCDNDVRROPYNNIVISGGITMFPGIARRISKEVV-
```



```
RERCPEALFKP-ALIGLDEAPGFHEMTFQS INKCDIDVRRDLIYGNIVLSGGTTMFKNLPERIGMEII
```



```
M,
```

Fig. 3. Comparison of the amino acid sequences of act2 (ac2) and actins from diverse species: The original alignments were made with FASTDB using a McLachlan structure genetic matrix and a window size of 60 (15). Minor modifications were made manually. Amino acids are depicted by the conventional one-letter code. Residues that are similar or identical to the act 2 sequence are shaded. Similarity is defined as ( $\mathrm{E} / \mathrm{D}$ ), $(N / Q),(K / R),(A / G),(S / T),(Y / F)$, and (L/I/M/V). The residues in actin that interact with the nucleotide ATP ( n ), $\mathrm{Ca}^{2+}(\mathrm{c})$, and myosin ( m ) and that are involved in actin-actin interactions (a) as determined by x-ray crystallography and other methods $(16,17)$ are indicated above the act 2 sequence. The regions corresponding to the four subdomains of actin are indicated below the actin sequences. oxy, Oxytricha nova macronuclear actin (18); tet, Tetrahymena thermophila actin (19); try, Trypanosoma brucei actin (20); mai, maize actin (21); pom, Sch. pombe actin (6); hum, human cytoplasmic $\boldsymbol{\gamma}$-actin (22). Gaps required for alignment are indicated by dashes.
cDNA and genomic sequences revealed the presence of four introns in the act2 gene (Fig. 2). Introns 1, 3, and 4 have sequences that match the consensus for $5^{\prime}$ and $3^{\prime}$ splice sites and the branch point in Sch. pombe (14). Intron 2 is unusual in having a $T$ in the first position of the CTA/GAC/T branch-point consensus sequence. The act 2 introns are located in the codons for amino acids Asn-12 and Gly-20, and after codons encoding for Val-196 and Thr-396. According to our alignment (Fig. 3), the corresponding positions in actins are at amino acids $12,20,164$, and 349 . Although a total of $>30$ intron positions have been described in actin genes from a variety of organisms, none is equivalent to those found in the act2 gene (23). Dibb and Newman (24) have proposed that most introns were not present in the ancestral actin gene but have been gained during the evolution of the actin gene family through insertion into a consensus sequence, C/AAGR. While the act 2 introns do not appear to have inserted within this sequence, the growing number of intron positions in actin genes supports the hypothesis that insertions are likely to have occurred. From analysis of multiple cDNAs and Northern blotting (data not shown), two alternative polyadenylylation sites were identified in the act2 gene, 26 and 165 base pairs downstream of the UAA stop codon. The majority of act 2 mRNAs use the downstream site.

In order to determine whether or not act 2 is essential for viability, a null allele of the gene was obtained by the one-step gene disruption method (25) (Fig. 1). Following gene disruption in the diploid strain SP826, tetrad analysis of 30 asci revealed 24 cases of $2: 2$ segregation where two spores developed into Ura ${ }^{-}$colonies and two spores failed to germinate, 5 cases with one colony, and 1 with three colonies. No Ura ${ }^{+}$colonies were recovered from any of the tetrads analyzed. These results indicate that the presence of an intact act 2 gene is essential for germination to occur in Sch. pombe. In addition, they indicate that act 2 is likely to have a function distinct from that of act1. They do not, however, eliminate the remote possibility that overexpression of act 1 may compensate for loss of act2 expression.

The act2 cDNA has a 427 -amino acid open reading frame that encodes a protein with a calculated $M_{r}$ of 47,376 and pI
of 5.7, compared with a typical actin with $M_{\mathrm{r}}$ of 42,000 and pI of 5.2. We have estimated the abundance of act2 through codon usage analysis. In Sch. pombe, highly expressed proteins tend to be encoded by different and fewer codon types than low-abundance proteins. These differences are measured in terms of a codon bias index (CBI) (26). We analyzed the codon bias of both actl and act2 genes from Sch. pombe and found that the actl gene $(\mathrm{CBI}=0.70)$ had a value similar to those of genes encoding highly expressed enzymes such as triose-phosphate isomerase, while act 2 $(C B I=0.18)$ was similar to genes encoding regulatory proteins such as cdc25 and weel. These results indicate that act 2 is unlikely to represent a large portion of the actin-based structures detected in yeast by phalloidin or by anti-actin antibodies (10).

When the act 2 amino acid sequence was compared with the National Biomedical Research Foundation Protein Identification Resource (version 24) and GenBank (version 64) data bases, the large group of actin sequences possessed alignment scores significantly above any other in the data base (Fig. 3). Amino acid identity to actins ranged from $35 \%$ to $40 \%$ (based on the length of actin) whereas similarity as defined in Fig. 3 was $50-55 \%$. The act2 protein was no more similar to Sch. pombe actin (act1) than to actins from a variety of other organisms, supporting the possibility that it diverged from actin early in eukaryotic evolution and will be found in other species. The protozoan actins are among the most divergent, possessing only $70-75 \%$ amino acid sequence identity to actins from animals, plants, and fungi. However, they have the same level of identity with act2 ( $35-40 \%$ ) as the other actins. This observation is also in agreement with an ancient divergence of the act2 and actin genes.

Although act 2 and actin have clearly evolved from a common ancestral gene, it remains to be determined whether these proteins share any functional properties. The elucidation of the x-ray crystallographic structure of rabbit skeletal muscle actin (16) has allowed us to make some theoretical predictions regarding possible functional similarities and differences between these proteins. Actin possesses two
domains each containing two subdomains. Subdomains 1 and 3 each contain a five-stranded $\beta$-sheet that includes a $\beta \alpha \beta$ unit. These structures are largely responsible for the ability of actin to bind ATP and divalent metal cation. Alignment of the amino acid sequence of act 2 with that of actin indicates that residues which interact with ATP and metal ion are highly conserved with the possible exception of residue 16, which has a substitution of tyrosine for methionine. The essential structural features of subdomains 1 and 3 also appear to be conserved, indicating that act2 is likely to possess ATPase activity (Fig. 3). The residues thought to be involved in actin-actin contacts (17) are generally not highly conserved in act2. Eight of nine predicted actin-actin interaction sites exhibit at least one nonconservative substitution or insertion in act2 and, in particular, a large insertion near the amino terminus disrupts a region thought to be important in filament formation (Fig. 3). Other actins from diverse species, including that of Tetrahymena, which copolymerizes with vertebrate muscle actin (27), have far greater conservation of residues necessary for actin-actin interaction. These observations do not, however, rule out the possibility that act2 is incorporated into filaments. Among the sites in actin that are thought to interact with myosin, those near the carboxyl terminus are conserved in act2, whereas the aminoterminal sites are divergent. It will be interesting to determine whether act 2 interacts with a myosin isoform.

Several proteins that interact with actin, including the capping/severing proteins gelsolin, villin, fragmin, and severin, share short regions of sequence similarity with actin (reviewed in ref. 28). While act2 is considerably more similar to actin than these proteins, its apparent low abundance makes it a candidate for an actin capping/severing factor. It is also possible that act2 has a function independent of actin. Despite lacking obvious sequence similarity, actin and the ATPase domain of $70-\mathrm{kDa}$ heat shock cognate protein have nearly identical structures (29). It is thought that these proteins, along with hexokinase, may define a class of proteins that use the energy generated from cleavage of ATP to drive conformational change.

We thank Dallan Young for the Sch. pombe cDNA library, David Beach for the Sch. pombe genomic library and strain SP826, Tomohiro Matsumoto for advice on working with yeast, and David Frendeway, Mark Pittenger, and Fred Southwick for comments on the manuscript. We also wish to acknowledge Madeline Wisnewski, Phil Renna, and Jim Duffy for help with the preparation of the manuscript. This work was supported by National Institutes of Health Grants RO1-GM43049 and PO1-CA46370 (to D.M.H.) and American Cancer Society Institutional Grant IRG-153G (to J.P.L.M.). D.M.H. is an Established Investigator of the American Heart Association.

1. Pollard, T. D. \& Cooper, J. A. (1986) Annu. Rev. Biochem. 55, 987-1035.
2. Meagher, R. B. \& McLean, B. G. (1990) Cell Motil. Cytoskeleton 16, 164-166.
3. Rubenstein, P. A. (1990) BioEssays 12, 309-315.
4. DeNofrio, D., Hoock, T. C. \& Herman, I. M. (1989) J. Cell Biol. 109, 191-202.
5. Ng, R. \& Abelson, J. (1980) Proc. Natl. Acad. Sci. USA 77, 3912-3916.
6. Mertins, P. \& Gallwitz, D. (1987) Nucleic Acids Res. 15, 7369-7379.
7. Shortle, D., Haber, J. E. \& Botstein, D. (1982) Science 217, 371-373.
8. Drubin, D. (1990) Cell Motil. Cytoskeleton 15, 7-11.
9. Adams, A. E. M. \& Pringle, J. R. (1984) J. Cell Biol. 98, 934-945.
10. Alfa, C. E. \& Hyams, J. S. (1990) J. Cell Sci. 96, 71-77.
11. Saikai, R. K., Gelfand, D. H., Stoffel, S., Scharf, S., Higuchi, R., Horn, G. T., Mullis, K. B. \& Erlich, H. A. (1988) Science 239, 487-491.
12. Wright, A., Maundrell, K., Heyer, W. D., Beach, D. \& Nurse, P. (1986) Plasmid 15, 156-158.
13. Grimm, C., Kohli, J., Murry, J. \& Maundrell, K. (1988) Mol. Gen. Genet. 215, 81-86.
14. Mertins, P. \& Gallwitz, D. (1987) EMBO J. 6, 1757-1763.
15. Brutlag, D. L., Dautricourt, J. P., Marlik, S. \& Relph, J. (1990) Comput. Appl. Biosci. 6, 237-245.
16. Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. \& Holmes, K. C. (1990) Nature (London) 347, 37-44.
17. Holmes, K. C., Popp, D., Gebhard, W. \& Kabsch, W. (1990) Nature (London) 347, 44-49.
18. Greslin, A. F., Loukin, S. H., Oka, Y. \& Prescott, D. M. (1988) DNA 7, 529-536.
19. Cupples, C. G. \& Pearlman, R. E. (1986) Proc. Natl. Acad. Sci. USA 83, 5160-5164.
20. Ben Amar, M. F., Pays, A., Tebabi, P., Dero, B., Seebeck, T., Steinert, M. \& Pays, E. (1988) Mol. Cell. Biol. 8, 2166-2176.
21. Shah, D. M., Hightower, R. C. \& Meagher, R. B. (1983) J. Mol. Appl. Genet. 2, 111-126.
22. Erba, H. P., Gunning, P. \& Kedes, L. (1986) Nucleic Acids Res. 14, 5275-5294.
23. Cresnar, B., Mages, W., Muller, K., Salbaum, J. M. \& Schmitt, R. (1990) Curr. Genet. 18, 337-346.
24. Dibb, N. J. \& Newman, A. J. (1989) EMBO J. 8, 2015-2021.
25. Rothstein, R. (1983) Methods Enzymol. 101, 202-211.
26. Russell, P. (1989) in Molecular Biology of the Fission Yeast, eds. Nasim, A., Young, P. \& Johnson, B. F. (Academic, San Diego), pp. 244-267.
27. Hirono, M., Tanaka, R. \& Watanabe, Y. (1990) J. Biochem. 107, 32-36.
28. Tellam, R. L., Morton, D. J. \& Clarke, F. M. (1989) Trends Biochem. Sci. 14, 130-133.
29. Flaherty, K. M., McKay, D. B., Kabsch, W. \& Holmes, K. C. (1991) Proc. Natl. Acad. Sci. USA 88, 5041-5045.

[^0]:    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. $\S 1734$ solely to indicate this fact.

[^1]:    *To whom reprint requests should be addressed.
    ${ }^{\dagger}$ The sequence reported in this paper has been deposited in the GenBank data base (accession no. M81068).

