MOLECULAR AND CELLULAR BIOLOGY, Nov. 1989, p. 5215–5218
0270-7306/89/115215-04$02.00/0

Vol. 9, No. 11

NOTES

Spore Coat Genes SP60 and SP70 of *Dictyostelium discoideum*

KATHY L. FOSNAUGH AND WILLIAM F. LOOMIS*

Center for Molecular Genetics, Department of Biology, University of California, San Diego, La Jolla, California 92093-0322

Received 16 May 1989/Accepted 24 July 1989

We cloned and sequenced the genes for two of the major proteins found in spore coats of *Dictyostelium discoideum*. The predicted translation product of each of these genes starts with a hydrophobic signal sequence that is subsequently cleaved. Expression of these spore coat genes is coordinate in prespore cells.

When *Dictyostelium discoideum* cultures reach a high population density and deplete the environment of food sources, they stop growing and the cells embark on a developmental pathway leading to the formation of fruiting bodies, in which a ball of spores is held several millimeters in the air by a cellular stalk (2, 12). Some of the best-characterized prespore-specific proteins are the spore coat proteins, which are synthesized only in prespore cells after 10 h of development, accumulate in prespore vesicles, and are secreted to make up the outer layer of spores (4–6, 15). We have previously isolated cDNA clones that correspond to two of the spore coat proteins, SP96 and SP60 (6, 11). We now report the isolation of genomic fragments that include the genes coding for SP60 and SP70 and confirm the assignment of the cDNAs to these proteins. The coordinate temporal and cell-type-specific regulation of transcription of these genes as well as the association of their products in disulfide-cross-linked complexes stored within specialized prespore vesicles makes it important to understand the structure of these genes in detail.

A partial *Sau*3A digest of *Dictyostelium* genomic DNA cloned in a derivative of plasmid pAT153 was screened with probes prepared from plasmids pSP1.8 and p2H3 (10, 11). Of about 50,000 colonies, we found several reactive with one or the other of the probes. Those with large inserts were further analyzed. One of the clones recognized by pSP1.8 carries a 5.5-kilobase (kb) insert that we excised and subcloned in plasmid pGEM-1 to generate pSP60 for further mapping and sequencing. A clone recognized by p2H3 carries a 3.5-kb insert that was subcloned in pGEM-1 to generate pSP70. Analysis of fragments generated from these inserts following single and double digestion with various restriction enzymes (13) generated the maps shown in Fig. 1. The restriction sites in the genomic fragments were unambiguously aligned with those in their respective cDNA clones (Fig. 1). When the genomic clones were used to probe restriction enzyme digests of total *Dictyostelium* DNA, unique restriction maps that were consistent with the maps of the cloned portions were generated from the fragment sizes, indicating that the fragments are derived from single-copy genes. pSP60 carries more than a kilobase at either end of the SP60 gene while pSP70 carries 2 kb at the 5′ end but does not extend past the gene (Fig. 1).

RNAs prepared from pSP60 and pSP70 were used to probe Northern blots of RNA taken throughout development (Fig. 2). The results confirmed that these genomic probes recognized the same 1.8- and 2.2-kb mRNAs recognized by their respective cDNA clones and established the orientation of

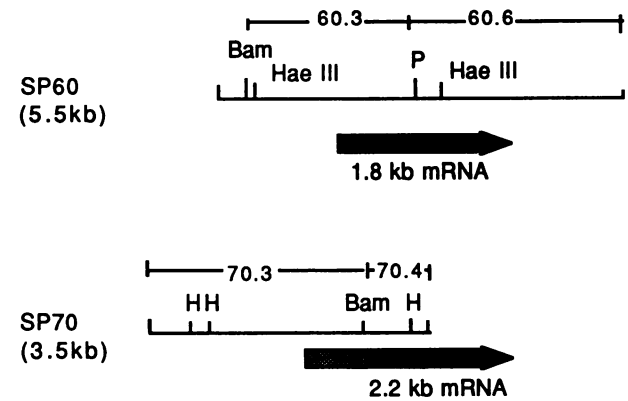

FIG. 1. Restriction site maps of SP60 and SP70. For SP60, the 5.5-kb insert in pAT153 (18) that reacts with pSP1.8 was subcloned in the *Bam*HI site of pGEM-1 and digested with *Sau*3A, *Bam*HI (Bam), *Hae*III, and *Pst*I (P). From single and double digests, a unique map of the sites was constructed. The relative positions of the *Pst*I and *Hae*III sites were congruent with those in the cDNA clone pSP1.8, oriented as shown here (11). The heavily hatched portion of the mRNA was sequenced in pSP1.8. The lightly hatched portion was predicted from the size of the mRNA (1.8 kb). Portions of the genomic clone were subcloned in pGEM-1 for sequencing (p60.3 and p60.6). For SP70, the 3.5-kb insert in pAT153 that reacted with 2H3 was subcloned in the *Bam*HI site of pGEM-1 and digested with *Sau*3A, *Hin*cII (H), and *Bam*HI. A unique map of the sites that was congruent with the sites in 2H3 oriented as shown here was constructed (10; data not shown). The heavily hatched portion of the mRNA was sequenced in the cDNA clone 2H3. The lightly hatched portion was predicted from the size of the mRNA (2.2 kb). Portions of the genomic clone were subcloned in pGEM-1 for sequencing.

* Corresponding author.

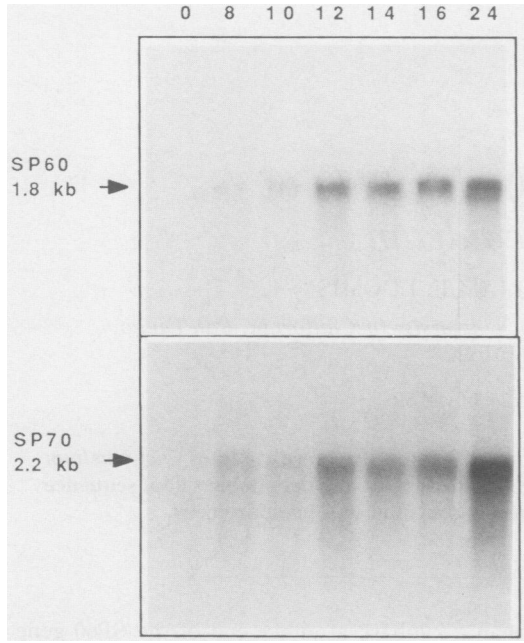

FIG. 2. Developmental accumulation of SP60 and SP70 mRNA. Northern blots of RNA taken at various times (shown at the top of each lane, in hours) during development of *D. discoideum* AX4 were hybridized to riboprobes generated from pSP60.3 or pSP70.3. Constant amounts of RNA were loaded in each well. The blots were also probed with a clone that recognizes actin mRNA to confirm that equal amounts of mRNA were present in all lanes (data not shown).

transcription within each fragment. The mRNAs are not present until hour 8 of development, gradually accumulate until hour 12 and then remain at high levels throughout the remainder of development.

Subclones of pSP60 and pSP70 and their cDNAs were sequenced (17) and found to be identical in the portions that overlapped (Fig. 3 and 4). The sequence of 2H3 differed from that reported by Gomer et al. (10), in which the open reading frame stopped shortly after the *Bam*HI site. Both SP60 and SP70 have short open reading frames initiated at AUG codons preceded by several adenines. An intron with a high A+T content flanked by canonical splice sites is spliced out of both genes, as determined by sequencing cDNAs generated from poly(A)$^+$ RNA of 18-h-developed cells by polymerase chain reaction with primers that flank the region (9, 16). The long open reading frames of both of these genes end at the translation termination signal (UAA) most commonly found in *Dictyostelium* genes (19).

The first exon of both SP60 and SP70 codes for a hydrophobic signal sequence that is cleaved off, leaving the previously determined sequences at the N termini of the processed spore coat proteins (Fig. 5) (7). The predicted products of both genes contain a central region with a high proportion of cysteine residues that could be involved in the intermolecular cross-links that form a complex of SP96, SP70, and SP60 within prespore vesicles (4). Within this region, there are several repeats related to the CPEGHECK sequence that are also found repeated in SP96 (unpublished data), suggesting that all three of these spore coat proteins have evolved from a common precursor.

There are a single glycosylation site (the NAT sequence) and nine serine residues in a row in the N-terminal region of SP70 that are likely to be the sites for the known glycosylation and phosphorylation of this spore coat protein (Fig. 5) (1, 4, 15). Immediately following the serine series in SP70, there are five exact repeats of an 11-amino-acid sequence (GGSTTGSHTTT). Such repeated sequences could facilitate self-polymerization of SP70 as it forms the outer layer of the spore coat. The predicted molecular weights of the

```
AATATTTAAAAATTCTAATAAAAAGTTTTTTATATTTAGTAAATTTGT
AAAATCAATTTGTAACAAAAACTAGTAATTTAAAAAAAAAAAAAACGA
AAAAAAAAAAAA ATGAAGATTTTATCATTATTAGTTGTAGGTGCCTTA
                Met Lys Ile Leu Ser Leu Leu Val Val Gly Ala Leu
TGTATGGGAGGTAAAGTTTATGGTGAAGTTAATGGTGATTGGAACAAT
Cys Met Gly Gly Lys Val Tyr Gly Glu Val Asn Gly Asp Trp Asn Asn
AATGGTGATTGGAACAATAATGGCGATTGGAACAATAATGGTGATTGG
Asn Gly Asp Trp Asn Asn Asn Gly Asp Trp Asn Asn Asn Gly Asp Trp
AACAATAATGGTCCAATTTTAG\GTAAGTTTTATTTTAAACTATATAC
Asn Asn Asn Gly Pro Ile Leu
ATAAAAATAATAAAATAAAAAAATTTGGTATTAATACTTTTTTTTTTT
TTTTTTTTTTTTATTTATTATTATTAAATCTTGACTTTTTTATATAG/
ACTGTTCAACACTTCAATGTCCACCAAGATATCATTGTGAAGTAAAT
Asp Cys Ser Thr Leu Gln Cys Pro Pro Arg Tyr His Cys Glu Val Asn
AATGGTAATAGACAATGTGTTGAAGATCAAATTACATTGCCACCATTT
Asn Gly Asn Arg Gln Cys Val Glu Asp Gln Ile Thr Leu Pro Pro Phe
GATAAATGTGATAATGTCCATTGTCCAAAAGGATTTAATTGCAAATAT
Asp Lys Cys Asp Asn Val His Cys Pro Lys Gly Phe Asn Cys Lys Tyr
GATTGGGAAAAAGATCTTGCTCTTTGTGTTCCATGGAGACCATATCCA
Asp Trp Glu Lys Asp Leu Ala Leu Cys Val Pro Trp Arg Pro Tyr Pro
CCAGTTTGTAGAACTAGATGTCCAGAAGGTCATGAATGTAAAGTTGAT
Pro Val Cys Arg Thr Arg Cys Pro Glu Gly His Glu Cys Lys Val Asp
GAATGGGGTAAAGAATGTTGCGTAAAGATCAAATGTGATGATATTTGT
Glu Trp Gly Lys Glu Cys Cys Val Lys Ile Lys Cys Asp Asp Ile Cys
GACTTGCGTTGTCCAAAGGGTCATGAATGCAAGATCAAACATGATGGT
Asp Leu Arg Cys Pro Lys Gly His Glu Cys Lys Ile Lys His Asp Gly
AGTAAATGCTGTGTCCGTTCATGGAGACCAAGACCACATAAACCACAT
Ser Lys Cys Cys Val Arg Ser Trp Arg Pro Arg Pro His Lys Pro His
CCACGTCCACCAATCTGCAGATTAAGATGTCCACCAGGTCATGAATGC
Pro Arg Pro Pro Ile Cys Arg Leu Arg Cys Pro Pro Gly His Glu Cys
AAACATGATGAACATGGTAAAGAATGTTGCGTTAAAAAACGTCATCAT
Lys His Asp Glu His Gly Lys Glu Cys Cys Val Lys Lys Arg His His
GATAGATGTGACCTCAAATGTAAGAGAGGTTATGAATGTAAAATCAAA
Asp Arg Cys Asp Leu Lys Cys Lys Arg Gly Tyr Glu Cys Lys Ile Lys
CATGATGGCTCCAAATGTTGTGTCAAGAGAACTCCAAAGAGACCATGT
His Asp Gly Ser Lys Cys Cys Val Lys Arg Thr Pro Lys Arg Pro Cys
TGTAAACCAAATTCATGTGCCAGAGATGAAAAATGTGTTGCCACCAAA
Cys Lys Pro Asn Ser Cys Ala Arg Asp Glu Lys Cys Val Ala Thr Lys
GACAGAATCATTTGTGTAAAACCAACTTGTGATAATACTCGTTGTCCT
Asp Arg Ile Ile Cys Val Lys Pro Thr Cys Asp Asn Thr Arg Cys Pro
CCAAATTACCATTGCATATGTGGTGATAAAATTGATGGTGTAAAATGT
Pro Asn Tyr His Cys Ile Cys Gly Asp Lys Ile Asp Gly Val Lys Cys
GTTCCAGATTGTAAGAAAGCAAGATGTGATGATGTCGAATATCCAGAT
Val Pro Asp Cys Lys Lys Ala Arg Cys Asp Asp Val Glu Tyr Pro Asp
TTCCATAGATGTGTTGAAAGACGTGGTGGTATCTTAAGTTGTGAATTT
Phe His Arg Cys Val Glu Arg Arg Gly Gly Ile Leu Ser Cys Glu Phe
GATCCACCAAGACAACCAAGATCCCTTGATTGGGCTGAAAATGAAAAT
Asp Pro Pro Arg Gln Pro Arg Ser Leu Asp Trp Ala Glu Asn Glu Asn
GATGATCGTGACTATGATGATCGTGACTATGATGATGATGAATATGAT
Asp Asp Arg Asp Tyr Asp Asp Arg Asp Tyr Asp Asp Asp Glu Tyr Asp
GGTGTTATGATGGCGGCCGGTGACGGTGACTATGATGGTGATTATGAT
Gly Val Met Met Ala Ala Gly Asp Gly Asp Tyr Asp Gly Asp Tyr Asp
GGTGATTATGATGATGACAATTATTATGGTGATGATGACTATGATAAT
Gly Asp Tyr Asp Asp Asp Asn Tyr Tyr Gly Asp Asp Asp Tyr Asp Asn
GATTGGGACAATGATAATGATTGGGGTAATGATTGGGATAATGATTGG
Asp Trp Asp Asn Asp Asn Asp Trp Gly Asn Asp Trp Asp Asn Asp Trp
GACAATGAAGATGGTGATAATTGGAATGATGATGACTTCCAAGATGCA
Asp Asn Glu Asp Gly Asp Asn Trp Asn Asp Asp Asp Phe Gln Asp Ala
AATGATGAATGGGACTATTAAGCCATTAATTTTTTAAAAAAAAAATAA
Asn Asp Glu Trp Asp Tyr END                          END
TAATAATATTAATACAAATAAGCCATTAATTTTTTAAAAAAAAAATAAT
AATAATATTAATACAAATAATTTTTTAAAAAAAAAAAAAGAAAAAAAA
AAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAATATCTATTAATCG
```

FIG. 3. Sequence of SP60. The transcribed regions of genomic clone pSP60 and the cDNA clone SP1.8 that covers the last 935 bases were sequenced from a set of progressive deletions (Erase-a-Base; Promega Biotec, Madison, Wis.). Analyses of sequences were assisted by the DNA Strider program on a Mac II computer (14). The two sequences are identical except for the poly(A) tail at the 3′ end of SP60. The *Eco*RV site (GATATC) and the central *Pst*I site (CTGCAG) that were used to generate subclones pSP60.3 and pSP60.6 are underlined. The deduced amino acid sequence is indicated below the nucleotide sequence, except in the region that is spliced out of the finished mRNA. The three-letter code for amino acids is used. The reading frame starts at the first ATG codon and ends at the first of two in-phase termination sites. This sequence has been deposited in GenBank (accession no. M26239).

SP60

MKILSLLVVGALCMGGKVYGEVN GDWNNNGDWNNNGDWNNNGDWNNNG
PILDCSTLQCPPRYHCEVNNGNRQCVEDQITLPPFDKCDNVHCPKG
FNCKYDWEKDLALCVPWRPYPPVCRTRCPEGHECKVDEWGKECCV
KIKCDDICDLRCPKGHECKIKHDGSKCCVRSWRPRPHKPHPRPPI
CRLRCPPGHECKHDEHGKECCVKKRHHDRCDLKCKRGYECKIKHD
GSKCCVKRTPKRPCCKPNSCARDEKCVATKDRIICVKPTCDNTRCP
PNYHCICGDKIDGVKCVPDCKKARCDDVEYPDFHRCVERRGGILSC
EFDPPRQPRSLDWAENENDDRDYDDRDYDDDEYDGVMMAAGDGD
YDGDYDGDYDDDNYYGDDDYDNDWDNDNDWGNDWDNDWDNEDG
DNWNDDDFQDANDEWDY

SP70

MRILKLAALSCLLFIAPSLS INCDGLSKDQCEQNFPQCQILTAKSCCGESK
SYCAERDSNDCLASKISCKKDPQGNIYEFWSSCTPSSGFTDFIPSN
ATCSSLNCNAQQMSCKYVQQACHETSCCPDIPQCQIPATGGGPATGS
ATGQGTSGGTPGSCDKVNCPNGYICTIVNQLAVCVSPSSSSSSSST
TGSHTTTGGSTTGSHTTTGGSTTGSHTTTGGSTTGSHTTTGGSTTGS
HTTTGGSTTGPTCGNVNCPRGYHCEVRGSQAVCVADEFDSCANVDCG
SGYHCKNGECIRDKVECDACDQINCPKVIIVYLNQKVAGLVPIRKRH
WQLHPEHCGGRPNKIKVICVPSPKGTCKTVQCPKGYKCKLYADGPT
CVKIEKPKCLTCKDMHCESNGLLCVLTPQKKTDEECCPIIPICINPS
TIAASTIATTTASTRHSTASTIASLVTGTTSGGGGMGSFGGRSDEESS
DPNAILGLFEDDIFWGDNDEYFSDNYRYVDEEDIDEDFNGNDEFEI
DVEGDFQEDDFEEEYAFY

FIG. 5. Amino acid sequences of spore coat proteins SP60 and SP70. The predicted amino acid sequences were deduced from the nucleic acid sequences of pSP60 and pSP70. The single-letter code for amino acids is used. Signal sequences that are removed from the processed spore coat proteins are shown in boldface at the start of SP60 and SP70. The next 20 amino acids in the predicted products are identical to the amino acids found by direct sequencing of purified SP60 and SP70 (7). The prespore motif (CPEGHECK) and its variants are underlined. The asparagine (N-97) in the glycosylation acceptor site of SP70 is in boldface.

primary products of SP60 and SP70 are 46,892 and 54,448, respectively.

Using the clones described in this paper, we have recently shown that cells dissociated from aggregates and shaken at 10^6/ml in buffer rapidly lose SP60 and SP70 mRNA unless 100 μM cyclic AMP is added to the buffer (K. L. Fosnaugh and W. F. Loomis, manuscript in preparation). This seems to be a common characteristic of prespore-specific genes. It has been shown that several of these genes are repressed in prestalk cells by the morphogen 1-(3,5-dichloro-2,6-dihydroxy-4-methoxyphenyl)-1-hexanone (DIF) that accumulates following aggregation (8). We expect that the spore coat genes are also repressed in prestalk cells by DIF acting through a common *trans*-acting protein. However, prespore cells must be insensitive to DIF, since they express the spore

FIG. 4. Sequence of SP70. The transcribed region of genomic clone pSP70 and the cDNA clone 2H3 that starts 36 bases after the only *Bam*HI site (GGATCC) (underlined here) were sequenced. The two sequences are identical in the shared region (664 bases). The cDNA extends 907 bases past the *Sau*3A site at which the genomic clone ends. The last 25 bases of the sequence are likely to be derived from the posttranscriptionally added poly(A) tail. The deduced amino acid sequence is indicated below the nucleotide sequence, except in the region that is spliced out of the finished mRNA. The three-letter code for amino acids is used. The reading frame starts at the first ATG codon and ends at the first of three in-phase termination sites. This sequence has been deposited in GenBank (accession no. M26238).

coat genes even though they are exposed to somewhat higher levels of DIF than prestalk cells are (3). It is possible that cells differentiating along the prespore pathway become insensitive to DIF. Such differential sensitivity to DIF of prespore and prestalk cells forms the basis of a model for cell type divergence that has been proposed recently for *Dictyostelium discoideum* (W. F. Loomis, Prog. Mol. Subcell. Biol., in press).

We thank Joseph Dynes and Richard Firtel for the gift of their partial *Sau*3A genomic *Dictyostelium* bank that contained 5- to 7-kb size-fractionated fragments cloned into the *Bam*HI site of a modified pAT153 cloning vector. We are also grateful to Mark Floyd for technical assistance.

This study was supported by Public Health Service grant GM23822 from the National Institutes of Health.

LITERATURE CITED

1. **Akalehiywot, T., and C.-H. Siu.** 1983. Phosphorylation of spore coat proteins of *Dictyostelium discoideum*. Can. J. Biochem. Cell Biol. **61**:996–1001.
2. **Bonner, J. T.** 1967. The cellular slime molds. Princeton University Press, Princeton, N.J.
3. **Brookman, J., K. Jermyn, and R. R. Kay.** 1987. Nature and distribution of the morphogen DIF in the *Dictyostelium* slug. Development **100**:119–124.
4. **Devine, K., J. Bergmann, and W. F. Loomis.** 1983. The spore coat proteins of *Dictyostelium discoideum* are packaged in prespore vesicles. Dev. Biol. **99**:437–446.
5. **Devine, K., J. Morrissey, and W. F. Loomis.** 1982. Differential synthesis of spore coat proteins in prespore and prestalk cells of *Dictyostelium*. Proc. Natl. Acad. Sci. USA **79**:7361–7365.
6. **Dowds, B. A., and W. F. Loomis.** 1984. Cloning and expression of a cDNA comprising part of the gene coding for a spore coat protein of *Dictyostelium discoideum*. Mol. Cell. Biol. **4**:2273–2278.
7. **Dowds, B. A., and W. F. Loomis.** 1986. Hexapeptide repeat structure in *Dictyostelium* spore coat protein. Biochem. Biophys. Res. Commun. **135**:336–339.
8. **Early, A., and J. G. Williams.** 1988. A *Dictyostelium* prespore-specific gene is transcriptionally repressed by DIF *in vitro*. Development **103**:519–534.
9. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85**:8998–9002.
10. **Gomer, R. H., S. Datta, and R. A. Firtel.** 1986. Cellular and subcellular distribution of a cAMP-regulated prestalk protein and prespore protein in *Dictyostelium discoideum*: a study on the ontogeny of prestalk and prespore cells. J. Cell Biol. **103**:1999–2015.
11. **Hong, C., and W. F. Loomis.** 1988. Regulation of SP60 mRNA during development of *Dictyostelium discoideum*. Biochim. Biophys. Acta **950**:61–65.
12. **Loomis, W. F.** 1975. *Dictyostelium discoideum*: a developmental system. Academic Press, Inc., New York.
13. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Marck, C.** 1988. "DNA Strider": a "C" program for the fast analysis of DNA and protein sequences on the Apple MacIntosh family of computers. Nucleic Acids Res. **16**:1829–1836.
15. **Orlowski, M., and W. F. Loomis.** 1979. Plasma membrane proteins of *Dictyostelium*: the spore coat proteins. Dev. Biol. **71**:297–307.
16. **Sakai, R. K., D. H. Gelfand, B. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
17. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
18. **Twigg, A. J., and D. Sherratt.** 1980. Trans-complementable copy-number mutants of plasmid ColE1. Nature (London) **283**:216–218.
19. **Warrick, H.** 1987. Codon frequency in *Dictyostelium discoideum*. Methods Cell Biol. **28**:497–500.