Vol. 135, No. 1, 1986 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 26, 1986 Pages 336-339

HEXAPEPTIDE REPEAT STRUCTURE IN DICTYOSTELIUM SPORE COAT PROTEIN

Barbara C.A. Dowds* and William F. Loomis+

Department of Biology, B-022,
University of California, San Diego,
La Jolla, California 92093

Received January 23, 1986

SUMMARY: The sequences of the NH_2-termini of two spore coat proteins of Dictyostelium discoideum have been determined. One of them (SP60) consists of perfect hexapeptide repeats of the sequence Gly-Asp-Trp-Asn-Asn-Asx-. The sequence has some homology to the parvovirus capsid protein which does not display periodicity. The NH_2-terminal sequence of the second protein, SP70, contains a modified amino acid in two positions and like SP60 is highly hydrophilic and acidic.

Dictyostelium grows as vegetative cells which aggregate and differentiate into spores and stalks in response to starvation. Spore coat proteins are first synthesised about 14 hours after this developmental program is initiated, at the time when differentiation of prespore and prestalk cells is beginning (1). They are located in prespore vesicles at this stage of development but are found on the outer surface of the spore coat 10 hours later when the spore and stalk cells have matured (2). Spores germinate in the presence of nutrient and the spore coats can be easily separated from the emerged vegetative amoebae at this time. When spore coats are analysed on two-dimensional gels, three major proteins are detected (1). They resolve as broad streaks or clouds but SP96 and SP70 can, however, be separated into doublets on 1-D gels. We have sequenced the proteins in order to determine their heterogeneity and to facilitate synthesis of oligonucleotide probes for screening gene banks. The availability of an SP96-specific antiserum

*Present address: Department of Genetics, Trinity College, Dublin 2, Ireland.

+To whom correspondence should be addressed.

0006-291X/86 $1.50

has facilitated the cloning of an SP96 cDNA (3). However, cloning of the SP70 and SP60 genes has not been possible up to now, due to lack of suitable antisera.

Initial attempts to sequence the 3 spore coat proteins following purification on 2-D gels failed due to blocking of the NH_2-termini. The NH_2 terminus of SP96 was also blocked when it was eluted from 1-D gels and the lower band of the SP96 doublet was found to be a degradation product of SP96 (Dowds, unpublished). However, neither SP70 nor SP60 were blocked when separated on 1-D gels and the NH_2-terminal sequence could be reproducibly determined (Table 1). Identical sequences were found for the upper and lower band of the SP70 doublet suggesting that they differed elsewhere in their sequence or in post translational modification. The two bands have the same mobility on a pH gradient and there is no evidence that one is a degradation product of the other (Dowds, unpublished). However, SP70 is known to be glycosylated (1) and carries a modified amino acid in two locations in the NH_2-terminal sequence (Table 1). We have ruled out hydroxyproline and hydroxylysine as possible identities of this residue.

The NH_2-terminal sequence of SP60 is interesting in that it consists of at least three repeats of the highly hydrophilic sequence Gly Asp Trp Asn Asn Asx (Table 1). There is some indication from the sequencing results that this hexapeptide repeats at least one more time in the NH_2-terminal region. The amino acid composition of SP60 (Table 2) shows an unusually high content of aspartate and asparagine such that the hexapeptide repeat could occur many times in this protein.

Many structural proteins consist of repeating sequences which reflect their various plastic physical properties. Such periodicities are displayed by silk fibroin, keratin, resilin, collagen, elastin and the Plasmodium circumsporozoite proteins (4,5). However, with the exception of the last example, the repeated structures are poorly specified, or, in the case of fibroins, very simple. While most periodic proteins are major structural proteins, a few other proteins are also repeated. Examples are the maize

Table 1. NH_2-terminal amino acid sequences of SP70 and SP60

SP70	SP60
1. Ile	1. Gly
2. Asn	2. Asp
3. X	3. Trp
4. Asp	4. Asn
5. Gly	5. Asn
6. Leu	6. Asn
7. Ser	7. Gly
8. Lys	8. Asp
9. Asp	9. Trp
10. Gln	10. Asn
11. X	11. Asn
12. Glu	12. Asn
13. Gln	13. Gly
14. Asn	14. Asp
15. Phe	15. Trp
16. Pro	16. Asn
17. Gln	17. Asn
18. Ala	18. Asp
19. Gln	19. Gly
20. Ile	20. Asp
21. Leu	21. Trp
	22. Asn

Spore coats ($5X10^9$) were prepared by the method of Lam and Siu (8). They were solubilised and the spore coat proteins were resolved on one-dimensional Laemmli gels (9) and stained briefly in Coomassie Brilliant Blue (10). The proteins were electroeluted (10), precipitated in acetone (11) and the pellet dissolved in 50 µl water. Small aliquots were re-electrophoresed on one-dimensional and two-dimensional gels before the remainder was sequenced on a gas-phase sequenator (Applied Biosystems Model 470A) using the standard "no vac" program supplied by the manufacturer.

The unidentified amino acid (X) at positions 3 and 11 of the SP70 sequence eluted between glycine and alanine under our HPLC conditions.

storage protein, zein (6), protamines and the anti-freeze protein found in certain antarctic fish (4). These repeated amino acid sequences share the property of self-polymerisation and consequently they form strong or flexible structures in most cases. These are the properties expected of a surface structural protein, such as a spore coat protein. It is interesting to note that the NH_2-terminal sequence of SP60 shares homology in 8 amino acids out of a 14 residue stetch of another encapsulating protein, the parvovirus capsid protein (7), (R. Doolittle, personal communication) although the latter has a non-repeated structure.

While the complete sequence of SP70 and SP60 and that of the gene for SP96 will provide more definite insights into the structure of the spore coat proteins, the partial sequences provide glimpses of interesting

Table 2. Amino acid composition of the spore coat proteins (mol %)

Amino acid	SP96	SP70	SP60	Dictyostelium discoideum 18 hr developed cells (12)
Asx	9.2	11.0	23.6	12.5
Thr	12.2	9.4	2.8	5.1
Ser	15.5	11.3	4.1	8.3
Glx	8.1	10.5	7.9	13.6
Pro	7.0	6.9	10.9	8.5
Gly	11.4	12.5	12.3	9.3
Ala	9.7	5.5	2.7	10.2
Cys	1.2	2.2	0.1	1.5
Val	4.3	3.7	3.6	3.3
Met	0.4	1.0	0.2	1.9
Ile	6.6	5.6	3.0	2.4
Leu	3.2	3.3	3.0	7.4
Tyr	1.6	2.0	2.3	4.9
Phe	1.2	2.9	2.1	5.7
His	1.7	3.6	5.4	2.7
Lys	4.1	5.0	8.7	8.2
Arg	2.5	3.1	7.1	4.1

Spore coat proteins(0.5 n moles each) were prepared as described in the Legend to Table I. Phenylthiohydantoid derivatives of amino acid were identified on an IBM cyano HPLC column using a Perkin Elmer series 4 liquid chromatograph and a LC 85B spectrophotometric detector equipped with a 1.4 ul flow cell.

structure. Meanwhile, the short peptide sequences allow us to screen libraries for the cloned genes.

ACKNOWLEDGMENTS

We thank Dr. Matt Williamson for running the sequenator and Dr. Russell Doolittle for searching the NEWAT data bank for sequences homologous to SP70 and SP60. This work was supported by N.I.H. Grant GM23822.

REFERENCES

1. Devine, K.M., Morrissey, J.H. and Loomis, W.F. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 7361-7365.
2. Devine, K.M., Bergmann, J.E. and Loomis, W.F. (1983) Dev. Biol. 99, 437-446.
3. Dowds, B. and Loomis, W.F. (1984) Mol. Cell. Biol. 4, 2273-2278.
4. Doolittle, R.F. (1979) In "The Proteins" (Neurath, H. & Hill, R.L., eds.), Vol. IV, pp. 1-118, Academic Press, New York.
5. Enea, V., Ellis, J., Zavala, F., Arnot, D.E., Asavanich, A., Masuda, A., Quakyi, I. and Nussenzweig, R. (1984) Science 225, 628-629.
6. Pedersen, K., Devereux, J., Wilson, D.R., Sheldon, E. and Larkins, B.A. (1982) Cell 29, 1015-1026.
7. Rhode, S.L. III and Paradiso, P.R. (1983) J. Virol. 45, 173-184.
8. Lam, T.Y. and Siu, C.H. (1981) Dev. Biol. 83, 127-134.
9. Laemmli, U.K. (1970) Nature 226, 680-685.
10. Hunkapiller, M.W., Lujan, E., Ostrander, F. and Hood, L.E. (1983) Methods Enzymol. 91, 227-235.
11. Hager, D.A. and Burgess, R.R. (1980) Anal. Biochem. 109, 76-86.
12. Freeze, H. and Loomis, W.F. (1977) J. Biol. Chem. 252, 820-824.