Two Proteins of the *Dictyostelium* Spore Coat Bind to Cellulose in Vitro†

Yunyan Zhang,‡ Ross D. Brown, Jr.,§ and Christopher M. West*,‡

Department of Anatomy and Cell Biology, University of Florida College of Medicine, and Food Science and Human Nutrition Department, Institute of Food and Agricultural Sciences, Gainesville, Florida 32610

Received April 9, 1998; Revised Manuscript Received May 22, 1998

ABSTRACT: The spore coat of *Dictyostelium* contains nine different proteins and cellulose. Interactions between protein and cellulose were investigated using an in vitro binding assay. Proteins extracted from coats with urea and 2-mercaptoethanol could, after removal of urea by gel filtration, efficiently bind to particles of cellulose (Avicel), but not Sephadex or Sepharose. Two proteins, SP85 and SP35, were enriched in the reconstitution, and they retained their cellulose binding activities after purification by ion exchange chromatography under denaturing conditions to suppress protein−protein interactions. Neither protein exhibited cellulase activity, though under certain conditions SP85 copurified with a cellulase activity which appeared after germination. Amino acid sequencing indicated that SP85 and SP35 are encoded by the previously described *psp*B and *psv*A genes. This was confirmed for SP85 by showing that natural M_r polymorphisms correlated with changes in the number of tetrapeptide-encoding sequence repeats in *psp*B. Using PCR to reconstruct missing elements from the recombinogenic middle region of *psp*B, SP85 was shown to consist of three sequence domains separated by two groups of the tetrapeptide repeats. Expression of partial *psp*B cDNAs in *Escherichia coli* showed that cellulose-binding activity resided in the Cys-rich COOH-terminal domain of SP85. This cellulose-binding activity can explain SP85's ultrastructural colocalization with cellulose in vivo. Amino acid composition and antibody binding data showed that SP35 is derived from the Cys-rich N-terminal region of the previously described *psv*A protein. SP85 and SP35 may link other proteins to cellulose during coat assembly and germination.

In the cellular slime mold *Dictyostelium discoideum*, spores differentiate from prespore cells as they rise up the stalk of the forming fruiting body (*1*). Each spore is surrounded by its own coat, and the collection of spores is held at the top of the stalk, as an aqueous suspension in the fluid interspore matrix, by surface tension. Each coat is composed of 450 fg of nine major protein types, a galactose-rich polysaccharide referred to as a galuran, and about 330 fg of cellulose (refs *2*−*4*, this report). The proteins are susceptible to dissociation by denaturation after reduction of disulfide bonds; i.e., they are not irreversibly covalently cross-linked (*4*, *5*). The *Dictyostelium* coat is a useful experimental system for understanding how glycoproteins and cellulose interact to form, at the cell surface, a protective two-dimensional extracellular layer with defined dimensions and properties of patency, distensibility, water permeability, and impermeability to larger molecules (*3*, *6*−*10*).

Formation of the *Dictyostelium* spore coat begins with the secretion of most, if not all, of its future proteins, which are pre-made and stored together in vesicles in prespore cells (*3*, *7*, *11*). The galuran accompanies these proteins. Co-immunoprecipitation studies have suggested that some of the proteins form a soluble complex before secretion (*11*−*15*). After secretion, the distribution of most of the proteins is unlocalized between neighboring spores until cellulose, emanating from the plasma membrane, is deposited. Thus coats are assembled from the secretory products of several neighboring spores (*16*). After the appearance of cellulose, specific protein domains are found either external to, or internal to and partially overlapping with the cellulose layer, as suggested by sensitivity to proteases and immunogold labeling (*3*, *4*, *17*−*19*). These observations imply that coat protein complexes do not become incorporated into a coat structure until cellulose appears. Protein−cellulose interactions may trap proteins, and also guide cellulose morphogenesis and influence its accumulation. Intermolecular disulfide cross-linking may stabilize the final structure (*4*).

A role for coat proteins as a permeability barrier for large molecules such as other proteins has derived from genetic analysis of four proteins which are associated in some way with the outer layer (*20*−*22*). Disruption of their genes resulted in an increased penetration of the coats by a protein tracer probe (*20*, *21*),[1] demonstrating a functional contribution by each protein to a macromolecular permeability barrier. In each case, the remaining coat proteins, except notably for SP35 in the triple disruption strain TL56, were incorporated into the coat normally (*4*). A contributing role for O-linked sugars on coat glycoproteins has also been inferred from similar analyses of glycosylation mutants (*3*). Since cellulose

† This work was supported in part by Grants MCB-9316897 and MCB-9730036 from the National Science Foundation.

* Author to whom correspondence should be addressed at Department of Anatomy and Cell Biology, 1600 SW Archer Road, University of Florida College of Medicine, Gainesville, FL 32610-0235. Phone: 352-392-3329. Fax: 352-392-3305. E-mail: <westcm@college.med.ufl.edu>.

‡ University of Florida College of Medicine.

§ Institute of Food and Agricultural Sciences.

[1] West, C. M., and Andersen-Davis, T., unpublished data for strain DP87−.

S0006-2960(98)00801-0 CCC: $15.00
Published on Web 07/03/1998

walls of plants are ordinarily permeable to proteins (*23*, *24*), this suggested that a disulfide cross-linked protein complex, apposed to the outer surface of the cellulose layer, constitutes a protein permeability barrier.

To investigate how coat proteins interact with cellulose in *Dictyostelium*, coat protein was selectively extracted from coats and reconstituted with cellulose. Coat proteins were isolated from coats rather than prespore vesicles because a previous study indicated that only a small fraction of coat proteins bound cellulose when isolated from prespore cells (*15*). The present study shows that it was feasible to efficiently renature binding of total coat protein to cellulose and that SP85 and SP35 bound cellulose after partial purification. SP85 and SP35 were found to be encoded by the previously identified *psp*B (14E6; ref *25*) and *psv*A (EB4; ref *26*) genes. After using PCR[2] to fill in the missing sequence from the middle of the *psp*B coding region, SP85 was found to consist of three sequence domains separated by two clusters of tandem tetrapeptide sequence repeats, which may serve as interdomain spacers. Recombinant expression of partial cDNAs in *Escherichia coli* showed that cellulose-binding activity resided in the COOH-terminal domain, which contained cysteine-repeat motifs also present in SP35. Since the full-length p58 product of *psv*A was previously shown to disappear during spore coat formation, SP35 may be formed by proteolytic processing. These findings probably explain the colocalization in vivo of SP85 and cellulose as determined ultrastructurally (*3*, *19*), and suggest a model by which coat proteins associate with the cellulose component of the coat via specialized cellulose-binding proteins which may be regulated.

EXPERIMENTAL PROCEDURES

Cells. Wild-type strains X22, WS380B, WS576, and NC-4 were passaged on lawns of *Klebsiella aerogenes* grown on SM agar plates (*4*). The axenic strain Ax3 was grown on HL-5 medium (*4*). Spores were allowed to form on SM agar plates following depletion of bacteria, and collected by slapping into the lid of the Petri plate. These were recovered in H_2O, filtered to remove stalks, and centrifuged, as described (*4*). The supernatant was saved as the interspore matrix fraction.

Purification of Spore Coats and Extraction with Urea. Coats were isolated either from post-germination cultures or directly from intact (i.e., pre-germination) spores. For germination, spores from 100 10-cm plates were suspended at 2×10^7 mL^{-1} in KP buffer (10 mM potassium phosphate, pH 6.5), activated at 42 °C for 30 min, and allowed to germinate in rotary suspension overnight. Coats were purified by differential and density gradient centrifugation as described (*4*). The culture supernatant was saved as the germination supernatant. For isolation of coats from pre-germination spores, spores from 400 plates were suspended at 7.5×10^8 mL^{-1} in KP, and lysed in a 5 °C French pressure cell at 20 000 psi. Coats were centrifuged at 9500*g* for 5 min and resuspended in KP with or without 0.1% (v/v) NP-40. Coats were then purified as described above. After purification, coats were extracted with 8 M urea, 20 mM Tris-HCl, pH 7.4, with 5% (v/v) 2-ME, 1 mM phenylmethylsulfonyl fluoride, 10 μg of leupeptin mL^{-1}, and 18.5 μg of aprotinin mL^{-1}, for 3 min in a boiling water bath. Extracted proteins and residual coats were separated by centrifugation at $18 \times 10^3 g$ for 20 min. Coats were resuspended and washed by centrifugation in 50 mM NH_4-Ac, pH 7.4. In some cases, the urea extract (supernatant) was dialyzed using Spectrapor cellulose ester membrane (MWCO = 10×10^3) against 50 mM NH_4Ac, pH 7.4, or desalted over a PD10 cartridge (Pharmacia, Piscataway, NJ), which contained 9 mL of Sephadex G-25 equilibrated in 50 mM NH_4Ac, pH 7.4. Soluble protein concentration was estimated by a commercial Coomassie Blue dye-binding assay (Pierce), with BSA as a standard.

Monosaccharide Composition Analysis. Samples were hydrolyzed either by incubation in 0.5 mL of 4 M trifluoroacetic acid for 4 h at 100 °C (mild acid hydrolysis), or by heating for 1 h in 0.5 mL of 100% trifluoroacetic acid, followed by dilution with an equal volume of water and additional incubation at 100 °C for 3 h (strong acid hydrolysis). Mild hydrolysis conditions were typical for monosaccharide analysis of glycoproteins (*27*), and the strong hydrolysis conditions were those used for wood cellulose (*28*). Samples were then taken to dryness in a vacuum centrifuge and reconstituted in 17 mM NaOH. Sugars were analyzed on a Dionex PA-1 column by high-pH anion-exchange chromatography (*27*), using a Dionex ED-40 pulsed amperometric detector.

Ion-Exchange Chromatography. Urea extracts were applied at 22 °C to a TSK DEAE−5PW (8 × 75 mm) column (Toyo Soda, TosoHaas, PA) on an LKB GTi HPLC system at 0.75 mL min^{-1}, preequilibrated in 50 mM NH_4Ac, pH 6.0, with or without 6 M urea as indicated. The column was typically eluted with a gradient of 50 mM NH_4Ac to 1.5 M NH_4Ac, pH 6.0. The DEAE flow-through fraction of some trials was supplemented with PMSF, leupeptin, and aprotinin as above, adjusted to pH 5.0 with H_3PO_4, and applied at 22 °C to a TSK SP-5PW (5 × 5 mm) column (Toyo Soda) at 0.5−0.75 mL min^{-1}, preequilibrated in 6 M urea, 20 mM sodium phosphate, pH 5.0. The column was eluted with a gradient up to 1 M NH_4Ac in the same solution, then equilibrated in 4 M urea, 0.25 M NH_4Ac, 20 mM sodium phosphate, pH 5.0, and eluted further at 0.25 mL min^{-1} with a second gradient from pH 5.0 to pH 8.0 in the same solution. Fractions (0.5 mL) were collected. Ion-exchange columns were regenerated after each use by successive treatment with 0.1 N NaOH and 30% HAc.

Avicel Column Chromatography. The DEAE flow-through fraction from pre-germination coats was concentrated in a Centricon-10 centrifugal ultrafiltration device (Amicon, Waverly, MA) and desalted on a PD-10 cartridge equilibrated in water, to remove urea and 2-ME. Hydrated Avicel (Type PH-101, FMC Corporation, Philadelphia, PA) was de-fined, loaded into columns (1 mL bed volume), and equilibrated in 50 mM NH_4Ac, pH 6.8. Protein was applied, incubated for 1−2 h at 22 °C, and eluted under gravity flow sequentially with 5 mL each of 50 mM NH_4Ac, pH 6.8; H_2O; 8 M urea, 50 mM NH_4Ac, pH 6.8; and 2% (w/v) SDS, 50 mM NH_4Ac, pH 6.8; each eluent contained 5% (v/v) 2-ME.

Centrifugal Cellulose-Binding Assay. Urea extracts were desalted either by gel filtration or dialysis (see above). DEAE fractions were diluted in H_2O so that the NH_4Ac

[2] Abbreviations: EGF, epidermal growth factor; 2-ME, 2-mercaptoethanol; PCR, polymerase chain reaction.

concentration was ≤ 0.1 M. SP-column fractions were concentrated in a Microcon-30 centrifugal ultrafiltration cartridge (Amicon) down to 20−180 μL, and diluted 50−200-fold with 50 mM NH_4Ac, pH 6.8. The protein sample to be tested (10−70 μL), and 300−3000 μg of Avicel, from a 100 mg/mL stock suspension in 50 mM NH_4Ac, pH 6.8, were brought to a final volume of 120 μL with 50 mM NH_4Ac, pH 6.8, and incubated in a 1.5 mL polypropylene microcentrifuge tube on a gyratory shaker for 16 h at 22 °C. After centrifugation at 10000*g* for 1 min, the nonbound fraction was removed with a pipet, and the pelleted bound fraction was washed 3× with 50 mM NH_4Ac, pH 6.8, by centrifugation. The pellet was resuspended in SDS−PAGE sample loading buffer with 5 mM DTT, boiled for 3 min, and centrifuged again. Equivalent proportions of the nonbound and bound (SDS-released) fractions were compared by SDS−PAGE.

Electrophoresis, Western Blotting, Microsequencing, and Composition Analysis. Samples were dried in a vacuum centrifuge to reduce volatile salts and reconstituted in H_2O. Electrophoresis was performed on 7−15% linear gradient SDS−polyacrylamide gels, or two-dimensional O'Farrell gels, followed by staining with Coomassie Brilliant Blue R-250 or silver staining reagents, as described (*4*). For Western blotting, either previously fixed and Coomassie Blue-stained gels equilibrated in 2 × 30 min washes of H_2O and one wash of top buffer (below), or unfixed gels, were electrotransferred to 0.45 μm pore diameter nitrocellulose membranes (Schleicher & Scheull), or Immobilon P^{sq} PVDF membranes (Millipore), in a BioRad TransBlot SD semi-dry transfer cell. The top buffer was 50 mM glycine, 50 mM Tris-HCl, pH 8.5, and 0.1% (w/v) SDS; the bottom buffer was 50 mM glycine, 50 mM Tris-HCl, pH 8.5, prepared in 20% (v/v) MeOH. For antibody probing, nitrocellulose membranes were incubated in primary antibody diluted in 5% nonfat dry milk, followed by alkaline phosphatase-conjugated secondary antibodies and colorimetric detection as described (*4*). For Edman degradation studies, samples were first boiled at 100 °C for 3 min in SDS−PAGE sample buffer (62 mM Tris-HCl, pH 7, 2% (w/v) SDS, 10% (v/v) glycerol) with 50 mM DTT, followed by addition of iodoacetamide to 100 mM and incubation in the dark for 1 h at 22 °C. The SDS−PAGE gel was prerun with 10 mM Na thioglycolate in the upper buffer (*29*). After SDS−PAGE and electroblotting, proteins were stained with Coomassie Blue and destained. Corresponding PVDF membrane fragments were introduced onto an ABI model 494 Procise sequenator. For internal sequencing, the stained gel band was partially dried, rehydrated with 20 mg ml^{-1} CNBr in 70% formic acid, macerated, incubated at 37 °C in the dark for 16 h, and then subjected again to SDS−PAGE using a Tris-tricine buffer system (*30*), followed by electroblotting.

For amino acid composition analysis, protein prepared as for Edman degradation was hydrolyzed in 6 N HCl with norleucine as an internal standard. PTC derivatives were prepared and analyzed as described (*31*).

DNA Analysis. pG219 and p14E6Pst73#1, which contained *psp*B DNA (*25*), were generously provided by R. A. Firtel and J. Coffman-Powell and amplified in *E. coli* strain DH5α. For gDNA, 10^5−10^6 cells growing axenically or on bacteria were washed 2× with water. The cell pellet was resuspended in 50 mM KCl, 1.5 mM $MgCl_2$, 20 mM Tris-HCl, pH 8.4, at 10^4 cells per μL. NP-40 (10% v/v) and 50 mg/mL proteinase K were added to final concentrations of 0.5% and 0.05 mg/mL, respectively, incubated at 56 °C for 45 min, and then incubated at 95 °C for 10 min. Oligonucleotides were synthesized at the UF ICBR DNA Synthesis lab, and used to prime the PCR amplification of plasmid DNA or gDNA from growing *Dictyostelium* cells. The primers were calculated to have melting points in the range of 68−72 °C, and contained, at their 5′-ends, ten-nucleotide (10-nt) extensions containing a *Bam*HI-site. Template (2−3 μL) was diluted to 15−20 μL in a cocktail which contained, as final concentrations, 50 mM KCl, 1.5 mM $MgCl_2$, 20 mM Tris-HCl, pH 8.4, 0.5 mM dNTPs, 0.5 μM oligonucleotide primers, and 0.5 unit Taq polymerase (Promega). Amplification used the following program: 94 °C, 3 min; [94 °C, 45 s; 49 °C, 90 s; 72 °C, 2−3 min], 30−35 cycles; 72 °C, 7 min. PCR products were separated on a 1% GTG Seakem agarose (FMC Products) gel, purified with a Qiagen gel extraction kit, and sequenced using PCR primers in ABI Prism Dye Terminator cycle sequencing protocols. The fluorescently labeled extension products were analyzed on an Applied Biosystems (Perkin-Elmer Corp.) Model 373 Stretch DNA Sequencer operated by the UF ICBR DNA Sequencing lab.

pspB Expression in E. coli. cDNAs were generated by PCR from *psp*B cDNA in pSP73 (p14E6Pst73#1) and subcloned into pET11a (Novagen), which drives expression in the cytoplasm with an N-terminal T7-epitope tag. After confirmatory sequencing, the plasmids were transformed into *E. coli* strain BL21(DE3). Protein expression was induced in clones in LB medium grown at 30 °C at $OD_{600} = 0.6$, by the addition of 1 mM IPTG. Induced cells were sonically lysed and centrifuged at 35000*g* × 15 min at 4 °C. Insoluble fractions were solubilized in urea/DTT and diluted in the absence or presence of glutathione as done for other disulfide-containing proteins (*32*, *33*). The urea solutions were then desalted on PD-10 columns as above.

Viscometric Cellulase Assay. Extracts and column fractions were incubated with calibrated volumes of carboxymethylcellulose at pH 5.0 at 37 °C, as described (*34*). Less activity was detected at pH 7.0.[3] The time required to drain 1 mL of the solution was determined in seconds. Shorter drain times correlate with decreased viscosity of the carboxymethylcellulose solution and increased endoglucanase activity.

In Situ Cellulase Assay in Gels. Extracts and column fractions were subjected to SDS−PAGE as described above, except that samples were heated at 37 °C, as solubilization at 100 °C resulted in much lower activity.[3] After electrophoresis at 22 °C, gels were equilibrated in 50 mM potassium phosphate, pH 5.0, for 3 × 10 min. Cellulase activity were detected by incubating the gel in 1% (w/v) carboxymethylcellulose in 50 mM potassium phosphate, pH 5.0, for 2 h at 37 °C, rinsing the gel with 50 mM potassium phosphate, pH 5.0, twice for 10 min each, and staining with 1% (w/v) Congo Red in the same buffer for 20 min, followed by destaining with 1 M NaCl until bands were clearly revealed, as described (*35*). Gels were photographed through a type 40A Kodak Wratten filter. Proteins were nonspecifically

[3] Data not shown.

Table 1: Monosaccharide Composition Analyses of Post-Germination Coat Extracts[a]

	pmol sugar 10^6 coats^{-1} (fg coat^{-1})		
	before urea extraction	after urea extraction	
sugar type		coat residue	released material
D-glucose	630	1800 (330)	1.6
D-galactose	25 (4.5)	30	1.2
D-galactosamine[b]	14 (3.1)	13	5.0
D-glucosamine[b]	59 (13)	7.6	53
D-mannose/xylose[c]	13 (2.3)	15	0.95

[a] Samples were degraded in 4 M TFA, or 100% TFA followed by dilution in H_2O (D-glucose only), and aliquots were chromatographed on a Dionex PA-1 column and quantitated with a pulsed amperometric detector. [b] Conditions of hydrolysis lead to de-*N*-acetylation of amino sugars. [c] Mannose and xylose were not differentiated in the analysis.

stained by Congo Red, and cellulase activity was detected as regions of negative staining.

RESULTS

Isolation and Extraction of Coats from Pre-germination and Germinated Spores. Coats released naturally from spores by germination were purified by density gradient centrifugation. To analyze coats before they were altered by germination, spores were lysed in a French pressure cell and coats were purified exactly as for their germinated counterparts. These coat preparations were homogeneous and indistinguishable from germinated coats under phase contrast optics.[3] Examination of sections of embedded material in the transmission electron microscope showed that the coats maintained their three-layered organization with cellulose in the middle, and nearly always retained their attachment to the plasma membrane,[3] as previously observed (*36*). The protein compositions of the pre-germination and post-germination coats were similar within the resolution of SDS−PAGE analysis,[3] indicating that any individual plasma membrane proteins or potential contaminating proteins lacked the abundance of the major coat proteins.

Treatment with 8 M urea and 5% 2-ME at 100 °C followed by centrifugation was previously shown to solubilize 85% of the protein of post-germination coats (*4*). Monosaccharide composition analysis before and after urea extraction was performed to determine whether the polysaccharides of the coat were also extracted. Except for glucose, the monosaccharide composition of the intact coat was in reasonable agreement with the content of the sum of the urea-released material and the urea-residue (Table 1). Comparison of the galactose content of the urea extract compared to the extracted residue showed that less than 5% of total coat galactose was released, indicating that the previously described galuran, a galactose-rich polysaccharide (*3*), was not solubilized. Similarly, most galactosamine (72%) was not extracted, consistent with the presence of this sugar in the galuran. Most mannose/xylose (94%) was retained in the extracted coat residue, whereas most glucosamine (87%) was extracted, suggesting that this amino sugar is largely associated with protein, possibly as GlcNAc. Strong acid hydrolysis yielded >50 times more glucose than did mild acid hydrolysis, consistent with earlier studies that glucose is present as cellulose (*3*). Less than 0.25% of total glucose was extracted by hot urea/2-ME, showing that cellulose also

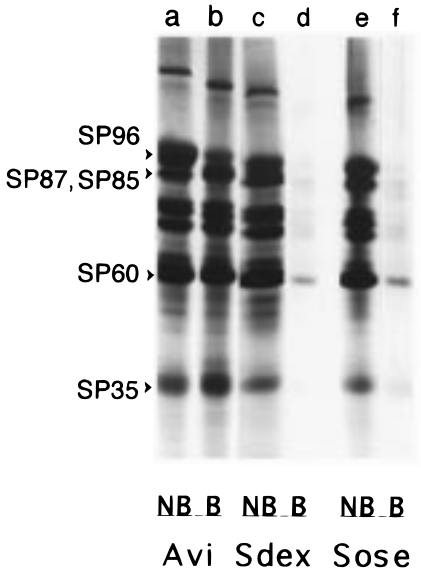

FIGURE 1: Cellulose binding of a urea extract from pre-germination spore coats. Extracted protein was separated from urea by gel filtration, and 30 μg of total protein was incubated with 3 mg of Avicel (Avi), lanes a,b; Sephadex G-75 (Sdex), lanes c,d; or Sepharose-4B (Sose), lanes e,f. Nonbound (NB) and bound (B) fractions were separated by centrifugation and then run on the same SDS−PAGE gel followed by Coomassie Blue staining. Image processing was identical for sister NB and B fractions. Positions of key proteins are indicated.

remained insoluble. The total level of cellulose as glucose was 330 fg coat^{-1}, compared to 450 fg protein coat^{-1} (*4*). About one-third of this amount of glucose was measured in the coats before extraction. The reason for this discrepancy is not known, but may have resulted from interference by other substances in the nonextracted coat.

When urea and 2-ME were removed by dialysis or gel filtration on a PD10 column, >90% protein remained soluble after centrifugation at 100000*g*.[4] Thus, coat proteins do not re-form an insoluble complex under these conditions. Reconstitution of an insoluble complex might require cellulose, which is investigated below.

Extracted Coat Proteins Reassociated with Cellulose after Renaturation. To test the potential requirement of cellulose for protein to form an insoluble complex, urea extracts were desalted by gel filtration to remove urea, and incubated with Avicel, which consists of homogeneous, 5−10 μm diameter particles of crystalline and amorphous cellulose. After centrifugation, the bound and nonbound fractions were analyzed by SDS−PAGE and staining with Coomassie Blue. Proteins were compared from both pre-germination and post-germination coats.

After incubating 30 μg of desalted protein from pre-germination coats with 3 mg of Avicel, comparison of the bound and nonbound fractions showed that all major coat proteins bound partially (Figure 1a,b). Binding of the coat extract to Avicel was specific, as binding to polysaccharide particles of Sephadex G-75 or Sepharose CL-4B was negligible (Figure 1c−f). Bands at M_r positions 97 000 and 35 000, relative to M_r standards included on the same gel, bound somewhat more efficiently to Avicel than the other coat proteins when the band intensities are compared between

[4] Mao, J., and West, C. M., unpublished data.

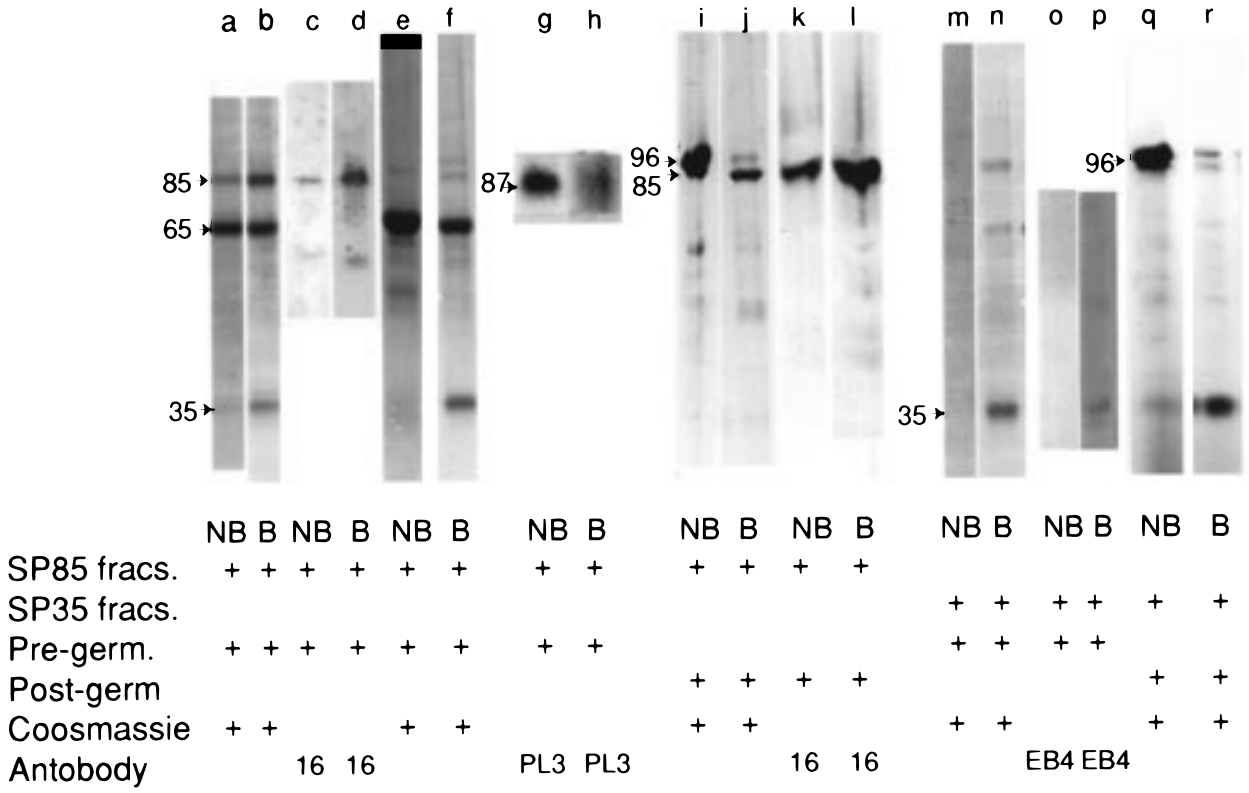

FIGURE 2: Cellulose binding of selected FPLC fractions from spore coats. Urea extracts were fractionated by FPLC as indicated, and individual fractions were tested for binding to Avicel as in Figure 1. Lanes a−h: The pre-germination urea extract analyzed in Figure 1 was subjected to DEAE−FPLC and the flow-through was applied to an SP-column and eluted with salt and then with increasing pH. Avicel-binding of selected fractions enriched in SP85 (a−d), SP65 (e,f), and SP87 (g,h) is shown, with detection by Coomassie Blue to show total protein in the fraction or Western blotting to confirm protein identity as indicated. Lanes i−l: A post-germination coat urea extract was dialyzed and subjected to DEAE−FPLC. SP35 is missing because it bound to the dialysis membrane.[5] Avicel-binding of a fraction containing SP96 and SP85 (they copurified when from post-germination coats) is shown both by Coomassie Blue (i,j) and Western blotting for SP85 (k,l). Lanes m−p: Avicel-binding of a fraction from the experiment shown in lanes a−h which was enriched in SP35. Lanes q,r: Avicel-binding of a fraction enriched in SP35 and SP96, from a urea-extract of post-germination coats applied directly to a DEAE column. Positions of key proteins are indicated.

the bound and nonbound fractions (Figure 1a,b). These proteins are shown below to be the previously described SP85 and SP35, respectively. The apparent M_r of SP85 differs from the value implied by its name (85 000) because the migration of this glycoprotein, whose predicted polypeptide mass is only 57 094 (see below), depends on the gel system used and the strain from which it is isolated (*4*, *15*). The coat protein names used here are as in refs *11* and *17*. The abundant outer-layer coat protein SP96, whose identity was confirmed using mAb 83.5 (*3*),[3] bound least efficiently, which mimics its normal inefficient incorporation during coat formation, being found mostly in the interspore matrix (*4*, *5*). Similar results were obtained for urea extracts of germinated coats.[3]

SP85 and SP35 Bound Cellulose after Partial Purification from Coats. To determine whether individual proteins such as SP85 and SP35 have intrinsic binding activity or whether binding is dependent upon a mixture of all proteins as suggested by a previous study (*15*), proteins were partially purified and assayed for Avicel binding.

Based on 2-D gel electrophoresis (*4*), DEAE chromatography at pH 6 was predicted to adsorb all major coat proteins except for the neutral species SP87, SP85, SP65, and SP35. As expected, bands corresponding to these proteins appeared in the flow-through fraction from the pre-germination coat extract applied to a column equilibrated without urea.[3] SP96, detected with mAb 83.5, eluted at 0.6−0.7 M NH_4Ac,[3] consistent with its acidic apparent pI (*3*). Elution of the other coat proteins required 6 M urea[3] (*3*, *37*).

The DEAE flow-through fraction was further resolved on an SP-cation exchange column equilibrated in 6 M urea at pH 5.0. The column was eluted first with a salt gradient in 6 M urea, followed by an ascending pH gradient in 4 M urea. SDS−PAGE analysis showed that SP35, SP65, SP85, and SP87 eluted sequentially in partially overlapping fashion in the salt gradient.[3] Similar protein bands also eluted near the end of the pH gradient. Without urea in the elution gradients or at pH 6.0, these proteins eluted more broadly and were extensively degraded.[3] Extensive degradation was also observed if the original urea extract was applied to the SP-column before passage over the DEAE column.[3]

To test binding of these proteins to Avicel, individual FPLC fractions were concentrated in 0.5 mL Microcon centrifugal ultrafiltration cartridges so that urea was diluted to 50−100 mM in the assay. Minimal binding to the cellulosic ultrafiltration membrane occurred in 6 M urea. In fractions from pre-germination coat extracts which were highly enriched in SP85, SP85 bound cellulose efficiently, as shown by the greater amount of the SP85 band in the bound fraction (Figure 2, lane b) compared to the corresponding nonbound fraction (lane a). The bound:nonbound ratio was a more valid measure of binding than the absolute

Table 2: NH_2-Terminal and Internal Sequences of Coat Protein Fractions[a]

protein band	treatment	NH_2-terminal sequence found	matching DNA sequence (Gene/Acc.#)
SP85/SP87	CNBr	(M_{158})[b]PLQLGYGANGKNGDNGISV (2)[c] (W_{178})LIYGYTIVDI (Q_{307})VFCVPDCP	*psp*B (14E6)/S72639
	CNBr	(A_{105})KIKECA (M_{202})VNDVATCLASTTGGXGLPG (D_{274})XXXDVDCPDGFXXEXKD (D_{314})PCKDVTCPDGFHCECKDGKT	*psp*D (PL3)/U25144
SP35	CNBr	(D_{22})SVNFGRG (N_{25})FGRGGGWGSGNQGVLCGTHYCPPGSTCESKH (2) (M_{166})VNGYLQCLFSNGTAPFACGLQTCVPPQL (2)	*psv*A (EB4)/P08798
	none	(D_{22})SVNFGR	

[a] Protein bands from SDS−PAGE gels were either electroblotted (treatment = none), or treated with CNBr in 70% formic acid and the peptides purified again by SDS−PAGE followed by electroblotting. Genes corresponding to the amino acid sequences found are listed in the right-hand column. [b] Residue in parentheses is deduced from the gene sequence; its location is relative to the start Met. [c] Sequence obtained twice, from different SDS−PAGE bands.

amount of bound material, because most coat proteins exhibited a basal level of binding to Avicel and other polysaccharide matrixes which was proportionate to the total amount of the protein. The identity of the Coomassie Blue-stained band as SP85 was confirmed by Western blot detection with mAb 16.1 (lanes c,d). Other fractions enriched in SP65 (lanes e,f) or SP87 (lanes g,h) showed that these proteins bound inefficiently. However, in fractions which contained both SP65 and SP85, SP65 binding to cellulose was also detected (lanes a,b), which may indicate an interaction between SP65 and SP85. SP85 isolated from post-germination coats also bound cellulose efficiently (lanes k,l), despite the fact that in this case SP85 bound to the DEAE column and partially coeluted with SP96. Coomassie Blue analysis of binding in this fraction showed that SP96 did not bind cellulose (lanes i,j), while the slightly more rapidly migrating SP85 band, though overlapping with SP96 in the unbound lane (lane i), was selectively enriched in the bound fraction. Another fraction containing SP96 also did not bind cellulose (lanes q,r).

SP35 from fractions of pre-germination coats which were highly enriched in this protein bound efficiently to Avicel, as determined by Coomassie Blue staining (lanes m,n). Similar results were observed in fractions from post-germination coats (lanes q,r) which were enriched in SP35, which coeluted with SP85 from this extract.

Among the purified proteins examined, cellulose-binding seemed to be specific for SP85 and SP35, since it was not observed for other coat proteins. Binding seemed to be intrinsic to SP85 and SP35, since no other proteins were present at similar levels in the FPLC fractions. In addition, binding occurred whether they initially flowed through (pre-germination) or bound to (post-germination) the DEAE column, suggesting that other molecules, which would be unlikely to copurify in both circumstances, are unlikely to be required.

To verify that SP85, SP65, and SP35 from pre-germination coats were equivalent in charge to their previously described counterparts from post-germination coats (*4*), they were examined by 2-D gel electrophoresis. The DEAE flow-through fraction was concentrated, desalted, and applied to an Avicel column. The column was eluted sequentially with water, 1 M NH_4Ac, 8 M urea, and 0.1% SDS, each containing 2-ME. SDS−PAGE analysis showed that the majority of SP85, SP65, and SP35 eluted in the urea fraction, with the remainder eluting in the flow-through and SDS fractions.[3] Analysis of the urea-eluted fraction by 2-D O'Farrell gel electrophoresis under denaturing conditions revealed that SP85, SP35, and SP65 each migrated to the neutral end of the gel,[3] similar to their behavior in extracts of post-germination coats (*4*). The presence of SP65 confirmed the interaction with SP85 seen in the binding assays on the SP-column fractions (Figure 2, lanes g,h). Thus, differential binding to DEAE by SP85 and SP35 from pre- and post-germination coats could not be attributed to changes in intrinsic charge.

SP85 Is Encoded by the pspB Gene. Gel bands from fractions enriched in SP85 were treated with CNBr to generate internal peptides, as sequence could not be obtained by Edman degradation of the intact protein. SDS−PAGE-purified CNBr-peptides yielded four sequences (Table 2) belonging to the previously described prespore cell-specific *psp*B (14E6) gene (*25*), and four from the *psp*D (PL3 or SP87) gene (*22*). Since the SP87 protein could be detected in the SP85-enriched fraction and did not bind Avicel (Figure 2, lanes g,h), these data suggested that SP85 is encoded by *psp*B.

The previously published sequence of *psp*B (*25*) was derived from the genomic library plasmid pG219, amplified in *E. coli*. Our sequencing of this plasmid added additional nucleotides at positions 1723 and 1769 (see Figure 4 for sequence), changed the assignment of nucleotides at positions 802 and 1734, and extended the sequence in the 3′-direction starting at position a in the schematic diagram in Figure 3B. An identical sequence was found in the cDNA library plasmid p14E6Pst73#1 represented in Figure 3C, except for differences in two areas: (a) the central region encoding the TXPP tetrapeptide repeats, and (b) the 3′-end of the coding region (starting at arrow b in Figure 3B) where there is a novel substitution starting at position 1877 (Figure 4). Consistent with the otherwise identical sequences of the two DNAs, Southern blot analysis (*25*), physical mapping (*38*), and genetic analysis (*39*) suggested that there is only a single *psp*B locus. To examine the origin of the different 3′-sequences, oligonucleotide primers corresponding to each 3′-sequence (P7 for p14E6Pst73#1; P8 for pG219) were combined with the primer P1 matching a shared upstream sequence (see Figure 3A for primer positions) and used to amplify native gDNA from strain A×3 in a PCR reaction. Each primer pair amplified DNA of the expected length from

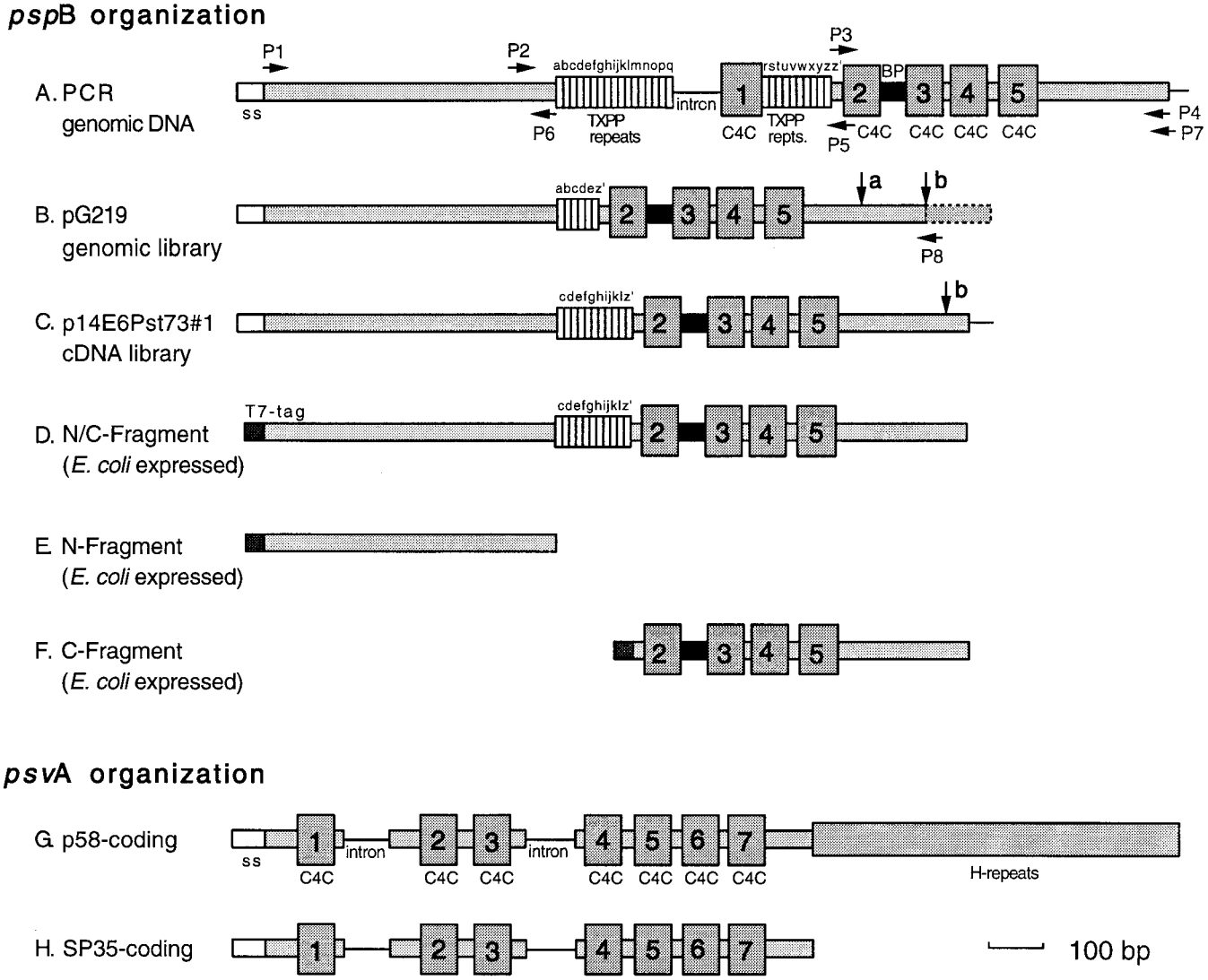

FIGURE 3: Organization of the *psp*B and *psv*A genes, their derived constructs, and inferred protein products. (A) This diagram is a composite of sequence derived by PCR amplification of *psp*B from strain Ax3 gDNA between primers P1 and P7, and cDNA sequence from part C below. (B) Schematic diagram of our our sequence determination of *psp*B in pG219, derived from a genomic library (*25*). (C) Schematic diagram of our sequence determination of *psp*B in p14E6Pst73#1, derived from a cDNA library (*25*). (D) Construct of combined N- and C-terminal *psp*B domains expressed in *E. coli*. (E) Construct of N-terminal *psp*B domain expressed in *E. coli*. (F) Construct of C-terminal *psp*B domain expressed in *E. coli*. (G) Comparative diagram of the p58 protein product of *psv*A (from refs *26* and *41*). (H) Diagram of the SP35 protein product of *psv*A (see text). Tall blocks, C4C-type Cys-repeats, which are numbered in order; mid-height blocks, tandem TXPP-repeats, which are lettered in order; short blocks, other domains; lines, untranslated sequences; ss, signal sequence; solid fill, position of basic proline (BP) motif sequence; a, denotes position of previously deduced stop codon (*25*); b, denotes position of sequence divergence between pG219 and p14E6Pst73#1; →, oligonucleotide primers, which are underlined in Figure 4, except for P8 in panel B, which has the sequence 5′-GTTAAGAGCATCTTCCAAAACCCTA.

its homologous but not its nonhomologous plasmid.[3] The primer pair P1/P7 amplified a product from gDNA, but P1/P8 did not. However, this product was about 500 bp longer than expected,[3] which is examined below. The divergent 3′-sequence in pG219 is probably noncontiguous *Dictyostelium* DNA. It is unlikely to represent an alternative exon, as (a) no intron was seen at this position in the genomic library DNA, (b) introns are short in *Dictyostelium* so the PCR reaction would be expected to bridge the hypothetical intervening sequence(s) yielding a slightly longer product, (c) if a hypothetical intron were located at the point of divergence between the two sequences, primer P4 would have been unable to hybridize, as it bridges this position, and (d) the new Figure 4 sequence, but not the pG219 sequence, has been found in the Tsukuba cDNA sequencing project.

As described above, PCR amplification of gDNA using P1 and P7 yielded a product of 2.1 kb which was 0.5 kb longer than expected. A similar discrepancy was also observed using P1 and P5 (shown below in Figure 5A), but not P1/P6 and P3/P4, which localized the difference to the coding region between P2 and P3 encoding the TXPP tetrapeptide repeats, where the two library sequences also diverged (Figures 3A−C). Direct sequencing of the PCR-amplified gDNA between P1 and P4 revealed sequence identity with p14E6Pst73#1 except for 550 nt of additional, novel sequence starting at nucleotide 613 (Figure 4), within the tetrapeptide repeat region. This sequence included 458 nt in-frame with the previously identified ORF, and 92 nt of intron-like sequence (Figures 3A and 4). The putative intron consisted primarily of homopolymeric stretches of A's and T's, began with a GA sequence at the second position

```
                                                                P1→
222  ATGAGACTATTAAGTGTTCTTTTAATAGGATTCCTCTGTTTAGCAGGTACCTATGCACAA  281
1    M  R  L  L  S  V  L  L  I  G  F  L  C  L  A  G  T  Y  A  Q    20

282  AAATACCAATTGTCACCGGCTTATAATGATCCATATTTAACCGATGATAAAACAGGAACA  341
     K  Y  Q  L  S  P  A  Y  N  D  P  Y  L  T  D  D  K  T  G  T    40

342  CATGATTTCTGGGTTCAAAATGCGTCTTTACCAGTTTTCTATGGTTTCCATGATTGGAAT  401
41   H  D  F  W  V  Q  N  A  S  L  P  V  F  Y  G  F  H  D  W  N    60

402  TTCCAAGATAATTCTGGTATAATGGAAATTAATGGTGATGAAATGCATATCACAGGTAAA  461
41   F  Q  D  N  S  G  I  M  E  I  N  G  D  E  M  H  I  T  G  K    80

462  ATTTATCCAACTGTTAATATGGGTGATTGCCACAGATACAATGTTGATTTAGTATTCAAG  521
81   I  Y  P  T  V  N  M  G  D  C  H  R  Y  N  V  D  L  V  F  K   100

522  AAAGATAAATCTGGTAATGTTATGCCAAAGAAAGAGCTAAGAGAAAGTGCATACGTTCCA  581
101  K  D  K  S  G  N  V  M  P  K  K  E  L  R  E  S  A  Y  V  P   120

582  CATGGACCAATTGACCCAGCCACTTGGAAATATTATACTTTTGTTCAAGGTAAATGGACT  641
121  H  G  P  I  D  P  A  T  W  K  Y  Y  T  F  V  Q  G  K  W  T   140

642  GGTTTTGGTTGTGACCCTCAAAATGTCGTCTTTTCCGGTGCTGAAGGTGGTATGCCATTA  701
141  G  F  G  C  D  P  Q  N  V  V  F  S  G  A  E  G  G  M  P  L   160
                                                    P2→
702  CAACTCGGTTATGGTGCAAATGGTAAAAATGGTGATAATGGTATCTCTGTTTGGTTAATT  761
161  Q  L  G  Y  G  A  N  G  K  N  G  D  N  G  I  S  V  W  L  I   180
                                                ←P6
762  TATGGTTACACTATTGTTGATATCAATTGTAACATTCGTCCAATCATTACACAATCACCA  821
181  Y  G  Y  T  I  V  D  I  N  C  N  I  R  P  I  I  T  Q  S  P   200

822  ACTCAACCACCAACACAACCACCAACTTATCCACCAACACAACCACCAACACAACCACCA  881
201  T  Q  P  P  T  Q  P  P  T  Y  P  P  T  Q  P  P  T  Q  P  P   220

882  ACACAACCACCAACTTATCCACCAACTTATCCACCAACTTATCCACCAACTTATCCACCA  941
221  T  Q  P  P  T  Y  P  P  T  Y  P  P  T  Y  P  P  T  Y  P  P   240

942  ACTTATCCACCAACTCATCCACCAAGTTATCCACCAACTTATCCACCAACACAACCACCA 1001
241  T  Y  P  P  T  H  P  P  S  Y  P  P  T  Y  P  P  T  Q  P  P   260

1002 ACTCAACCACCAGTACAAGgtaaagtattattttaaattatataaaatccataaaattaa 1061
261  T  Q  P  P  V  Q  D                                          267

1062 aaactcatttgttaatttctttttttttttttttttttaaaaaaaataataagATTGTTCAAC 1121
268                                                     C  S  T    270

1122 TTTAGAATGCCCAGAAGGTTTCCATTGTGAAATTGTAAATAATCGTAGAACATGTGTTTG 1181
271   L  E  C  P  E  G  F  H  C  E  I  V  N  N  R  R  T  C  V  C   290

1182 TGATACTACAGTTCCACCAACTCATCCACCAACTCAATCACCAACTTATCCACCAACTCA 1241
291   D  T  T  V  P  P  T  H  P  P  T  Q  S  P  T  Y  P  P  T  Q   310

1242 ACCACCAACTCAACCACCAACTTATCCACCAACTTATCCACCAACTTATCCACCAACACA 1301
311   P  P  T  Q  P  P  T  Y  P  P  T  Y  P  P  T  Y  P  P  T  Q   330
          P3→                    ←P5
1302 ACCACCAAGAGCATCATGCGATAATGTTAGATGTCCAAGAGGTTACCATTGTGAATGTAA 1361
331   P  P  R  A  S  C  D  N  V  R  C  P  R  G  Y  H  C  E  C  N   350

1362 TCATTGGGAAAATGTTGCTAGATGTGTAAGAAACGAAGAACCAACTCATAGACCAAAACC 1421
351   H  W  E  N  V  A  R  C  V  R  N  E  E  P  T  H  R  P  K  P   370

1422 ACCACACCATGGTTGTCGTCCAGATAGTTGCGCTAGAGGTGAGAAATGTATCTGTGTACG 1481
371   P  H  H  G  C  R  P  D  S  C  A  R  G  E  K  C  I  C  V  R   390

1482 TGGTAAGATCTATTGTATTCCAAAGACCACCTGTGATCAAGTACGTTGTCCAAGAAAACA 1541
391   G  K  I  Y  C  I  P  K  T  T  C  D  Q  V  R  C  P  R  K  H   410

1542 CCATTGTGAATGCAATAGAAAAGGTCAAGTTTTCTGTGTTCCAGATTGTCCAAAATTAAC 1601
411   H  C  E  C  N  R  K  G  Q  V  F  C  V  P  D  C  P  K  L  T   430

1602 TTGTAAACAAGTTGGTTGCCCAGAAAACCATGAATGCGTTTCAAGACGTGGTGAACTCCA 1661
431   C  K  Q  V  G  C  P  E  N  H  E  C  V  S  R  R  G  E  L  H   450

1662 TTGTGTATACGTTAGACCACCAACAGGTAGATGGGGTGATGATCTCTCAGATTTAGGTTA 1721
451   C  V  Y  V  R  P  P  T  G  R  W  G  D  D  L  S  D  L  G  Y   470

1722 CAATCAAGCAAGTGTTGAAGCTGCAATTGATGCTTATGAAGCCCAAAGGGGTCAAAGACC 1781
471   N  Q  A  S  V  E  A  A  I  D  A  Y  E  A  Q  R  G  Q  R  P   490

1782 ACATAATAATATTCCATCAAACAATTTACATCAAGAAAACAATGATAATTTGGGTGATGG 1841
491   H  N  N  I  P  S  N  N  L  H  Q  E  N  N  D  N  L  G  D  G   510
                                                              ↓
1842 TTTTTATGATGGTGTTGAAATTGGTTTCGCTGACGGAGGTTTCGACGTAAACGATCTCAA 1901
511   F  Y  D  G  V  E  I  G  F  A  D  G  G  F  D  V  N  D  L  N   530
        ←P4    ←P7
1902 TGGTTTTTAAAAAAAAAAAAAAAAAAATAATTATTAAATTTTTA                  1944
531   G  F  Stop                                                    532
```

FIGURE 4: Revised sequence and deduced translation of the *psp*B coding region from strain Ax3. The sequence corresponds to the PCR-derived map shown in Figure 3A, and has Genbank accession number AF066071.[6] Partial *psp*B sequences (corresponding to nucleotides 777−1308) from strains WS576 and WS380B (see text) have accession numbers AF066072 and AF066073, respectively. The sequence between primers 1 and 7 is derived from PCR amplification of cellular DNA (see text). The sequence upstream of primer 1 and downstream of primer 4, and the primer sequences themselves, are derived from p14E6Pst73#1. Nucleotide numbering at the 5′-end is as in ref *25*. The sequences of oligonucleotide primers used are underlined. Amino acid sequences shown in Table 2 are in bold. The vertical arrow marks the position of sequence divergence between p14E6Pst73#1 and pG219 (see text).

of the codon triplet, and ended with an AT, thus conforming to the typical *Dictyostelium* intron pattern. As observed for other coat protein genes (*4*), this intron is positioned at the junction between distinct sequence motifs suggestive of

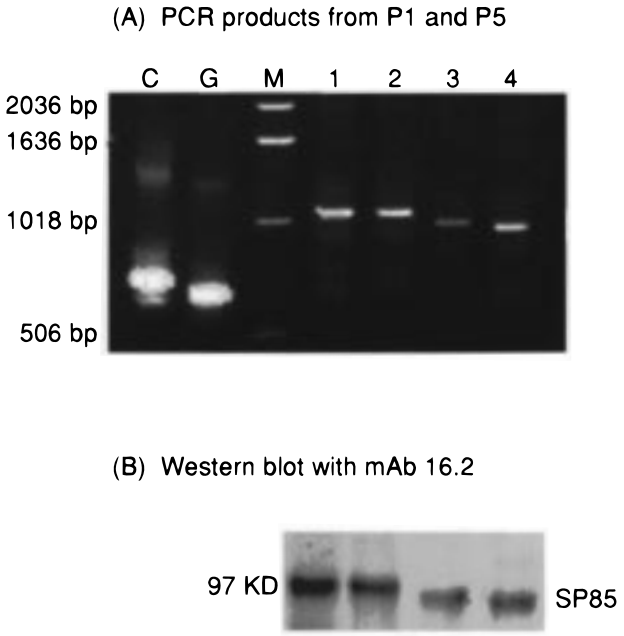

FIGURE 5: Strain-specific length polymorphisms in *psp*B and SP85. (A) DNA from vegetative (growing) cells was amplified using primers P1 and P5, and compared on a 1.2% agarose gel by ethidium bromide staining. Templates: C, *psp*B cDNA plasmid p14E6Pst73#1; G, *psp*B genomic DNA plasmid pG219; lane 1, strain Ax3 (derived from NC-4); lane 2, strain NC-4; lane 3, strain WS380B; lane 4, strain WS576. Lane M contains an M_r ladder, whose lengths are given at the left. (B) 10^6 spores from each strain were extracted and resolved by SDS−PAGE under reducing conditions, Western blotted, and immunoprobed with mAb 16.2. The apparent M_r of SP85 from strain Ax3 is given on the left.

separate folding domains. Conceptual translation of the additional sequence showed that it encoded a C4C-type Cysteine-repeat motif nested within 21 TXPP tetrapeptide repeats, motifs which are also present elsewhere in the protein. Inclusion of the new sequence does not cause the total known sequence to exceed the length of the mRNA as determined by Northern blotting (*25*). It is surmised that amplification of *psp*B library DNA in *E. coli* resulted in loss of internal sequence resulting from recombination of the highly repetitive DNA encoding the TXPP tetrapeptide repeats. The inferred recombination events deleted nts 870−1296 from *psp*B in pG219, and nts 810−843 and 954−1296 from *psp*B in p14E6Pst73#1 (Figures 3A−C and 4), thus explaining the length discrepancies between the two plasmids in this region.

The revised sequence suggests that SP85 consists of three domains separated by spacer regions consisting of 17 and 10 TXPP repeats, respectively (Figure 3A). The middle domain is comprised of a single EGF-like C4C-repeat. Four additional C4C-repeats occur in the C-terminal domain.

SP85 displays M_r polymorphisms in wild-type strains WS380B and WS576 relative to NC-4 and its axenic derivative Ax3 (Figure 5B; ref *39*). PCR analysis of gDNA from these strains using P1 and P5 revealed a corresponding variation in the lengths of *psp*B DNA (Figure 5A), verifying that SP85 is encoded by *psp*B. Sequence analysis of the PCR products revealed that *psp*B from all three strains diverged in the region of TXPP tetrapeptide repeats g−q (Figure 3a), such that 2 and 3 fewer repeats were encoded in WS380B and WS576, respectively. These were not simple deletions from the Ax3 sequence, but may have been

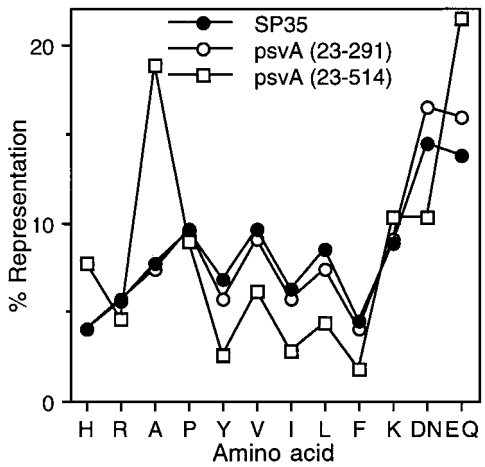

FIGURE 6: Amino acid analysis of SP35. The Avicel-bound subfractions of salt gradient-eluted SP-fractions containing SP35 were pooled and resolved by SDS−PAGE (see Figure 2n), and electroblotted onto a PVDF membrane. The Coomassie Blue-stained SP35 band was cut from the PVDF electroblot and acid hydrolyzed. PTC-derivatives were prepared and analyzed by HPLC to quantitate the amino acids released. Filled circles: values which could be reliably quantitated given the amount of material available are reported. Open circles: theoretical values for codons 23−291 of the *psv*A gene. Open squares: theoretical values for codons 23−514 of *psv*A.

deletions from a hypothetical parent strain with at least four more repeats than found in Ax3. Thus, this region of *psp*B, which was recombinogenic in *E. coli*, also displays heterogeneity in different wild-type strains of *Dictyostelium*, as predicted previously (*39*), and also described for another *Dictyostelium* gene encoding repetitive mucin-like repeats (*40*). Spontaneous recombination is infrequent, however, as no differences were seen between strain NC-4 and its derivative Ax3.

SP35 Is Encoded by the psvA Gene. CNBr-peptides from pre-germination SP35 gel bands and SP35 from post-germination coats yielded six amino acid sequences (Table 2). These sequences were each found in the 5′-half of the previously described prespore cell-specific *psv*A (EB4) gene, and defined SP35's N-terminus after signal peptide cleavage. Since *psv*A appears to be a single-copy gene based on Southern blot analysis (*41*) and physical mapping (*38*), SP35 is likely encoded by *psv*A. However, the apparent M_r of SP35 was smaller than both the predicted value for the *psv*A gene product p58, which is 53 015, and its previously reported value, 58 000 (*26*). To analyze the difference, SP35 was probed with antisera specific for the NH_2-terminal and COOH-terminal regions of p58. Only the C-antibody, generated against a peptide corresponding to the Cys-rich NH_2-terminal part of p58, recognized SP35 (Figure 2, lanes o,p). This suggested that SP35 is encoded by the 5′-part of the *psv*A coding region. This was supported by amino acid composition analysis, which can be a sensitive measure of protein identity (*42*) and showed a close match to the predicted composition of codons 23−291 of *psv*A (Figure 6). The predicted M_r, 28 834, is consistent with the measured apparent M_r of 35 000. Since p58 is absent from spores (*26*), and sequences from the 3′-coding region of p58 but not SP35 have been found in the Tsukuba cDNA cloning project, SP35 is probably proteolytically cleaved in the vicinity of codon 291 to yield SP35 during sporulation.

The p58 and SP35 products of the *psv*A gene are represented in Figure 3G,H. SP35, which corresponds to the N-terminal half of p58, consists primarily, as previously noted (*41*), of seven C4C EGF-like repeats.

A Recombinant SP85 Domain Isolated from E. coli Expressing Partial pspB cDNA Bound Cellulose. For further evidence that SP85 has intrinsic cellulose-binding activity, the construct shown in Figure 3D, which corresponds to the *psp*B derivative represented in the cDNA library plasmid p14E6Pst73#1, was expressed in *E. coli*. This construct encoded the N-terminal domain upstream of and including 10 of the tandem TXPP repeats, and the C-terminal domain downstream of but including the final TXPP repeat. Constructs encoding the N- and C-terminal domains separately (Figure 3E,F) were also expressed. The terminal domain constructs were devoid of TXPP repeats as they could not be expressed in *E. coli*.[3] These constructs lacked the *psp*B signal peptide, and encoded a T7-epitope tag at the N-termini.

E. coli clones expressing the N/C *psp*B cDNA were lysed and separated into soluble and insoluble (inclusion body) fractions. The inclusion body fraction contained a predominant protein by SDS−PAGE/Coomassie Blue analysis with an apparent M_r of 50 000 (Figure 7, lanes g−n). This was consistent with the expected M_r of 49 309, and the protein band reacted with the anti-T7 antibody. The soluble *E. coli* fraction contained lesser amounts of the M_r 50 000 protein and a smaller M_r 25 000 protein, whose level increased at the expense of the M_r 50 000 band in older extracts and during incubation in the absence of cellulose (lanes c−f). Since this smaller fragment carried the N-terminal epitope tag, it is likely to be a proteolytic derivative containing the N-terminal domain only. *E. coli* clones transformed with the N-terminal fragment construct contained an anti-T7 reactive protein with an apparent M_r of 24 000 in both the soluble fraction and in inclusion bodies (lanes s−z), in close agreement with the predicted M_r of 24 151. Clones expressing the C-terminal fragment contained an anti-T7 reactive protein of apparent M_r 21 000 in inclusion bodies (lanes o−r), close to the expected M_r of 21 688.

Both the soluble and urea-solubilized inclusion body fractions from the N/C-expressing strain were tested for binding to Avicel and Sephadex (Figure 7), as for coat proteins in Figure 2. The N/C-fragment from the soluble fraction bound Avicel strongly (lanes c,d), as determined by Western blotting with the anti-T7 epitope Ab. Binding was specific for cellulose, as the fragment did not bind Sephadex (lanes e,f), and was specific for the N/C-fragment, as other Coomassie Blue-stained bands, as well as the accumulated truncated N-terminal SP85 fragment (N′), failed to bind Avicel (lanes a,b). The N/C-fragment recovered from the insoluble fraction also bound Avicel efficiently and selectively, but only if it had been treated with glutathione (lanes g−n). This material showed partial binding to Sephadex (lanes m,n), as if it had not fully renatured. Approximately 2.5 μmol of the renatured N/C-fragment bound per gram of Avicel at saturation, which is comparable to the binding of known cellulose-binding proteins (*43*). These results confirmed that SP85 binds Avicel. The cellulose-binding activity of SP85 appeared to depend on protein folding, and did not require the middle domain containing the C4C-1-repeat embedded within the TXPP repeat cluster, or glycosylation.

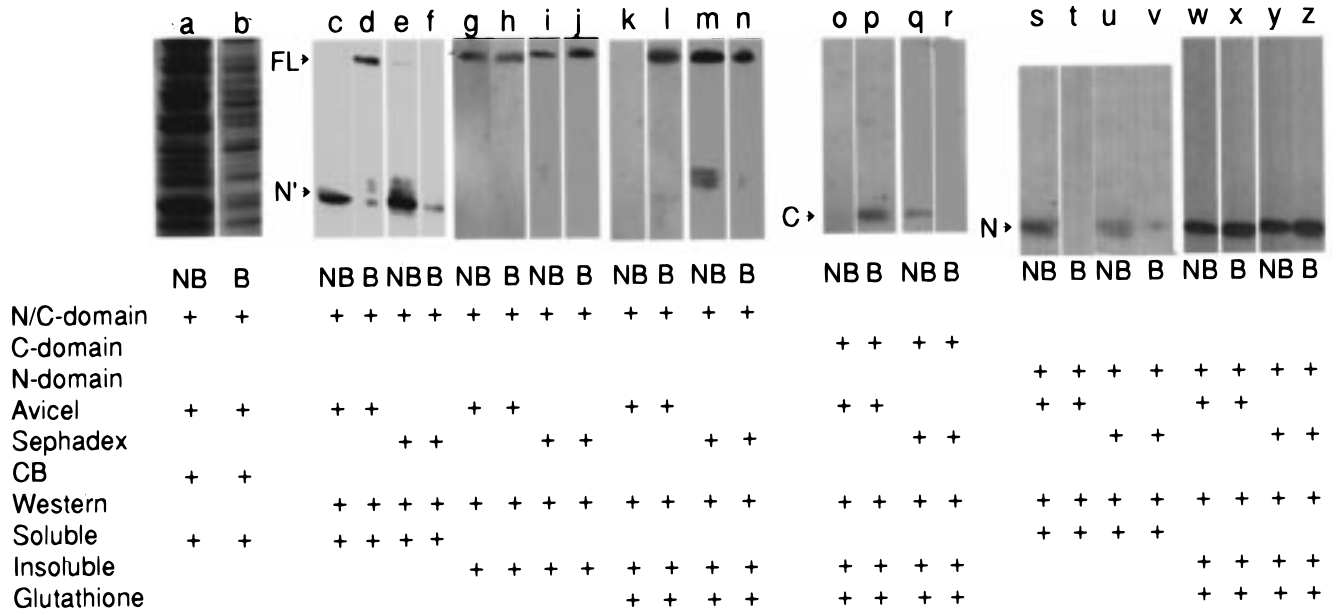

FIGURE 7: Binding of recombinant SP85 polypeptides to Avicel and Sephadex. Binding of N/C-construct (Figure 3D), and the separate N-(Figure 3E) and C-terminal (Figure 3F) domains, to Avicel or Sephadex was assayed as in Figure 1, with detection either by Coomassie Blue (CB) for total protein or anti-T7 antibody for the epitope tag. Extracts consisted either of the soluble, cytoplasmic fraction, or the insoluble fraction after solubilization in urea and renaturation in the presence or absence of glutathione as indicated. Lanes a−n: Binding of extracts expressing the N/C construct are shown; in some cases, especially when Avicel was absent, a truncated N-terminal (N′) fragment was also detected. Lanes o−r: Binding of extracts expressing the C-terminal fragment. Lanes s−z: Binding of extracts expressing the N-terminal fragment.

Table 3: Cellulase Activity during Germination and in Extracts of Post-Germination Coats

preparation	amount	drain time (s)[a]
H_2O		116
interspore matrix	13 μg protein	116
germination supernatant	6 μg protein	102
dialyzed urea extract of coat	6 μg protein	79
	30 μg protein	45
	60 μg protein	49
DEAE fractions from dialyzed urea extract[b]		
−2	100 μL	116
−1	100 μL	101
0	100 μL	60
+1	100 μL	61
+2	100 μL	78
DEAE fractions from non-dialyzed urea extract[b]		
−1	100 μL	118
0	100 μL	117
+1	100 μL	115
+2	100 μL	115
+3	100 μL	115
+4	100 μL	114
Trichoderma reesei cellulase	10 μg	59
	20 μg	46

[a] Cellulase activity was inferred from the reduction in viscosity of a solution of carboxymethylcellulose, which was measured as the drain time from a calibrated column; shorter drain times correlate with decreased viscosity and higher cellulase activity. [b] The fractions are numbered to facilitate comparison between different DEAE−FPLC trials. '0' corresponds to the fraction most enriched in SP85. Neighboring fractions are numbered relative to this position.

The expressed C-terminal fragment, found only in the insoluble fraction, also bound Avicel (lanes o−r), but only when it was treated with glutathione. The expressed soluble N-terminal fragment failed to bind Avicel (lanes s−v), consistent with observations on the N′-fragment above. The N-fragment derived from the insoluble fraction showed only nonspecific binding to Sephadex and Avicel (lanes w−z) even when treated with glutathione. Thus the Avicel binding of the N/C-protein could be ascribed to its C-terminal domain, and the nonspecific binding of some fractions to Sephadex could be ascribed to its N-terminal domain, which may not have fully renatured.

SP85 Copurified with a Cellulase after Germination. Since many known cellulose-binding proteins degrade cellulose (*43*−*46*) or bind to cellulases (*47*, *48*), we asked if SP85 or SP35 possessed or associated with cellulase activity in any of the extracts. Cellulase activity was measured by a viscometric assay (Table 3), which measures the digestion of carboxymethylcellulose in solution, and a zymogram assay (Figure 8), which measures the digestion of carboxymeth-

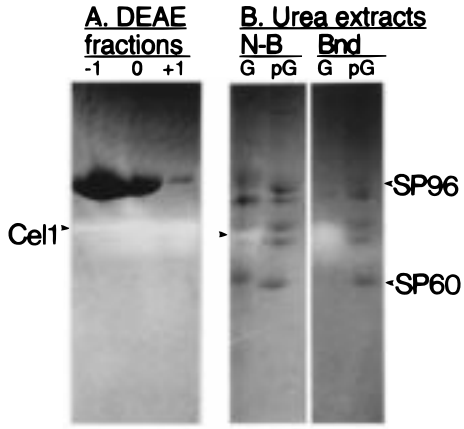

FIGURE 8: Zymogram analysis of cellulase activity in FPLC fractions and coat extracts. (A) DEAE−FPLC fractions of a dialyzed urea extract from post-germination coats were subjected to SDS−PAGE and incubated in 1% carboxymethylcellulose for 2 h. After staining with Congo Red and destaining, cellulase activity and protein were visualized as yellow (light in the reproduction) and red (dark in the reproduction) bands, respectively. Fractions 0 and +1 are most enriched in SP85, and −1 is most enriched in SP96. The fractions correspond to the similarly numbered fractions analyzed in Table 3. (B) Dialyzed urea extracts from post-germination spores (G) and gel-filtered urea extracts from pre-germination spores (pG) were separated into Avicel-bound (Bnd) and nonbound (NB) fractions, and analyzed as in panel A. The positions of the primary cellulase detected, Cel1, and the coat proteins SP96 and SP60 are indicated.

ylcellulose overlayed on an SDS−PAGE gel as revealed by negative staining with Congo Red, while proteins in the gel are positively stained.

Of the various spore-associated fractions, cellulase activity was detected only after germination, in the dialysis-renatured urea extracts of coats and the medium in which germination occurred (Table 3). As found previously (*49*, *50*), prior to germination, coat extracts (Figure 8B) and the interspore matrix from between spores (Table 3) exhibited little or no activity. Nearly all of the activity in the coat extract migrated with an apparent M_r of 75 000 and could bind to Avicel (Figure 8B), and is referred to as Cel1. Cel1, which is distinct from SP85 and SP35, is similar to a previously described cellulase, CelA, which is synthesized and secreted during germination (*51*).

Cel1 was detected after DEAE chromatography of the urea extract of post-germination coats when the extract was pre-dialyzed, coeluting with SP85 (Table 3). Since no activity was found in the absence of pre-dialysis (Table 3), coelution did not seem to reflect an intrinsic property of the cellulase activity. Zymogram analysis showed that the cellulase corresponded to Cel1 (Figure 8A). Thus SP85 and SP35 do not possess intrinsic cellulase activity, but SP85 may reconstitute a complex with Cel1 during dialysis which is stable to ion-exchange chromatography.

DISCUSSION

Reconstitution of an Insoluble Coat Protein Complex Required Cellulose. Spore coats could be isolated from intact spores after lysis in a French pressure cell, using the same density gradient protocol as for coats released naturally by germination. When total protein was extracted using urea and reducing agent, the cellulose and galuran were retained in the residual shell (Table 1). After removal of urea and reducing agent by gel filtration or dialysis, the protein largely remained soluble after centrifugation at 100000$g \times 30$ min.[4] This was consistent with earlier data showing that coat proteins secreted during suspension culture, in the absence of spore differentiation, are soluble (*18*). However, when incubated with insoluble fragments of purified Avicel cellulose under renaturing conditions, proteins associated to form an insoluble complex (Figure 1a,b). Incubation of coat proteins renatured by gel filtration with a 100-fold mass excess of cellulose captured half of the protein, showing that renaturation was efficient. The galuran, which was not extracted by urea (Table 1), was thus not required for cellulose-binding. All major protein species were bound (Figure 1a), suggesting that reconstitution was similar to the original process of coat formation. Proteins did not bind to Sephadex or Sepharose (Figure 1c−f), thus demonstrating a specificity consistent with the polysaccharide composition of the coat.

During the course of these experiments it was observed that proteins renatured by dialysis instead of gel filtration bound cellulose poorly although more efficiently than to Sephadex or Sepharose.[5] SP35 was greatly depleted after dialysis and could be recovered from the cellulosic dialysis membrane by washing with SDS and reducing agent,[5] and this eluate was greatly enriched in SP35 and SP85.[5] Indeed, these two proteins were also the most enriched in the cellulose-binding fraction of coat extracts renatured by gel filtration (Figure 1a,b). This suggested that binding of most coat proteins to cellulose might be indirect via SP35 and SP85, but this model remains to be tested. In the present study, it was necessary to avoid dialysis and ultrafiltration through cellulosic membranes, except in the presence of 4−8 M urea, to ensure protein recovery.

In another study, spore coat precursor proteins isolated from prespore cells exhibited poor binding to cellulose (Figure 1 of ref *15*). A possible explanation for the divergent results is that cellulose-binding may be developmentally regulated, as, e.g., p58 was not yet replaced by SP35. Alternatively, low binding may have been due to inappropriate disulfide bond formation during prespore cell lysis (*4*), or the presence of nonionic detergents, which might interfere with hydrophobic interactions which can underlie cellulose-binding (*52*−*54*).

SP85 Is Encoded by pspB. The M_r 85 000 coat protein which binds cellulose is equivalent to the previously identified SP85 coat protein (*4*), also known as PsB, based on similar apparent M_r, M_r polymorphisms in different wild-type strain isolates, apparent pI, and reactivity with mAbs 16.1 and MUD102. SP85 is encoded by *psp*B, based on a perfect match of internal SP85 peptide sequences with translated *psp*B sequences, and correlation of SP85 M_r polymorphisms with sequence variations in *psp*B, in different strains. *psp*B was previously concluded not to encode a coat protein based on the absence of differences in coats from a putative *psp*B$^-$ strain (*4*, *25*). Based on the new information in this report showing that a region comprising 550 nts was missing from the plasmid previously used to target the *psp*B

[5] Zhang, Y., and West, C. M., unpublished data.

[6] *psp*B sequences from strains Ax3, WS576, and WS380B have GenBank accession numbers AF066071, AF066072, and AF066073, respectively.

locus, it is not surprising that the previous gene targeting strategy (*25*) was not successful.

The DNA encoding the TXPP tetrapeptide repeats appears to be recombinogenic in *E. coli* as both previously reported library copies of *psp*B were missing the majority of the TXPP tetrapeptide repeats, including the embedded middle domain C4C repeat and intron (for the genomic library copy) (Figure 3B,C). This region also appears to be recombinogenic in *Dictyostelium*, as it is shorter by 2 or 3 TXPP repeats in wild-type strains WS380B or WS576, respectively. This was previously predicted from a genetic analysis of the *psp*B locus (*39*), which is shown here to be equivalent to the molecularly defined *psp*B (14E6) locus (*25*), having coincidentally been given the same name. Similar polymorphisms have been observed for another *Dictyostelium* protein containing tandem PTVT-tetrapeptide repeats, SP29 or PsA (*40*), and for vertebrate mucin genes (*55*).

Based on the revised sequence of *psp*B, SP85 is proposed to contain three major domains separated by 17 and 11 tandem TXPP tetrapeptide repeats, respectively (Figure 3A). The TXPP repeat clusters may serve as interdomain spacers as suggested for similar sequences in other extracellular proteins (*40*, *44*, *46*). The N-terminal domain bears no obvious resemblence to other database sequences. The middle domain consists of a single C4C-type repeat with a characteristic spacing of Cys-residues and predicted β-turns, and resembles the N-terminal subdomain of the EGF-motif (*4*, *26*). C4C-repeats are prevalent in all other known coat proteins (*4*) and one is also found in a stalk matrix protein, *sta*B (*56*). The cellulose-binding C-terminal domain contains four additional C4C-repeats, two of which are separated by a short basic proline-rich sequence which frequently separates C4C-repeats in other coat proteins (*4*). The C-terminal third of the C-terminal domain sequence is unique. The *mod*B glycosylation modification may be applied to the TXPP-repeat region as similar repeats are modified in this way in another prespore cell protein, SP29 (*40*, *57*).

*SP35 Is Encoded by psv*A. SP35 was suggested to be encoded by the single-copy (*38*, *41*) *psv*A locus based on the identity of sequences from internal peptides of SP35 with the conceptual translation product of *psv*A. However, the open reading frame of *psv*A predicts a protein product of M_r 53 015 after signal peptide cleavage, and an M_r 58 000 protein, p58, has been detected in prespore cells by antibodies raised against polypeptides corresponding to the N-terminal and the C-terminal halves (*26*). SP35 corresponds to the 5′-half of the coding region from codons 23 to approximately 291, based on its apparent M_r of 35 000, its amino acid composition (Figure 6), reactivity with an antiserum raised against the N-terminal half (Figure 2o,p) but not one against the C-terminal half, and the positions of the amino acid sequences obtained. *psv*A was previously concluded to not encode a coat protein based on the observation that p58 could not be detected after sporulation using the antibody against the C-terminal half (*26*). The simplest explanation is that, as prespore cells differentiate into spores, p58 is processed into SP35 which becomes a major component of the spore coat. Based on the anti-H antibodies used, this would not have been detected in the previous study. A strain in which *psv*A is thought to be disrupted exhibited no obvious phenotype (*26*) but, since an M_r 31 000 *psv*A-encoded protein similar to SP35 was still produced, an SP35-like protein was probably still produced by this strain.

SP85 and SP35 Bound Cellulose in the Absence of Other Protein Interactions. SP85 and SP35 each exhibited efficient cellulose-binding activities after partial purification from pregermination coats (Figure 2a−h, m−p). The renatured activities exhibited sharp specificity with respect to polysaccharide type. Binding was inhibited by urea, but not by 5 mM EDTA. Since the FPLC fractions most enriched for SP35 or SP85 did not contain other Avicel-binding proteins with similar abundance on Coomassie Blue-stained SDS−PAGE gels, it is likely that these proteins bind directly to Avicel. Cellulose-binding activity was also robust after purification from post-germination coats (Figure 2i−l, q−r), despite their binding to, rather than flowing through, the DEAE column. This also suggested that cellulose-binding was not due to a co-purifying protein. A polysaccharide component is unlikely to mediate binding as none seemed to be extracted from the coat by urea (Table 1). Avicel-binding was not a property of all proteins, as SP96 (Figure 2i,j,q,r), SP65 (Figure 2e,f), SP87 (Figure 2g,h), and the entire dialyzed coat extract[5] did not bind. Nonbinding does not seem to result from failure to renature, because these proteins were able to bind after gel filtration of the original coat extract (Figure 1a,b).

Since SP85 was not as completely purified from other proteins as SP35, SP85 domains were expressed from partial *psp*B cDNAs transfected into *E. coli*. The N/C construct containing the N- and C-terminal domains bound Avicel efficiently and specifically (Figure 7c−n), verifying that other coat components were not required. Binding activity could be ascribed to the C-terminal domain (Figure 7o−r). The C-terminal domain contains 4 tandem C4C-repeats with distinct sequences. The Cys-spacing of the C4C repeat is found in the fungal type I cellulose-binding domain (*46*), and the Cys-spacing of the fungal type I motif is present across two consecutive C4C repeats of SP85. These Cys-rich motifs may also be important for binding of SP85 to cellulose, as Avicel-binding of the recombinant C-terminal domain required renaturation in the presence of glutathione.

As cited above, the cellulose-binding activity of precoat proteins from extracts of prespore cells appeared to be relatively weak, but was reportedly specific for cellulose relative to other polysaccharide particles (*15*). This binding activity of total coat protein was found to be reduced even further in extracts from mutant strains which were lacking the coat proteins SP96 or SP70, and binding of SP85 was not observed. The results were interpreted to mean that a complete precoat complex is required for cellulose-binding activity. In contrast, proteins are incorporated normally into mutant coats in vivo (*4*), and we have observed that proteins extracted from mutant coats (strain TL56) exhibit efficient Avicel-binding activity in vitro.[5] While the reason for this discrepancy is not clear, the low level of cellulose-binding exhibited by the proteins isolated from prespore cells may not be directly comparable to the high levels of binding when isolated from mature coats.

SP85 and SP35 Interact with Other Coat Proteins. Extensive interactions among all coat proteins except for SP75 have been previously described based on co-immunoprecipitation studies performed on extracts of prespore cells (*11*−*14*), which contain all of the coat protein precursors and the galuran compartmentalized together in the prespore

vesicle (*3*). These results provide an explanation for how binding of SP85 and/or SP35 to cellulose might result in the binding of all other coat proteins (Figure 1a,b), although these interactions have not been shown to occur in vivo (*4*) and remain to be demonstrated to exist at the time of sporulation.

The possibility that coat protein interactions are regulated is suggested by several observations. Binding of SP85 and SP35 to the DEAE column after but not before germination is apparently not due to a change in these proteins, as their intrinsic apparent isoelectric points and M_rs do not vary. Partial coelution from the DEAE column with SP96 and Cel1 (Figure 8, Table 3) suggests that interaction of SP85 with other proteins may be developmentally regulated. A separate type of regulation is suggested by the apparent processing of p58 into SP35 before the coat is assembled.

Cel1, which was germination-specific, is probably equivalent to the previously described CelA (*44*, *51*). An interaction with SP85 would be reminiscent of the cellulosome complex found in anaerobic bacteria, which consists of cellulases and a noncatalytic, multidomain cellulose-binding protein called scaffoldin, which contains multiple binding sites for other proteins and helps target the cellulase complex to cellulose (*43*, *46*−*48*, *58*). An analogous interaction of CelA with SP85 may confer on it the ability, which it lacks in isolation (*44*), to degrade crystalline cellulose in coats.

Possible Functions of SP85 and SP35. The ability of SP85 to bind cellulose may explain its overlapping distribution with the cellulose layer of the coat in vivo as determined by immunoelectron microscopy (*3*, *19*). Since other coat proteins failed to bind cellulose after purification and mostly appear to reside in the outer regions of the coat beyond cellulose, it is possible that they bind indirectly via SP85 and/or SP35, perhaps by virtue of the previously described coat protein complex. Our unpublished observation that coat proteins bind cellulose poorly after SP35 and some SP85 were depleted by dialysis[5] also supports this model. The intimate association of coat glycoproteins with cellulose may be important for maintaining a permeability barrier to protect the coat from exogenous proteins or enzymes which might digest it. Finally, SP85 may help target CelA to cellulose after it is secreted during germination, thereby facilitating breakdown of the wall prior to emergence of the amoeba.

ACKNOWLEDGMENT

We are grateful to K. Williams, B. Yoder, D. Blumberg, M. Hildebrandt, W. Nellen, V. McGuire, S. Alexander, J. Coffman-Powell, R. A. Firtel, and W. F. Loomis for sharing cell strains and antibodies, to Scott Whittaker and Greg Erdos of the ICBR EM Core Laboratory for performing the EM studies, to Katalin Auffenberg for performing the viscometric analyses, to M. Gritzali for performing the sugar analyses, and to the UF ICBR Protein Chemistry Laboratory for performing Edman degradation and amino acid composition analyses.

REFERENCES

1. Richardson, D. L., Loomis, W. F., and Kimmel, A. R. (1995) *Development 120*, 2891−2900.
2. Freeze, H. H. (1997) in *New Comprehensive Biochemistry, Vol. 29B* (Neuberger, A. and Van Deenen, L. L. M., Eds.) pp 91−105, Elsevier Science, Amsterdam, The Netherlands.
3. West, C. M., and Erdos, G. W. (1990) *Dev. Genet. 11*, 492−506.
4. West, C. M., Mao, J., van der Wel, H., Erdos, G. W., and Zhang, Y. (1996) *Microbiology 142*, 2227−2243.
5. Wilkinson, D. G., and Hames, B. D. (1983) *Eur. J. Biochem. 129*, 637−643.
6. Cotter, D. A. (1981) *The Fungal Spore*: *Morphogenetic Controls* (Turian, G., and Hohl, H. R., Eds.) pp 385−411, Academic Press, New York.
7. Hemmes, D. E., Kojima-Buddenhagen, E. S., and Hohl, H. R. (1972) *J. Ultrastruct. Res. 41*, 406−417.
8. Kessin, R. H., Gundersen, G. G., Zaydfudim, V., Grimson, M., and Blanton, R. L. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 4857−4861.
9. Loomis, W. F. (1975) *Dictyostelium discoideum*: *A Developmental System*, Academic Press, New York.
10. Lydan, M. A., and Cotter, D. A. (1994) *FEMS Microbiol. Lett. 115*, 137−142.
11. Devine, K., Bergmann, J., and Loomis, W. F. (1983) *Dev. Biol. 99*, 437−446.
12. Devine, K. M., Morrissey, J. H., and Loomis, W. F. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7361−7365.
13. Watson, N., McGuire, V., and Alexander, S. (1994) *J. Cell Sci. 107*, 2567−2579.
14. Watson, N., Williams, K. L., and Alexander, S. (1993) *J. Biol. Chem. 268*, 22634−22641.
15. McGuire, V., and Alexander, S. (1996) *J. Biol. Chem. 271*, 14596−14603.
16. West, C. M., and Erdos, G. W. (1992) *J. Cell Biol. 116*, 1291−1300.
17. Orlowski, M., and Loomis, W. F. (1979) *Dev. Biol. 71*, 297−307.
18. Erdos, G. W., and West, C. M. (1989) *Exp. Mycol. 13*, 169−182.
19. West, C. M., and Erdos, G. W. (1988) *Cell Differ. 23*, 1−16.
20. Fosnaugh, K. L., Fuller, D., and Loomis, W. F. (1995) *Dev. Biol. 166*, 823−825.
21. Nakao, H., Yamamoto, A., Takeuchi, I., and Tasaka, M. (1994) *J. Cell Sci. 107*, 397−403.
22. Yoder, B. K., Mao, J., Erdos, G. W., West, C. M., and Blumberg, D. D. (1994) *Dev. Biol. 163*, 49−65.
23. Bolwell, G. P. (1993) *Int. Rev. Cytol. 146*, 261−324.
24. Carpita, N. C., and Gibeaut D. M. (1993) *Plant J. 3*, 1−30.
25. Powell-Coffman, J. A., and Firtel, R. A. (1994) *Development 120*, 1601−1611.
26. Hildebrandt, M., Humbel, B. M., and Nellen, W. (1991) *Dev. Biol. 144*, 212−214.
27. Kozarov, E., van der Wel, H., Field, M., Gritzali, M., Brown, R. D., and West, C. M. (1995) *J. Biol. Chem. 270*, 3022−3030.
28. Fengel, D., and Wegener, G. (1979) *Hydrolysis of Cellulose*: *Mechanisms of Enzymatic and Acid Catalysis*, *Adv. in Chem. Series 181* (Brown, R. D., and Jurasek, L., Eds.) pp 145−158, American Chemical Society, Washington, DC.
29. Dunbar, B., and Wilson, S. B. (1994) *Anal. Biochem. 216*, 227−228.
30. Ploug, M., Jensen, A. L., and Barkholt, V. (1989) *Anal. Biochem. 181*, 33−39.
31. Ozols, J. (1990) *Methods Enzymol. 182*, 587−601.
32. Orsini, G., and Goldberg, M. E. (1978) *J. Biochem. 153*, 3453−3458.
33. Buchner, J., and Rudolf, R. (1991) *Bio/Technology 9*, 157−163.
34. Shoemaker, S. P., and Brown, R. D. (1979) *Biochim. Biophys. Acta 523*, 147−161.
35. Mathew, R., and Rao, K. K. (1992) *Anal. Biochem. 206*, 50−52.
36. Cotter, D. A., Miura-Santo, L. Y., and Hohl, H. R. (1969) *J. Bacteriol. 100*, 1020−1026.
37. Gonzalez-Yanes, B., Mandell, R. B., Girard, M., Henry, S., Aparicio, O., Gritzali, M., Brown, R. D., Erdos, G. W., and West, C. M. (1989) *Dev. Biol. 133*, 576−587.
38. Kuspa, A., and Loomis, W. F. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5562−5566.

39. Smith, E., Gooley, A. A., Hudson, G. C., and Williams, K. L. (1989) *Genetics 122*, 59−64.
40. Gooley, A. A., Marshchalek, R., and Williams, K. L. (1992) *Genetics 130*, 749−756.
41. Hildebrandt, M., Saur, U., and Nellen, W. (1991) *Dev. Genet. 12*, 163−169.
42. Shaw, G. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5138−5142.
43. Goldstein, M. A., Takagi, M., Hashida, S., Shoseyov, O., Doi, R. H., and SegalI, H. (1993) *J. Bacteriol. 175*, 5762−5768.
44. Ramalingam, R., Blume, J. E., and Ennis, H. L. (1992) *J. Bacteriol. 174*, 7834−7837.
45. Tomme, P., Driver, D. P., Amandoron, E. A., Miller, R. C., Jr., Antony, R., Warren, J., and Kilburn, D. G. (1995) *J. Bacteriol. 177*, 4356−4363.
46. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., and Warren, R. A. J. (1991) *Microbiol. Rev. 55*, 303−315.
47. Bayer, E. A., Morag, E., Shoham, Y., Tormo, J., and Lamed, R. (1996) *Bacterial Adhesion*: *Molecular and Ecological Diversity* (Fletcher, M., Ed.) pp 155−182, Wiley Liss, New York.
48. Morag, E., Lapidot, A., Govorko, D., Lamed, R., Wilchek, M., Bayer, E. A., and Shoham, Y. (1995) *Appl. Environ. Microbiol. 61*, 1980−1986.
49. Jones, T. H. D., Renobales, M. D., and Pon, N. (1979) *J. Bacteriol. 137*, 752−757.
50. Roseness, P. A. (1968) *J. Bacteriol. 96*, 639−645.
51. Blume, J. E., and Ennis H. L. (1991) *J. Biol. Chem. 266*, 15423−15427.
52. Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T. A., Knowles, J. K. C., and Teeri, T. T. (1992) *Proteins*: *Struct., Funct. Genet. 14*, 475−482.
53. Sakka, K., Takada, G., Karita, S., and Ohmiya, K. (1996) *Ann. N.Y. Acad. Sci. 782*, 241−251.
54. Gilbert, H. J., Hall, J., Hazlewood, G. P., and Ferreira, L. M. A. (1990) *Mol. Microbiol. 4*, 759−767.
55. Verma, M., and Davidson, E. A. (1994) *Glycoconjugate J. 11*, 172−179.
56. Robinson, V., and Williams, J. (1997) *Differentiation 61*, 223−228.
57. Zachara, N. E., Packer, N. H., Temple, M. D., Slade, M. B., Jardine, D. R., Karuso, P., Moss, C. J., Mabbutt, B. C., Curmi, P. M. G., Williams, K. L., and Gooley, A. A. (1996) *Eur. J. Biochem. 238*, 511−518.
58. Matano, Y., Park, J. S., Goldenstein, M. A., and Doi, R. H. (1994) *J. Bacteriol. 176*, 6952−6956.

BI9808013