Five Genes Involved in Biosynthesis of the Pyruvylated Galβ1,3-Epitope in Schizosaccharomyces pombe N-Linked Glycans®§

Received for publication, March 31, 2004, and in revised form, June 1, 2004
Published, JBC Papers in Press, June 1, 2004, DOI 10.1074/jbc.M403574200

Ekaterina N. Andreishcheva‡§¶†‡‡§§, Jeremy P. Kunkel‡§¶, Trent R. Gemmill¶, and Robert B. Trimble‡¶**
From the ¶Wadsworth Center, New York State Department of Health, Albany, New York 12201-0509 and the ¶Department of Biomedical Sciences, School of Public Health, State University of New York at Albany, Albany, New York 12201

The N-linked galactomannans of Schizosaccharomyces pombe have pyruvylated Galβ1,3-(PvGal) caps on a portion of the Galα1,2- residues in their outer chains (Gemmell, T. R., and Trimble, R. B. (1998) Glycobiology 8, 1087–1095). PvGal biosynthesis was investigated by ethyl methanesulfonate mutagenesis of S. pombe, followed by the isolation of cells devoid of negatively charged N-glycans by Q-Sepharose exclusion and failure to bind human serum amyloid P component, which acts as a lectin for terminal PvGal residues. Mutant glycans were characterized by lectin binding, saccharide composition, exoglycosidase sensitivity, and NMR spectroscopy. Restoration of the cell surface negative charge by complementation with an S. pombe genomic library led to the identification of five genes involved in PvGal biosynthesis, which we designated pvg1Δ-pvg5Δ. Pvg1p may be a pyruvyltransferase, since NMR of pvg1Δ mutant N-glycans revealed the absence of only the pyruvyl moiety. Pvg2p-Pvg5p are crucial for attachment of the Galβ1,3- residue that becomes pyruvylated. Pvg3p is predicted to be a member of the β1,3-galactosyltransferase family, and Pvg5p-green fluorescent protein labeling was consistent with Golgi localization. Predicted Pvg1p and Pvg3p functions imply that Galβ1,3- is added to the galactomannans and is then pyruvylated in situ, rather than by an en bloc addition of PvGalβ1,3- caps to the outer chain. PvGp4-green fluorescent protein targeted to the nucleus, and its sequence contains a MADS-box DNA-binding and dimerization domain; however, it does not appear to solely control transcription of the other identified genes. Pvg2p and/or Pvg5p may contribute to an enzyme complex. Whereas a functional role for the PvGal epitope in S. pombe remains unclear, it is nonessential for either cell growth or mating under laboratory conditions.

Naturally all of the proteins found on the cell surfaces of almost all organisms are glycosylated. In most cases, the glycan moiety associated with these glycoproteins provide the surface with a net negative charge, which plays a significant role in the regulation of the cell influx/exflux processes and in cell-cell interactions such as cell aggregation and pathogenic adhesion. Whereas a wide range of acidic saccharides and saccharide modifications (e.g. muramic acid, hexuronic acids, sialic acids, succinate, lactate, and pyruvate) are found in bacterial polysaccharides (1–6), in general, sialylated and (less frequently) sulfated glycans are associated with vertebrate glycoproteins. In several yeast species (i.e. Saccharomyces cerevisiae, Candida albicans, Pichia holstii, and Pichia pastoris), phosphate groups present on extracellular glycans provide the negative cell surface charge (7–9). A notable exception is sialic acid α2,3- and α2,6-linked to penultimate β-galactopyranosyl residues in glycans from two isolates of C. albicans (10, 11). In contrast to the charged groups summarized above, Schizosaccharomyces pombe N-linked polysaccharides contain pyruvic acid 4,6-ketal-linked to Galβ1,3-(PvGal)1 (12, 13). Pyruvate has only recently been found in eukaryotic polysaccharides and glycoprotein oligosaccharides. Available examples include several species of green and red algae (14, 15), the nervous system of the sea slug Aplysia kurodai (16), and the marine sponge Microciona prolifera, in which PvGal is required for cellular recognition and adhesion (17). It is interesting to note the close structural similarity between the 4,6-ketal-linked pyruvate of the PvGal epitope and α2,6-linked sialic acid common in glycoproteins of higher eukaryotes (12, 13, 18).

Given that bacterial pyruvyltransferases have been identified in Xanthomonas sp. (19, 20), Rhizobium sp. (21, 22), and Bacillus anthracis (23), among other species, it was of interest to determine whether S. pombe has a genome-encoded functional analog of these proteins. However, initial data base searches could not identify any sequence homology between known bacterial pyruvyltransferases and any putative S. pombe proteins. Therefore, S. pombe strains devoid of negatively charged N-glycans were isolated following mutagenesis, and genetic complementation with an S. pombe genomic library was utilized to identify genes involved in the PvGal biosynthetic pathway.
Here we demonstrate that at least five different genes are required to synthesize the PvGal epitope. Although the inter-relationships between the encoded proteins remain to be completely established through future biochemical analyses, the results obtained suggest that PvGal1,3-synthesis is carried out by a coordinated enzymatic system in which the Galβ1,3-residues are added to the *S. pombe* galactomannans and are then subsequently pyruvylated.

**EXPERIMENTAL PROCEDURES**

**General—**Mallickrodt Baker (Paris, KY) supplied 37% formaldehyde solution, 2-propanol, and anhydrous methanol. Pharmco (Brookfield, CT) provided anhydrous ethanol. Unless otherwise noted below, all other simple reagents were obtained from Sigma. Solvents, buffer components, and chemicals were ACS reagent grade or better. Sigma also supplied the antibiotics O418, kanamycin, and ampicillin. New England Biolabs (Beverly, MA) was the source for DNA restriction enzymes. Modifications. Briefly, the commercial SAP solution (containing 1 mg of SAP Staining of Fixed Cells—

**Cell Lines and Cell Growth—**Fission yeast “wild-type” *S. pombe* strains (h*’ ura4–294) (ATCC 38436) and (h*’ ura4–leu1–32 ade6–704) (ATCC 96081) and *S. cerevisiae* strain (MATa ura3–52 his3–Δ200 trp1–Δ901 lys2–801 sec2–37 leu2–3,112, SEY6210) (a kind gift from Dr. Scott Emr via Dr. Neta Dean) were used in this study. Cell growth media and growth conditions are described in the Supplemental Material.

**Mutagenesis and Isolation of Mutants—**“Wild-type” *S. pombe* (h*’ ura4–294) was mutagenized using ethyl methanesulfonate (Sigma). Following mutagenesis, cells were collected by centrifugation, washed with DDW, and resuspended in 400 µl of lectin blotting buffer (50 mM sodium citrate, 100 mM 2-mercapto- ethanol, 1% SDS, pH 5.5). Suspensions were heated in a 100°C water bath for 5 min and pelleted by centrifugation. Soluble glycoproteins were precipitated from the decanted supernatant by the addition of ice-cold acetone to 80% (29). Glycoproteins were resuspended in 200 µl of lectin blotting buffer/mg of protein, denatured by boiling for an additional 5 min, and loaded at 600 ng/lane on 8% SDS-PAGE gels. Sample preparation; SDS-PAGE; electroblotting to Immobilon-P transfer membranes (Millipore); and lectin blotting with concanavalin A (for α-linked Man), GSL/BSL I-B3 (for β-linked Gal), and PNA (for β-linked Gal) were performed according to standard protocols. Q-Sepharose Fast Flow was the source for PAGE supplies. Vector Laboratories supplied biotinylated concanavalin A lectin and biotinylated GSL/BSL I-B3. Biotinylated PNA lectin was obtained from Roche Applied Science.

**Compositional Analysis—**Acid hydrolysis of the N-linked polysaccharides was performed as described (12, 28). Monosaccharides were separated by HPAEC on a Dionex (Sunnyvale, CA) CarboPac PA1 4 x 250-mm analytical column isocratically with 16 mM NaOH and were monitored by PAD using a Dionex ED50 Electrochemical Detector.

**Depyruvylation and Eosinylglycosidase Digestions—**Wild-type galactomannans were depyruvylated with mild acid hydrolysis as described (12, 29). The same time window that showed that Galβ1,3--in the wild-type and the depyruvylated galactomannans represented about 10% of the total Gal (28). To confirm the presence of unsubstituted Galβ1,3-residues in the *pvg*1 mutant, we incubated ~3–µmol Man equivalents (~50 nmol of total oligosaccharide) of wild-type and mutant N-linked glycans overnight with 20 milliunits of β-galactosidase (Sigma) in 40 µl of digestion buffer (50 mM sodium citrate, 5 mM CaCl2, pH 6.0) and with 300 milliunits of *Xanthomonas manihotis* β-galactosidase (expressed in *Escherichia coli*) (New England Biolabs) in 40 µl of digestion buffer (50 mM sodium citrate, pH 4.5). Following precipitation with ice-cold 60% methanol (29), supernatants were evaporated under vacuum, and Gal release was measured by HPAEC-PAD. The supernatant was reconstituted with 0.1% CHAPS, precipitated by HPAEC-PAD, washed twice with C water, and the amount of total injected N-glycans so that release of only 10% of Gal by α-galactosidase (~100 nmol) yielded a prominent peak.

**Oligosaccharide Preparation and NMR Spectroscopy—** *S. pombe* cells were grown to early stationary phase in YEEDS. Following disruption by vortexing with 0.5-mm acid-washed glass beads, cell debris was removed by centrifugation. Nonglycosylated proteins were precipitated by 60% saturation with solid ammonium sulfate. After centrifugation, the glycoprotein-containing supernatant fraction was equilibrated by dialysis with deglycosylation buffer (50 mM sodium citrate, pH 5.5), and oligosaccharides were released from the glycoproteins with endo-β-N-acetylglucosaminidase H (30) (expressed in *E. coli*). Deglycosylated proteins were removed by centrifugation, and released N-glycans were sized on an extra fine mesh Bio-Gel P-4 (Bio-Rad) column, and the large galactomannans eluting at the void volume were examined further. Samples were prepared, and one- and two-dimensional NMR spectra were recorded, processed, assigned, and analyzed as described.

**Genetic Methods and Plasmid Construction—**An *S. pombe* genomic library (32) (a kind gift from Dr. Tony Carr) was used for complementation of mutation and restoration of wild-type phenotype. Library plasmids recovered from *pvg*1 mutants with restored wild-type phenotypes were sequenced using appropriate primers (Table SI, Supplemental Material) to determine their *S. pombe* genetic content. *S. pombe* pvg1 his3–112 and ura3–52 alleles were identified through a Washington University ELAST (33, 34) comparison of these sequences against the annotated *S. pombe* GeneDB gene data base (35, 36) (available on the World Wide Web at www. geneDB.org/geneDB/pombe/index.jsp). All plasmids in this study were derived from the low copy *S. pombe* shuttle vector pUR19 (ars1, ura4+).
FIG. 1. PvGal epitope expression assays. A, Q-Sepharose binding assay of S. pombe cells. Negatively charged pvg<sup>−</sup> wild-type cells adhere tenaciously to positively charged beads, whereas uncharged pvg<sup>+</sup> mutant cells remain indifferent to them. B, SAP staining assay of fixed cells. SAP staining of wild-type cells is diagnostic for the presence of pyruvylated terminal Galp1,3-Gal epitope on the cell surface. Mutant strains were derived from the G418-resistant, GFP-expressing shuttle vector pSP2 (ars<sup>Ⅰ</sup>−, S. cerevisiae URA3) (ATCC 77298) (32) or the high copy S. pombe shuttle vector pSP2 (ars<sup>Ⅰ</sup>−, S. cerevisiae URA3) (ATCC 77298) (37). Details of cloning and expression plasmid construction are available in the Supplemental Material. The expression plasmids pUR–1–5-GFP and pSP–1–5-GFP were transformed into the respective pvg<sup>+</sup>–pvg<sup>−</sup> mutants and tested for complementation of mutation and restoration of wild-type phenotype as described above. The cellular location of the expressed Pvg(1–5)-GFP fusion proteins was determined by fluorescence microscopy.

A G418-resistant h<sup>−</sup> strain was designed by random integration of an EcoRI–Ascl kanMX6-GFP cassette into the S. pombe (h<sup>−</sup> urad<sup>4</sup> leave1–32 ade6–704) strain. Briefly, the pF6a-kanMX6-P3nmt1-GFP plasmid (38) (a kind gift from Dr. Jurgy Balder) was digested with EcoRI and Ascl, and a 3.5-kbp fragment containing the kanMX6 marker (39), together with the S657/GFP gene (40) under transcriptional control of the nmt1 promoter (41), was isolated using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). The cassette was transformed into the h<sup>−</sup> urad<sup>4</sup> leave1–32 ade6–704 strain and replica-plated on YEDS/G418 plates. G418-resistant, GFP-expressing h<sup>−</sup> colonies were used for mutagenesis and subsequent mating analysis.

Transcription Analysis by End Point RT-PCR—End point RT-PCR was performed to determine whether transcription of pvg1–pvg5 was significantly affected in any of the mutants and whether Pvg4p, a putative transcription factor (see “Results”), was involved in regulation of one or more of pvg1–pvg3 and pvg5. Briefly, nucleic acids were isolated from wild-type and pvg<sup>−</sup> mutant cells grown to early stationary phase in YEDS. Following DNase I (Invitrogen) digestion to remove genomic DNA, cDNA was synthesized from RNA transcripts using the SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen), utilizing SuperScript II RT and oligo(dT)<sub>12–18</sub>. PCR primer was performed with Platinum TaqDNA polymerase (Carlsbad, CA) and appropriate gene-specific primers for pvg1–pvg5 and act1 (act<sup>Ⅰ</sup>) control (Table I). Difficult amplification of pvg3 due to secondary structure effects necessitated the use of the forward primer pvg3-TMD, which did not incorporate the transmembrane domain coding sequence. Both positive controls (no DNase I) containing genomic DNA and negative controls (no SuperScript II RT) were utilized.

Isolation and Characterization of Mutants Defective in Synthesis of Negatively Charged N-Linked Poly saccharides—Since no phosphate or sulfate has previously been detected in S. pombe N-linked glycans (12, 13), pyruvylate appears to be the only negatively charged functional group on the cell surface. Based on this assumption, we isolated glycosylation mutants using the inability to bind to Q-Sepharose beads as the primary screen (Fig. 1A; c.f. wt and pvg<sup>−</sup>). Those mutants with highly underglycosylated or undergalactosylated galactomannans, which are known to be highly osmosensitive (42, 43), were effectively excluded from further examination by repeated DDW washes. We applied biotinylated SAP staining of fixed cells to further screen mutant colonies for the absence of pyruvylated Gal (Fig. 1B). For a fortuitous coincidence, serum amyloid P component exhibits calcium-dependent lectin-like affinity for the PvGal epitope, an apparently nonphysiological ligand (44, 45). About 70 colonies, unable to bind to Q-Sepharose beads and unreactive with SAP, were obtained from five separate ethyl methanesulfonate mutagenesis trials. We performed lectin blot analyses (not shown), with biotinylated concanavalin A and biotinylated GSL/BSL 1-B<sub>4</sub>, to roughly determine the composition of the glycan present on the glycoproteins of the SAP-negative strains. The mutant strains used for further investigation had overall levels of mannosylation and α-galactosylation similar to the wild-type strain, suggesting that the mutations only affected the relatively infrequent PvGal modification found in S. pombe N-linked glycans.
Transformation of these strains with an S. pombe genomic library (32) resulted in classification into five different groups, each complementable by one of five plasmids. Partial sequencing of their respective genomic inserts revealed that all five plasmids contained open reading frames predicted to encode proteins of unknown functions (Table II). When cloned into pUR-KAN (Fig. S1, Supplemental Material) for expression, each coding sequence, with its upstream promoter region, allowed full complementation of only its corresponding complementation group and restored expression of pyruvylated N-linked glycans, as determined by Q-Sepharose binding and SAP staining of fixed cells. We designated the five genes pvg1−
pvg5 and registered these names with the S. pombe GenDB data base (35, 36) (available on the World Wide Web at www.genedb.org/genedb/pombe/index.jsp).

Whereas ~98% of the large wild-type galactomannans are negatively charged (12, 13), the glycans from the five pvg− strains were totally incapable of binding Q-Sepharose beads. This suggested a complete, or nearly complete, lack of negatively charged functionalities on these glycans. Compositional analysis, by acid hydrolysis and HPAEC-PAD, of the large N-linked glycans from representatives of each of the five complementation groups showed that all had monosaccharide compositions similar to those of wild-type cells (not shown). This corroborated lectin blot results and confirmed that neither the disruption of the pvg gene products nor artifacts from the random mutagenesis dramatically impacted the metabolic pathways responsible for the bulk of N-linked glycosylation in these mutants and that the lack of glycan negative charge did not result from a complete Gal deficiency.

Wild-type S. pombe galactomannans are very large; the “average” composition of Endo H-released galactomannans is \( \text{Pv}_{5-6}\text{Gal}_{12-25}\text{Man}_{22-62}\text{GlcNAc} \) (13, 28). Approximately 80% of the Man residues of the outer chain poly-Mana,1,6-backbone are occupied by Galol,1,2-side chains, of which about 10% are capped by PvGalβ1,3- (Fig. 2) (13, 28). Unsubstituted β1,3-linked Gal is not found in the wild type. Therefore, about 8% of outer chain extension Man residues in toto are substituted by PvGalβ1,3Galol,2,1, and all pyruvate in S. pombe galactomannans is associated with this appendage.

Lectin blot analyses (not shown) of purified glycoproteins from the wild-type and four mutant strains, using concanavalin A and GSL/BSL I-B, indicated almost equivalent glycoprotein glycan profiles for all. PNA, which has a 10-fold higher affinity for β-linked Gal than for α-linked Gal (51), showed a significantly greater staining of pvg1− glycoproteins than of wild-type or pvg2−pvg5− glycoproteins, indicative of the presence of terminal Galβ1,3- (Fig. 3). However, incubation of pvg1− glycoproteins with β-galactosidase did not decrease PNA staining (not shown). To confirm this unexpected result, we prepared depyruvylated Galβ1,3-containing glycans by exposing wild-type galactomannans to mild acid hydrolysis. Digestion of these glycans with β-galactosidase yielded no appreciable Gal release, as determined by HPAEC-PAD monosaccharide analysis (not shown) and NMR oligosaccharide analysis (not shown). These results, taken together with those of PNA lectin blotting, imply that the Galβ1,3- residues are inaccessible to this enzyme.

Since the pvg1−pvg5− mutant galactomannans were observed to be roughly as abundant in Gal as wild-type galactomannans, the absence of the negative charge could be predicted to arise from either of two mutant structural phenotypes, one missing the terminal pyruvyl group but retaining the terminal Gal and the other devoid of the entire (relatively infrequent) PvGalβ1,3-cap. NMR analysis of wild-type and pvg1−pvg5− galactomannans confirmed the presence of both structural expec-
tations (Fig. 4). Wild-type and pvg1− glycans exhibited the 4.67-ppm anomic proton signal of Galβ1,3-(β) (shifted upfield to 4.64 ppm in pvg1−) and the 4.30-ppm C4-H resonance of 3-substituted Galol,1,2-(3G) (shifted upfield to 4.25 ppm in pvg1−). However, the signal for pyruvate protons (Pv) at 1.54 ppm was absent in the pvg1− glycans. The glycans from pvg2−pvg5− mutants displayed neither of these particular proton signals and are thus devoid of the entire PvGalβ1,3-cap.

All five pvg sequences were determined from the S. pombe gene data base to specify relatively small deduced proteins (Table II). All five genes lack introns, a somewhat surprising result, considering that 43% of S. pombe genes contain introns, and such genes have an average of ~2.2 introns/gene (35, 36). By comparison, only 5% of S. cerevisiae genes contain introns, and such genes have an average of ~1.1 introns/gene (35, 36). A collection of bioinformatic algorithms were queried with the deduced protein sequences of the five pvg genes. The predicted (unmodified) molecular weights and isoelectric points, cellular locations, transmembrane domains and topology, and functions of each of the encoded proteins are presented with their genome data base annotations (Table II). With the exception of Pvg4p, the proteins are predicted to have Type II-oriented single-pass transmembrane domains. Such proteins have a cytoplasmic N terminus and a C terminus within the lumen of a secretory compartment or within the periplasm for a yeast plasma membrane protein. Several neural net glycosylation prediction algorithms were also canvassed (Table III), although they were not trained specifically with yeast proteins.

To determine subcellular localization of Pvg1p−Pvg5p, we expressed C-terminal GFP fusions of each protein in the respective pvg1−pvg5− mutants as described above and in the Supplemental Material. The transformants exhibited growth rates similar to that of wild-type cells. Each of the Pvg1−pvg5− GFP fusion proteins complemented its corresponding lesion, as determined by Q-Sepharose binding and SAP staining of fixed cells. Images showing the localization of some of these proteins were obtained (Fig. 5) and are discussed below. However, low expression levels from the pUR-based plasmids or possibly post-translational processing events (e.g. misfolding of GFP) prevented sufficient fluorescence for localization of Pvg2p-GFP or Pvg5p-GFP.

To examine the effect of high copy expression of the C-terminal GFP-tagged proteins, we constructed pSP2-based expression plasmids as detailed in the Supplemental Material. The pSP-(1−3,5)GFPc transformants grew about half as fast as did their pUR-(1−3,5)GFPc counterparts. The pSP-4GFPc transformants were extremely slow growing and appeared to experience difficulties with cell division, giving rise to elongated cells. Three of the Type II transmembrane constructs (Pvg1,2,5p-GFP) successfully complemented their corresponding lesions. The Pvg3p-GFP fusion construct, specifying a probable β1,3-galactosyltransferase, as well as the extremely slow growing pSP-4GFPc transformant, failed to restore charge. With the exception of pSP-2GFPc, we were unable to obtain detectable GFP signal from the pSP-GFPc transformants. Thus, overexpression of Pvg-GFP fusion proteins from pSP2-based plasmids led in most cases to cellular disruption and misfolding of the Pvg and/or GFP portion.

**Characterization of pvg1** **Mutant**—A sequence similarity search by standard protein-protein BLAST (33, 55) showed that the predicted amino acid sequence of Pvg1p shares an ~300−350-aa stretch of about 30% identity and about 45% similarity with nearly 20 identified or predicted pyruvyltransferrases. These include the PssK “exopolysaccharide polymerization” proteins of *Rhizobium leguminosarum* bv. *trifolii* and *R. leguminosarum* bv. *viciae* (56) and the Streptococcus ther-
### TABLE II
Identity of *pvg1*-*pvg5* genes, putative functions, and bioinformatic results

| Gene name | Systematic name (open reading frame size, bp) | Deduced protein size | Predicted* $M_r$ (pI) | Predicted cellular locationb,c (+, low; ++, moderate; +++, high) | Predicted transmembrane domain and topologyd | Predicted* and putative function/annotatione | Number of mutants isolated |
|-----------|---------------------------------------------|----------------------|-----------------------|---------------------------------------------------------------|---------------------------------------------|-----------------------------------------------|---------------------------|
| *pvg1*    | SPAC8F11.10c (1206)                         | 401                  | 44,871.65 (5.20)      | Extracellular, including cell wall +++, Signal peptide +++, Anchor signal +++, Cleavage site (aa 30/31) ++ | Type II TMD (aa 12–29) | Hydrolase, transferase Pyruvyltransferase No *S. cerevisiae* ortholog | 3                         |
|           |                                             |                      |                      |                                                               |                              |                                               |                           |
| *pvg2*    | SPAC27E2.07 (1170)                          | 389                  | 45,065.48 (5.54)      | Golgi +/+ Endoplasmic reticulum ++ Signal peptide ++ Anchor signal ++ Cleavage site (aa 32/33) ++ | Type II TMD (aa 13–35) | Ligase                                        | >50                       |
|           |                                             |                      |                      |                                                               |                              |                                               |                           |
| *pvg3*    | SPBC1921-06c (1137)                         | 378                  | 43,834.13 (5.52)      | Endoplasmic reticulum ++ Signal peptide ++ Anchor signal ++ Cleavage site (aa 24/25) + | Type II TMD (aa 9–28) | Transferase, hydrolase Galactosyltransferase | 8                         |
|           |                                             |                      |                      |                                                               |                              |                                               |                           |
| *pvg4*    | SPBC317.01 (1119)                           | 372                  | 41,572.09 (7.34)      | Nuclear +++                                                | NA                         | Transcription regulation MADS-box/SMF-type TF | 3                         |
|           |                                             |                      |                      |                                                               |                              |                                               |                           |
| *pvg5*    | SPAC22F8.02c (1119)                         | 372                  | 43,086.70 (7.67)      | Cytoplasmic ++/Mitochondrial ++ Signal peptide ++ Anchor signal ++ Cleavage site (aa 35/36) ++ | Type II TMD (aa 13–35) | Transferase, oxidoreductase Unknown No apparent orthologs, possibly *S. pombe*-specific | 4                         |

*Unmodified (i.e. co- or post-translational modifications); average isotopic $M_r$, pI.

b PSORT II (prediction of protein cell localization) (46).

b SignalP2.0-HMM (prediction of signal peptide, signal anchor, and cleavage sites) (47).

b TMHMM 2.0 (prediction of transmembrane helices) (48).

b ProtFun 2.1 (prediction of cellular role, enzyme class, and gene ontology in eukaryotes) (49).

b Pfam 10.0 (prediction of protein family) (50).

S. pombe GeneDB (available on the World Wide Web at www.genedb.org/genedb/pombe/index.jsp).

b NA, not applicable.
mophilus EpsL protein (57), the latter including a shorter 118-aa portion with 39% identity and 56% similarity. Among the reported sequences were a number of “polysaccharide polymerization” and “exopolysaccharide biosynthesis” proteins, including those from Nostoc sp., Leuconostoc mesenteroides, Oenococcus oeni, Burkholderia fungorum, Actinobacillus pleuropneumoniae, and Bacillus subtilis. Reverse position-specific (RPS) BLAST (58) confirmed the presence of a ~300-aa conserved pyruvyltransferase (pfam4230) domain (Pfm protein family data base) (50). Of great interest and with the highest scoring standard BLAST sequence homology to Pvg1p (including a shorter 128-aa stretch with 43% identity and 57% similarity) is a “possible galactosyltransferase” from Cryptosporidium parvum. This relatively large (827-aa) predicted protein appears to be bifunctional, with an N-terminal galactosyltransferase (pfam01762) domain similar to that of Pvg3p and the identified C-terminal pyruvyltransferase domain.

Both one- and two-dimensional NMR analysis show the presence of unsubstituted Galβ1,3- residues in N-linked galactomannans of the pvg1- mutant (Fig. 4), indicating that 1) PvGal epitope synthesis occurs by subsequent transfer of the pyruvyl group to Galβ1,3- residues already present on the N-linked galactomannan chains and does not require the formation of some PvGal-containing precursor, and 2) Pvg1p is the only protein among the five discovered that is required for pyruvate addition and is putatively a pyruvyltransferase. A functionally active C-terminal GFP fusion of Pvg1p appears to localize to the cell periphery, where it may act just prior to glycoprotein secretion and incorporation into the cell wall (Fig. 5A).

The S. pombe GeneDB annotation for pvg1 indicates that there is no S. cerevisiae analog (Table II), as would be expected from the absence of pyruvate in S. cerevisiae glycoprotein mannos. Other bioinformatic results support both a cell membrane location and a pyruvyltransferase function for Pvg1p (Table II). A salient bioinformatic prediction is a cluster of two other highly homologous proteins with significant similarity to Sur1p, and one near the C terminus. Sur1p transfers Man to inositolphosphorylceramide (60) and possesses a 159-aa middle section having 26% identity and 42% similarity with a few glycosyltransferases, including yeast Och1p. Och1p is the α1,6-mannosyltransferase, with ho-
Five Genes Involved in S. pombe PvGal Epitope Biosynthesis

Table III

| Gene product name | Predicted N-linked glycosylation sitesa | Predicted O-GalNAc mucin-type glycosylation sitesb | Predicted O-β-GlcNAc/O-phosphate "yin-yang" sitesd |
|-------------------|----------------------------------------|-----------------------------------------------|--------------------------------------------------|
| Pvg1p             | Asn-38 + (X is Pro)                     | Ser-40, Ser-41, Thr-43, Ser-44, Ser-46, Ser-48, Thr-49 | NAa                                               |
|                   | Asn-211 ++                              |                                               |                                                  |
|                   | Asn-269 +                               |                                               |                                                  |
| Pvg2p             | Asn-230 ++                              | Ser-338                                      | NA                                               |
| Pvg3p             | Asn-53 ++                              | None                                         | NA                                               |
| Pvg4p             | NA                                     | N/A                                          | 5++++/11++/13 + O-β-GlcNAc glycosylation sites, of which 13 are predicted "yin-yang" sites |
| Pvg5p             | Asn-38 +                                | Ser-284                                      | NA                                               |
|                   | Asn-138 +                               |                                               |                                                  |

a NetNGly 1.0 (prediction of N-linked glycosylation sites in human proteins) (52).
b NetOGly 2.0 (prediction of O-GalNAc mucin-type glycosylation sites in mammalian proteins) (53).
c YinOYang 1.2 (prediction of O-β-GlcNAc glycosylation and "yin-yang" sites for eukaryotic intracellular/nuclear proteins) (54).
d NA, not applicable.

Fig. 5. Expression and localization of Pvg1p-Pvg4p C-terminal GFP-fusion proteins. Pvg1p-GFP exhibits a pericellular location; Pvg2p-GFP has a polar location; Pvg3p-GFP has a cytoplasmic location; and Pvg4p-GFP has a nuclear location. Expression was in pUR19, except for Pvg2p-GFP, which was in pSP2. The Pvg5p-GFP fusion expressed from both pUR19 and pSP2 successfully complemented the pvg5" mutation but failed to exhibit fluorescence.

modologs in both S. cerevisiae and S. pombe, responsible for the first step of outer chain elongation of N-linked oligosaccharides (61). An N-terminal transmembrane domain is predicted for budding yeast Och1p but not for the fission yeast homolog. RPS BLAST also revealed the presence of a conserved Och1p-related domain, which encompasses Och1- and Sur1-like proteins, as well as S. cerevisiae Hoc1p (homologous to Och1p) (62). Hoc1p is a component of one of two mannan polymerase complexes involved in outer chain a1,6-Man elongation (63, 64).

Sur1p, Och1p, and Hoc1p all possess the tripeptide DXD sequence that coordinates an Mn2+ ion, which plays a central role in catalytic utilization of a nucleotide saccharide donor in most, but not all, glycosyltransferases (65–67). However, when aligned against Sur1- and Och1-like proteins, Pvg2p shows a noticeable absence of the expected DXD motif, having SRS instead (aa 187–189). Indeed, of the five predicted Pvg proteins that we identified, only Pvg3p exhibits the DXD sequence.

The S. pombe GeneDB annotation for pvg2 indicates that it is possibly S. pombe-specific, with no apparent orthologs (Table II). The role of Pvg2p remains unknown, although NMR results show that it is essential for Galβ1,3- addition (Fig. 4). A C-terminal GFP fusion was localized primarily to the cell poles, consistent with deposition of newly synthesized galactomannans in growing cells (Fig. 5B). Paradoxically, other bioinformatic results predict localization of Pvg2p to the membrane of a secretory compartment (i.e. Golgi and endoplasmic reticulum), in apparent contradiction to empirical results (Table II). Pvg2p is also ascribed a ligase function of unknown specificity.

Characterization of pvg3\(^{-}\) Mutant.—Although standard protein-protein BLAST of the predicted amino acid sequence for Pvg3p indicated some homology to known and putative glycosyltransferases as well as to other proteins, it was indecisive. Three iterations of PSI BLAST clarified a strong homology to several hundred β1,3-glycosyltransferases, particularly with many glycoprotein (UDP-Gal: β-GlcNAc) β1,3-galactosyltransferases from humans, Old World hominoids, and a variety of mammals and higher eukaryotes. RPS BLAST confirmed that pvg3 encodes a protein with a conserved galactosyltransferase (pfam01762) domain of the glycolipid transferase family 31 (GT31) in the Carbohydrate Active Enzyme database (66). The GT31 family of over 200 structurally related inverting β1,3-glycosyltransferases, which contain a “GT-A” or “SpsA” Rossmann-like fold (66–68), is predominantly composed of (UDP-Gal: β-GlcNAc) β1,3-galactosyltransferases from species ranging from yeasts to plants to humans (69, 70). As mentioned earlier, of special interest is the presence in the GT31 family of a putative bifunctional enzyme from C. parvum, with an N-terminal galactosyltransferase domain similar to that of Pvg3p, and a C-terminal pyruvyltransferase (pfam04230) domain similar to that of Pvg1p described above.

Pvg3p is the only member of the GT31 family predicted to be present in S. pombe. This corresponds with the finding that galactomannans purified from pvg3\(^{-}\) cells lack both pyruvate and β1,3-linked Gal (Fig. 4), since the PvGal epitope contains the only β-linked Gal. The expressed Pvg3p-GFP fusion showed punctate labeling typical for localization within organelles of the cytoplasmic compartment (Fig. 5C), a pattern also seen for S. pombe Golgi-resident Gma12p, an α1,2-galactosyltransferase (71). These data, along with the homology to known β1,3-galactosyltransferases, strongly imply that Pvg3p is the β1,3-galactosyltransferase that adds Gal to the outer chain Galα1,2- residues of galactomannan.

Other bioinformatic results support both a secretory com-
partment location and a galactosyltransferase function for Pvg5p (Table II).

Characterization of pvg4− Mutant—Standard protein–protein BLAST of the deduced Pvg4p amino acid sequence indicated strong homology to over 100 MADS-box transcription factors, particularly homoeotic plant proteins. The MADS-box domain is an ancient and conserved contiguous sequence of ~56 amino acids found, generally N-terminally, in transcriptional regulation proteins in species ranging from yeasts to plants to humans (72). RPS BLAST confirmed the presence in Pvg4p of a MADS-box (pfam00319) domain.

MADS-box transcription factors bind to A/T-rich cis-acting regulatory elements, which are classified into two general linear (73–75). Members of the serum response factor (SRF)-like Type I subtype bind to the canonical “CARG” sequence 5′-CC(A/T)gG-3′, whereas those of the myocyte-specific enhancer factor 2 (MEF2)-like Type II subtype bind to the DNA sequence 5′-T(T/A)gATAR-3′. Three iterations of PSI BLAST returned Pvg4p homology to over 1000 MADS-box proteins in the MEF2-like Type II subfamily. A search for short, nearly exact matches confirmed this as the strongest alignment, particularly with the four human MEF2A, B, C, and D. The alignment, generated by ClustalW 1.8 (78), shows a strong homology in α-helix I and the first portion of β-strand I, implying that Pvg4p binds to a MEF2-like Type II MADS-box-binding sequence. Weak homology in the remainder of β-strands I and II indicates that Pvg4p has a distinctive dimerization area. On the C-terminal side of the MADS-box (not shown), α-helix II of Pvg4p may bind distinct accessory factors.

There are two other MADS-box transcription factor genes predicted to be in S. pombe, one (SPBC19G7.06) with no S. cerevisiae ortholog and of unknown function and the other the pheromone receptor Map1p, involved in mating-type determination (82, 83). The latter gene is a functional homolog of the S. cerevisiae transcription factor Mcm1p (79, 84) and is also homologous to the S. cerevisiae Arg80p (85). However, both Mcm1p and Arg80p belong to the SRF-like Type I MADS-box subfamily (75, 85). The other S. cerevisiae MADS-box transcription factors are encoded by SMP1 and RLM1, which belong to the MEF2-like Type II MADS-box subfamily (75). Rlm1p is activated by phosphorylation in the (Mpk1/Slt2) mitogen-activated protein kinase pathway (86, 87). Of particular relevance, most if not all of the target genes for Rlm1p are involved in cell wall biosynthesis and remodeling during growth and differentiation and as a result of cell integrity signaling induced by environmental stresses (88, 89). Similarly, Smp1p is phosphorylated in the stress-activated (Hog1) mitogen-activated protein kinase pathway (90). However, whereas the phosphorylation sites are similar for Rlm1p and Smp1p, and whereas the two proteins are known to heterodimerize, they do not appear to have overlapping functions (86, 88, 90).

Three PSI BLAST iterations of the non-MADS-box remainder of Pvg4p (i.e. aa 57–372) indicated very strong homology to dozens of DNA-directed RNA polymerase II (large subunit) proteins. A search of this portion for short, nearly exact matches revealed a midprotein sequence (aa 158–193) extremely rich in proline, histidine, and aromatic amino acids. Analysis of this sequence indicates several regions that might serve as ligand areas for proteins containing SH3 (Src homology 3) or WW (tryptophan-tryptophan) domains (91). Proteins containing these proline-rich motifs are involved or postulated as adaptors between membrane signal receptors and cytoplasmic or cytoskeletal factors, acting as mediators of intracellular signal transduction (91).

A functionally active Pvg4p-GFP fusion protein showed strong general nuclear localization with no apparent preferential labeling of nuclear substructures (Fig. 5D). One-dimensional NMR of N-linked glycans purified from pvg4− cells indicated that Pvg4p function is required for Galβ1,3-addition (Fig. 4). Other bioinformatic results confirm both a nuclear location and a transcription regulation function for Pvg4p (Table II). A striking prediction is the presence of nearly 20 O-linked β-GlcNAc sites (Table III). Many of these are putative “yin-yang” sites for the dynamic reciprocity between O-β-GlcNAc attachment and O-phosphorylation (92). Yin-yang sites have been shown to be involved as control nodes for cellular regulation in multicellular eukaryotes (93, 94), for example, the MADS-box transcription factor SRF (95). However, direct evidence of O-β-GlcNAc-modified proteins in simple eukaryotes remains elusive (93, 94).

Characterization of pvg5− Mutant—Standard protein–protein BLAST of the predicted Pvg5p amino acid sequence provided no insight into the possible function of Pvg5p in PvGal synthesis.
Neither three iterations of PSI BLAST nor a search for short, nearly exact matches contributed helpful information. However, an interesting observation is that a C-terminal 151-aa portion of Pvg5p shows 20% identity and 39% similarity with a C-terminal section of Pvg1p. Nevertheless, these proteins do not possess overlapping biosynthetic functions, since the pvg5 gene failed to complement the pvg1- mutant phenotype, and vice versa. Based on NMR data, Pvg6p is essential for Galβ1,3- addition, whereas Pvg1p is required for terminal pyruvylation.

As in the case of pvg2, the S. pombe GeneDB annotation for pvg5 indicates no apparent orthologs for the possibly S. pombe-specific gene (Table II). Other bioinformatic results suggest a cytoplasmic or mitochondrial location for Pvg6p (Table II), although the former location is inconsistent with the predicted Type II transmembrane domain. A transferase or oxidoreductase function is also proposed.

Transcription Analysis by End Point RT-PCR—End point RT-PCR of transcripts of each of the five pvg genes (and act1 control) in each of the five pvg- mutant strains indicated that the translatable regions of each gene were transcribed at a level and size comparable with that of the wild-type strain (Fig. 7). Most notably, Pvg4p does not appear to solely control transcription of any of the other four genes (pvg1–pvg3 and pvg5). Therefore, despite the presence of potential MEF2-like Type II MADS-box-binding promoter sequences upstream from pvg1 and pvg5, expression of the transcripts appears unaffected. Negative controls (no RT) indicated no genomic DNA contamination of the mRNA samples (not shown).

Mating Analysis—It has been reported that S. pombe strains greatly deficient in surface Gal are incapable of sexual conjugation (43). To test whether the PvGal epitope plays any role in the sexual behavior of S. pombe cells, we spot-mixed pvg- mutant strains of both mating types (h+ urad4–294) and G418-resistant (h+ urad4+ leu1–32 ade6–704), defective in synthesis of pyruvylated galactomannans, on malt extract medium plates. After 3 days of incubation at room temperature, each cross-pattern (h+ × h+, h+ × h−, h− × h+, h− × h−) showed an almost equal number of crescent-shaped zygotic asci. Cell mixtures from each cross were replated on CSM-Leu/G418 medium. After 2 days of incubation at 33 °C, each cross-pattern showed an almost equal number of colonies (Fig. 8). Microscopic examination indicated these colonies to be prototrophic haploid crossover recombinants; diploid cells were not observed. No colonies were observed in the control (uncrossed) quadrants of the plates. Therefore, neither the terminal pyruvate (i.e. pvg1- mutant) nor the complete PvGal epitope (i.e. pvg2–pvg5- mutants) of cell wall N-glycans appears to play any significant role in growth or mating (i.e. conjugation, meiosis, and sporulation). Interestingly, however, pvg1- cells with exposed Galβ1,3- residues showed an increased propensity to agglutinate at late stationary phase of growth in liquid medium.

DISCUSSION

We have identified in S. pombe at least five genes that contribute proteins that participate in the cell surface PvGal epitope synthesis pathway. Whereas we cannot at this point propose a complete mechanism for PvGal biosynthesis, the data are most consistent with a nearly quantitative addition of pyruvate directly to Galβ1,3- residues already present on the N-linked galactomannan chain rather than by a presynthesis of PvGal-containing intermediates and transfer of PvGal to terminal outer chain Galα1,2- residues. Since unsubstituted β1,3-linked Gal was found in pvg1- glycans (Fig. 4), we suggest, based on sequence homology, that PvGl1p is a pyruvyltransferase. Of the four additional proteins that we found to be required for Galβ1,3- addition (Pvg2p–Pvg5p), only the one containing the putative galactosyltransferase domain, Pvg3p, seems a likely candidate to be the required β1,3-galactosyltransferase. The fact that a predicted Carbohydrate Active Enzyme data base GT31-family
protein acts in the PvGal production pathway is particularly interesting, because the S. pombe Galβ1,3Gal epitope shares structural similarity with the mammalian glycoprotein glycan structures (Galβ1,3GlcNAc, Galβ1,3GlcNAc, GaINAcβ1,3Gal, GlcNacβ1,3Gal, and GlcNacβ1,3GalNAc), which are synthesized by other GT31 enzymes (66).

Whereas ascribing a regulatory role for Pvg4p, a predicted MADS-box transcription factor, in the transcription of pvg1–pvg3 and pvg5 seemed reasonable, the PvGal synthesis system does not appear to be solely under its control. No obvious canonical SRF-like Type I or MEF2-like Type II cis-acting regulatory elements for MADS-box binding were discovered in the promoter regions of the pvg2 and pvg3 genes, although each transcription factor possesses a distinct binding specificity. Potential MEF2-like Type II sequences were identified in the pvg1 and pvg5 promoters. However, expression analysis by end point RT-PCR, particularly the finding that transcription of pvg1–pvg3 and pvg5 is not significantly affected in the pvg4 mutant, does not clarify the situation. MADS-box-containing proteins specifically recruit other factors into multicomponent heteromeric regulatory complexes. It is reasonable to postulate that one or more transcriptional modulators that interact with Pvg4p as well as other Pvg4p target genes remain yet to be identified.

The functions of gene products Pvg2p and Pvg5p remain enigmatic. The annotated S. pombe genome suggests no apparent orthologs for pvg2 and pvg5, which may be S. pombe-specific (Table II). Both genes are predicted to encode transmembrane proteins, with Pvg2p experimentally located to the plasma membrane at the cell poles but Pvg5p not yet localized. Whereas bioinformatic algorithms predict a Golgi or endoplasmic reticulum location for Pvg2p, a cytoplasmic or mitochondrial location is proposed for Pvg5p (Table II). If Pvg2p and/or Pvg5p is located in the periplasmic space, one or both may interact with Pvg1p or with each other and have a role, perhaps as lectins, within a plasma membrane-bound enzyme complex that adds the terminal pyruvate to galactomannans. However, this does not fully explain why the absence of either protein results in the lack of the entire PvGal epitope and not merely the terminal pyruvate. Both or either of the two proteins may also coordinate at the plasma membrane directly or indirectly with the Pvg4p MADS-box-containing transcription factor, as part of a cell wall integrity signal transduction process. If Pvg2p and/or Pvg5p is indeed Golgi-resident, one or both may interact with the Pvg3p β1,3-galactosyltransferase in a manner akin to that of members of the S. cerevisiae mannosyltransferase activity complexes (63, 64).

The Bähler group (96) has used whole genome DNA microarrays to determine which S. pombe genes are involved in a core environmental stress response (i.e. genes with significantly changed expression in response to free radical, heavy metal, thermal, osmotic, or redox stress). None of the pvg genes showed significant changes during these stresses. Similarly, the same group has also identified genes having significantly altered expression during sexual differentiation of S. pombe (i.e. conjugation, meiosis, and sporulation). The PvGal epitope synthesis genes pvg2, -3, and -5 (and possibly also pvg1) were determined to be induced “early” in sexual differentiation (97). It is important to note that the RNA for the RT-PCR experiments that we have reported here was isolated from cultures in early stationary phase and not from cultures under the nutrient-depleted conditions required to instigate sexual differentiation. The Bähler group’s findings are especially interesting in light of subsequent findings that S. pombe-specific genes such as pvg1, -2, and -5, as compared with genes that are well conserved across species, are disproportionately induced during sexual development (98).

S. pombe mutants of Gal epimerase gps2 and UDP-Gal transporter gns1 are both viable but are osmosensitive with aberrant cell morphology (42, 43), indicating that the presence of Gal, and hence the PvGal epitope, is not essential. However, some possible recognition function for the epitope cannot be discounted. It was tempting to speculate that PvGal and/or the Galβ1,3Gal epitope is involved in the mating process. Mutant pvg1− cells do have an increased tendency, as compared with the pvg2− or pvg5− mutants, to agglutinate when reaching late stationary phase, which is consistent with the observations that galactose or wild-type galactomannans, but not mannose, glucose, or galactose-deficient gns1 mannan, inhibit calcium-dependent nonssexual flocculation (99, 100). In addition, mannose-dependent co-flocculation with Pediococcus damnosus, which occurs at high levels for S. cerevisiae but at very low frequency for wild-type S. pombe (due to shielding by terminal Gal and PvGal), was greatly enhanced in the galactose-deficient gns1 S. pombe (101). Although galactose has been reported to promote sexual flocculation (100) and conjugation (43), we demonstrated that cells totally devoid of pyruvate as well as those lacking the complete PvGal epitope mate with frequencies comparable with that of wild-type cells, at least under laboratory conditions. We hypothesize that there is functional discrimination between the penultimate Gal1,2- residue and the terminal Galβ1,3- residue, with the presence of penultimate Gal1,2- being necessary for sexual flocculation and with terminal Galβ1,3- involvement in nonssexual flocculation behavior. The addition of Pv to Galβ1,3Galα1,2- thus impedes nonssexual flocculation.

In other organisms pyruvylated saccharides appear to have various functional roles. Pyruvylolation of the secreted exopolysaccharides of Xanthomonas sp. contributes to the physical and chemical properties of xanthan gum (1, 19). In Rhizobium sp. and Sinorhizobium sp., acidic exopolysaccharides are necessary for symbiotic nodulation of root hairs for nitrogen fixation (22, 102). Decreases in pyruvic acid content correlate with ineffective nodule formation (103), whereas increases in pyruvyltransferase activity correlate with an increased level of PvGal and with root hair attachment (104). Lactic acid bacteria also secrete exopolysaccharides, and, in S. thermophilus, disruption of the putative pyruvyltransferase EpsL in rory strains results in nonviability (57). In apparent contradiction, nonropy strains do not possess EpsL but exhibit typical cell growth. A widely used strategy among bacteria (e.g. B. anthracis) may be the addition of pyruvate to peptidoglycan-associated polysaccharides, which appears to facilitate the anchorage of several cell surface proteins by noncovalent binding to an S-layer homology domain (23). However, Bacillus licheniformis polysaccharide contains pyruvylated Gal, but it is not covered by an S-layer lattice (105).

We do not expect to have identified all of the genes involved in or affecting PvGal synthesis. In order to select against gross glycosylation deformities, we intentionally and effectively excluded osmosensitive cells by the lack of an osmotic support (e.g. mannitol and KCl) in the mutant enrichment process. Mechanistic links appear to remain absent. In bacterial polysaccharide synthesis, cytoplasmic phosphoenolpyruvate is employed as the high energy pyruvate donor to produce ketal-linked pyruvylated monosaccharides in the saccharide repeating units of the pyrophosphoryl-lipid-anchored precursors (20, 21, 106–109). However, in S. pombe, as in all eukaryotic cells, glycoprotein glycan production is a highly compartmentalized process, in which attachment of terminal monosaccharides occurs in the lumen of the Golgi apparatus. We were
Five Genes Involved in S. pombe PvGal Epitope Biosynthesis

Unable to identify reports of either pyruvic acid or energetically favorable pyruvate intermediates in the lumens of the endoplasmic reticulum or Golgi. Indeed, any such compounds would probably require specific transporters to enable them to traverse the membrane from the cytoplasm. Conversion of pyruvate into a high energy intermediate within the Golgi would almost certainly require several unique luminal proteins. Given the uncertainty as to date and the structures of their N-linked glycans (Figs. 2 and 4), it seems reasonable to predict that there should be at least one pyruvate transporter and pyruvate-intermediate processing enzyme in the Golgi apparatus. Alternatively, based on our evidence that Pvg1p is located at the plasma membrane, pyruvylolation of terminal Galα1,3- in the outer chain extensions may occur there in the periplasm, where there may be a role for Pvg2p and/or Pvg5p.

Experiments to identify the relationships among the identified Pvg proteins and to further elucidate the mechanisms of PvGal epitope synthesis are ongoing. These include but are not limited to the search for possible additional genes and proteins; efforts to knock out, individually and in various combinations, the five identified pvg genes; in situ expression of GFP fusion or epitope-tagged proteins; cloning, purification, and further biochemical characterization of the proteins; development of assays to determine their biochemical and biological functions; and site-directed mutagenesis to examine structure-function associations.

Acknowledgments—We are indebted to Dr. Tony Carr for graciously providing the S. pombe genomic library and for helpful advice on plasmid recovery, and we gratefully acknowledge the gift of pFA6a-kanMX6 plasmids from Dr. Jürg Bähler. We acknowledge use of the Wadsworth Center Structural Biology NMR Facility and of the Wadsworth Center Advanced Light Microscopy Core Facility. We thank Dr. Alexei Khodjakov for performing the fluorescence microscopy in Fig. 5 and Dr. Alexey Khodjakov for performing the fluorescence microscopy in Fig. 5. Oligonucleotide synthesis and DNA sequencing were executed in the Wadsworth Center Molecular Genetics Core. We express our gratitude to Dr. John Cipollo for valuable suggestions during the course of this work. We thank Frank Abruscato for help with molecular techniques and Catherine Lubowsky for excellent technical assistance.

References

1. Becker, A., Katenz, F., Puhler, A., and Ielpi, L. (1998) Appl. Microbiol. Biotechnol. 50, 145–152
2. Chandrasekaran, R., Lee, E. J., Tai, S., Kafigt, N., and Zeevenhuizen, M. J. (1994) J. Carbohydr. Res. 261, 287–298
3. Leslie, M. R., Parolis, H., Parolis, L. A. S., and Petersen, B. O. (1998) Carbohydr. Res. 309, 95–101
4. Behault, G. M., Choy, Y. M., Dutton, G. S. S., Funnell, N., Stephen, A. M., and Yang, M. T. (1973) J. Bacteriol. 113, 1345–1347
5. Cerantola, S. Hounsey, J. D., Seongs, C. M., and Montouzie, H. (2000) FEMS Microbiol. Lett. 185, 243–246
6. Gemmill, B. J. and Trimble, D. (1995) Biochim. Biophys. Acta 1246, 227–237
7. Cutler, J. E. (2001) Med. Mycol. 39, Supplement 1, 75–86
8. Pardis, L. A., Duus, J. O., Parolis, H., Meldal, M., and Bock, K. (1996) Carbohydr. Res. 293, 161–177
9. Alviano, C. S., Travassos, L. R., and Schauer, R. (1999) Glycoconjug. J. 16, 454–554
10. Soares, P. M. A., Soares, R. M. de A., Alviano, D. S. A., Angluster, J., Alviano, C. S., and Travassos, L. R. (2000) Biochim. Biophys. Acta 1474, 262–268
11. Gemmill, B. J. and Trimble, D. R. (1996) J. Bacteriol. 271, 2954–2959
12. Gemmill, B. J. and Trimble, D. R. (2004) in The Molecular Biology of Schizosaccharomyces pombe: Genetics, Genomics, and Beyond (Egel, R., ed.) pp. 13–25, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
13. Cottarell, G., Beach, D., and Deutschel, U. (1993) Curr. Genet. 23, 547–548
14. Bähler, J., Wu, J.-Q., Longtine, M. S., Shah, N., McKee, A., III, Steever, A. B., Wach, A., Philippova, P., and Pringle, J. R. (1998) Yeast 14, 953–961
15. Longtine, M. S., McKee, A. J., Demari, D. J., Shah, N. G., Wach, A., Brachat, A., Philippova, P., and Pringle, J. R. (1998) Yeast 14, 953–961
16. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) Mol. Biol. Cell 6, 845–846
17. Tanaka, N., Komoni, M., Osumi, M., and Takegawa, K. (2001) Yeast 17, 1003–1014
18. Hind, C. R. K., Collins, P. M., Renn, D., Cook, R. B., Caspi, D., Baltz, M. L., and Peps, M. B. (1984) J. Exp. Med. 159, 1058–1069
19. Hind, C. R. K., Collins, P. M., Baltz, M. L., and Peps, M. B. (1985) Biochem. J. 225, 107–111
20. Nakai, K. (2000) Adv. Protein Chem. 54, 277–344
21. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Int. J. Neural Syst. 8, 581–599
22. Keogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001) J. Mol. Biol. 305, 567–580
23. Jensen, L. J., Gupta, R., Blom, N., Devos, D., Tamames, J., Kemsir, C., Nielsen, H., Stærfeldt, H. H., Rapacki, K., Workman, C., Andersen, C. A. F., Knudsen, S., Krogsgaard, A., Rodrigues, A., and Brunak, S. (2002) J. Mol. Biol. 319, 1257–1274
24. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) Nucleic Acids Res. 32, D138–D141
25. Swamy, M. J., Gupta, D., Mahanta, S. R., and Surulir, A. (1991) Carbohydr. Res. 215, 59–67
26. Gupta, R., and Brunak, S. (2002) PLoS. Comput. Biol. 8, 310–322
27. Jensen, L. J., Lund, O., Herlufsen, N., Gøker, A. A., Williams, K. L., and Brunak, S. (1998) Glycoconjug. J. 15, 115–130
28. Blom, N., Gammeltøft, S., and Brunak, S. (1999) J. Mol. Biol. 294, 1351–1362
29. Elkon, C. B., Madsen, T. S., Hahn, M., Wang, B. Y., Miller, W. J., and Li, J. D. (1997) Nucleic Acids Res. 25, 3389–3402
30. Sadykov, M. R., Ivashina, T. V., Kanapin, A. A., Shliapnikov, M. G., and Repovich, R. A. (1999) FEBS Lett. 453, 297–304
31. Stengel, F., Neisser, J.-R., and Moll, B. (1996) J. Bacteriol. 178, 1680–1690
32. Marchler-Bauer, A., Anderson, J. B., DeWeese-Scott, C., Fedorova, N. D., Geer, L. Y., He, S., Hurwit, D. I., Jackson, D. J., Jacobs, A. R., Lanczycki, C. J., Liebert, C. A., Liu, C., Madje, T., Marchuk, G. H., Mazurier, R., Nikolayska, A. N., Panchenko, A. R., Rao, B. S., Shoemaker, B. A., Simonyan, V., Song, J., Thiesen, P. A., Vasudevan, S., Wang, Y.,
