Structure
Article

The Conformation and Function of a Multimodular Glycogen-Degrading Pneumococcal Virulence Factor

Alicia Lammerts van Bueren,[1,5] Elizabeth Ficko-Blean,[1] Benjamin Pluvinage,[1] Jan-Hendrik Hehemann,[3,6] Melanie A. Higgins,[1] Lehua Deng,[2] A. David Ogunniyi,[4] Uwe H. Stroeher,[4] Nahida El Warry,[1] Robert D. Burke,[1] Mirjam Czjzek,[3] James C. Paton,[4] David J. Vocadlo,[2] and Alisdair B. Boraston[1,*]
[1]Biochemistry & Microbiology, University of Victoria, P.O. Box 3055 STN CSC, Victoria, BC V8W 3P6, Canada
[2]Department of Chemistry, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada
[3]Station Biologique de Roscoff, Marine Plants and Biomolecules, CNRS-UPMC Paris 6, BP74, 29682 Roscoff, France
[4]Research Centre for Infectious Diseases, School of Molecular and Biomedical Science, University of Adelaide, South Australia 5005, Australia
[5]Present address: Netherlands Cancer Institute (NKI-AVL), Cell Biology II–B6, Plesmanlaan 121, 1066CX Amsterdam, The Netherlands
[6]Present address: Biochemistry and Microbiology, University of Victoria, P.O. Box 3055 STN CSC, Victoria, BC V8W 3P6, Canada
*Correspondence: boraston@uvic.ca
DOI 10.1016/j.str.2011.03.001

SUMMARY

SpuA is a large multimodular cell wall-attached enzyme involved in the degradation of glycogen by the pathogenic bacterium *Streptococcus pneumoniae*. The deletion of the gene encoding SpuA from the bacterium resulted in a strain with reduced competitiveness in a mouse model of virulence relative to the parent strain, linking the degradation of host-glycogen to the virulence of the bacterium. Through the combined use of X-ray crystallography, small-angle X-ray scattering, and inhibitor binding, the molecular features involved in substrate recognition by this complex protein are revealed. This uniquely illustrates the complexity of the active site, the conformational changes incurred during carbohydrate binding by this protein, and the interaction and cooperation of its composite modules during this process. New insight into the function of this particular pneumococcal virulence factor is provided along with substantial contributions to the nascent framework for understanding the structural and functional interplay between modules in multimodular carbohydrate-active enzymes.

INTRODUCTION

One of the largest class of carbohydrate-active enzyme is the glycoside hydrolases, members of which use a hydrolytic mechanism to cleave glycosidic bonds (Zechel and Withers, 2001), and through this activity participate in processes of particular relevance such as those that impact the global carbon cycle (e.g., microbial degradation of plant cell wall materials; Bayer et al. [1998]) and those that impact human health (Gloster and Davies, 2010). With respect to the latter, one noteworthy area of examination regards the role of glycoside hydrolases in the destruction of host glycans by microbial invaders of the human body. *Streptococcus pneumoniae*, a prevalent and serious bacterial pathogen, is one such microbe that relies on its use of extracellular glycoside hydrolases to process a variety of host glycans for optimal virulence (King, 2010; Higgins et al., 2009).

The extracellular pneumococcal glycoside hydrolase virulence factors all possess complex multimodular structures, a property generally considered as the presence of multiple structurally and functionally distinct units within a single polypeptide. One module of a glycoside hydrolase houses the catalytic machinery; in addition to this, such an enzyme may contain numerous ancillary modules that confer carbohydrate-binding activity (carbohydrate-binding modules or CBMs), protein-protein interactions, cell-wall anchoring, and/or modules of unknown function. A question critical to understanding the biological function of carbohydrate-active enzymes is how the functions of the modules are spatially coordinated to achieve a specific outcome. This question is most powerfully addressed by the analysis of three-dimensional structure. However, the structural heterogeneity of flexible, multimodular proteins has complicated the application of structural methods to these proteins, and thus, there are comparatively few complete structures of multimodular carbohydrate-active enzymes relative to the number of single-module structures. The best represented class of multimodular glycoside hydrolase for which structures have been determined are those from glycoside hydrolase family 13. Examples such as the pullulanase from *Bacillus acidopullulyticus* (Turkenburg et al., 2009), SusG from *Bacteroides thetaiotaomicron* (Koropatkin and Smith, 2010), the pullulanase (KpPulA) from *Klebsiella pneumoniae* (Mikami et al., 2006), and α-amylase 1 from *Thermoactinomyces vulgaris* (Abe et al., 2004) all reveal the relative dispositions of some or all of their composite modules, including the variety of CBMs found in these proteins. Although structures such as these are highly informative, they rarely capture cooperation between the composite modules. Recent advances in the application of solution small-angle X-ray scattering (SAXS) to protein structure, particularly combined with X-ray crystallography (Putnam et al., 2007), make it a complementary approach that more generally enables the complete

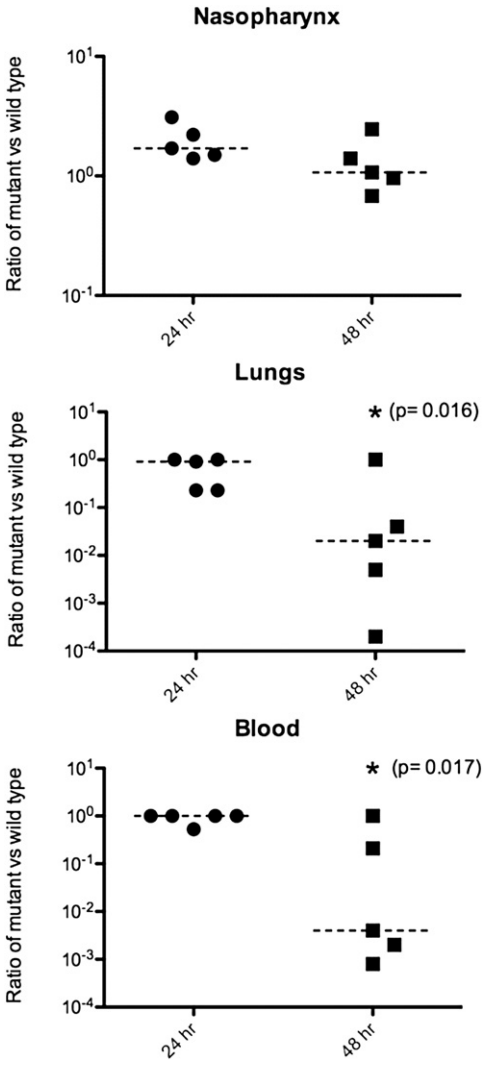

Figure 1. In Vivo Competition Index Analysis of *S. pneumoniae* D39 Δ*spuA* Mutant in Three Different Host Tissues at Two Different Times Post-Intranasal Inoculation

Each symbol indicates the CI for an individual animal. The dashed line represents the mean CI of five animals. Those mean CI values that are significantly different from one are indicated.

analysis of proteins with complex modular architectures, including carbohydrate-active enzymes (Ficko-Blean et al., 2009; Hammel et al., 2004), which should ultimately improve our understanding of multimodular proteins.

SpuA is a cell wall-attached protein that, like many of its GH13 family members, comprises multiple distinct modules (Bongaerts et al., 2000; van Bueren et al., 2007). This enzyme hydrolyzes the α-1,6-branchpoints of glycogen, and the growth of *S. pneumoniae* on an exogenous supply of this α-glucan is dependent upon the presence of a functional *spuA* gene (Abbott et al., 2010). A unique fused dyad of noncatalytic family 41 CBMs at the N terminus of SpuA binds tightly to glycogen, including glycogen in the context of type II alveolar cells in lung tissue (van Bueren et al., 2007). SpuA is part of a group of complementary proteins involved in α-glucan metabolism, which are all considered part of a candidate core pneumococcal genome (Obert et al., 2006), and many of which have been identified as putative virulence factors in large-scale screens (Abbott et al., 2010; Hava and Camilli, 2002). The conservation among pathogenic streptococci of SpuA homologs, which collectively comprise subfamily 12 of glycoside hydrolase family 13, suggests the significance of glycogen degradation to streptococcal virulence (Santi et al., 2008; Stam et al., 2006). The structure of a catalytically active truncated version of the *S. agalactiae* homolog of SpuA has been determined (Gourlay et al., 2009); however, the mode of glycogen recognition by this complex, multimodular class of protein remains unknown. This knowledge gap and the possible relevance of SpuA to pneumococcal virulence prompted us to confirm the protein's role in virulence. A thorough structure-function analysis of the enzyme, incorporating X-ray crystallography in combination with SAXS, reveals the structural changes required of SpuA to recognize glycogen and the unprecedented role of a CBM in contributing to the architecture of the enzyme's active site.

RESULTS

SpuA Is a Virulence Factor

To confirm the results of the large-scale signature-tagged mutagenesis studies indicating that SpuA is important to the pneumococcus-host interaction, we assessed the ability of a nonpolar deletion replacement *spuA* mutant of *S. pneumoniae* D39 (serotype 2) to compete with the wild-type D39 strain in a mouse model. The similar numbers of wild-type and *spuA* mutant bacteria isolated from the nasopharynx of intranasally (i.n.) inoculated mice at 24 and 48 hr postinfection suggest that SpuA plays little or no role in colonization (Figure 1). However, the significantly reduced numbers of *spuA* mutant bacteria in the lungs and blood relative to wild-type 48 hr postinfection ($p = 0.016$ and 0.017, respectively) confirm a role for SpuA in invasive disease. This observation interpreted in light of the glycogen-degrading properties of SpuA provides robust evidence for the link between pneumococcal virulence and the ability of this bacterium to attack exogenous glycogen. Nevertheless, it remains unclear if SpuA is catalytically active on glycogen present in the cells of its host. We approached this question by using immunofluorescence microscopy analysis of a human type II alveolar cell line (A549), which is a common model for the analysis of adherence and invasion by *S. pneumoniae*. When labeled and used as a probe for glycogen, *Thermotoga maritima* CBM41, a high-affinity α-1,4 glucan-specific binding protein (Lammerts van Bueren et al., 2004), gave intense fluorescent foci corresponding to glycogen granules in all of the samples (Figure 2). Cells left untreated or treated with an inactive nucleophile mutant of SpuA had a diffuse fluorescence signal in addition to the foci (Figures 2A and 2C). Treatment of fixed and permeabilized cells with recombinant SpuA completely removed the diffuse fluorescence, but not the foci, resulting in an overall decrease in fluorescence (Figures 2B and 2D). An analysis of the size of the glycogen granules (foci) using area, circumference, and feret diameter indicated no change in the glycogen granules upon treatment with SpuA (see Figure S1 available online). This suggests that SpuA is relatively inactive on the intracellular glycogen granules and,

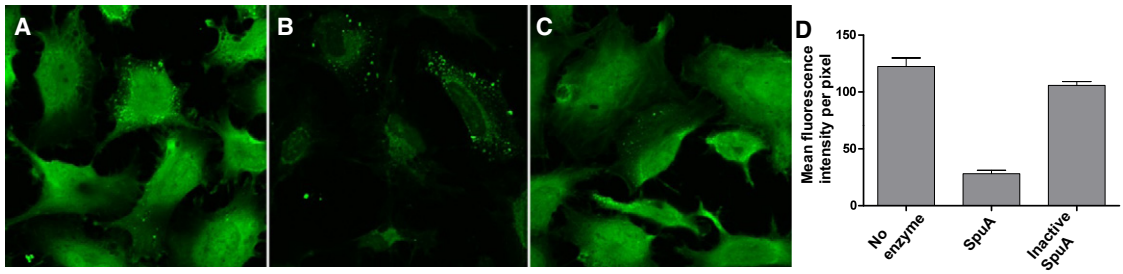

Figure 2. Degradation of Intracellular Glycogen in A549 Cells by Recombinant SpuA

(A) Fixed and permeabilized A549 cells probed with *T. maritima* CBM41 (TmCBM41), an α-glucan specific probe.

(B) Fixed and permeabilized A549 cells treated with SpuA then probed with TmCBM41.

(C) Fixed and permeabilized A549 cells treated with catalytically inactive SpuA then probed with TmCBM41.

(D) Quantification of cytoplasmic fluorescence. Error bars represent the mean (± standard error of the mean) fluorescence per unit area for a sampling of five randomly selected regions of cytoplasm.

A supplemental analysis is found in Figure S1.

instead, is quite active on a more dispersed form of glycogen present in the cytoplasm of host cells.

The X-ray Crystal Structure of SpuA Reveals Its Complex Modular Architecture

The X-ray crystal structure of a truncated recombinant form of SpuA (SpuAΔCBM) comprising only the catalytic domain (the central $(\alpha/\beta)_8$ barrel and its two flanking immunoglobulin-type modules) was initially determined to 1.85 Å resolution (Figure S2). The structure of a nearly full-length recombinant form of SpuA in complex with maltotetraose was subsequently determined to 2.25 Å resolution. This crystal structure reveals the organization of six of native SpuA's seven domains: the two CBMs (CBM41-1 and CBM41-2); a linker domain of which only approximately 65% of its 96 residues could be modeled (residues 362–367, 375–381, 391–402, and 412–418 were disordered could not be built); and the catalytic domain (Figure 3). The CBMs are folded over the top of the $(\alpha/\beta)_8$ barrel such that the CBM41-1 module caps the active site and creates a deep cleft between the CBM and the $(\alpha/\beta)_8$ barrel (Figures 3B and 3C). In this cavity was clear electron density into which could be built two molecules of maltotetraose (Figures 3B and 3C). A second maltooligosaccharide was found associated with CBM41-2, though only three of the four glucose units in the bound maltotetraose molecule could be modeled (Figure 3B and 3C).

Structural Insight into the Hydrolysis of the α-1,6-Branchpoints in Glycogen

The two maltotetraose molecules in the active site were bound in a slightly offset parallel fashion (Figure 4A). The top and most exposed maltotetraose molecule was bound quite superficially in the active site, yet maintained several hydrogen-bonding interactions with protein side chains in addition to classic aromatic side chain-pyranose ring interactions (Figure 4A). The second tetrasaccharide was bound deeply in the active site, wedged between the $(\alpha/\beta)_8$ barrel and the CBM, and made extensive interactions with the protein, particularly with respect to interactions with aromatic amino acid side chains (Figure 4A).

The canonical catalytic center of the α-amylase family (glycoside hydrolase family 13) comprises a trio of carboxylate groups provided by Asp or Glu side chains (MacGregor et al., 2001; Turkenburg et al., 2009). In SpuA, Asp778 is found as the catalytic nucleophile, Glu807 as the acid/base residue, and Asp895 as a stabilizing residue in the −1 subsite, which proposed to make interactions with the O_2 and O_3 hydroxyls of the sugar. The reducing end glucose unit of the more deeply buried maltotetraose is located near this catalytic center; however, it appears slightly out of position relative to the placement of the catalytic residues (Figure 4A). To provide additional clarity to the occupation of the "minus" subsites, we sought to obtain a complex of SpuAΔCBM with the compound 4-O-α-D-glucopyranosylmoranoline (G1M) (Maruo et al., 1993). Although G1M was a comparatively weak inhibitor of SpuA, we were able to determine a 2.37 Å resolution structure of a complex with this compound (Figure 4B). An overlay of this with the maltotetraose complex reveals the small shift in the position of the inhibitor relative to the maltotetraose molecule such that G1M engages Asp778 and Asp895 in a manner consistent with their proposed functions (Figure 4C). Although G1M lacks the glycosidic oxygen (O1), its conformation and positioning suggest that an α-configured glycosidic oxygen (O1) of an intact α-glucoside residue, adopting the same conformation in the −1 subsite as G1M, would be in position to hydrogen bond with Glu807, the general acid/base residue, and be α-1,6-linked to a glucose residue in the +1 subsite. Further support for this is given by the overlay of the SpuA complexes with the maltotetraose complex of the *Klebsiella pneumoniae* pullulanase, KpPulA (Mikami et al., 2006), which reveals similar positioning of the sugar components and a catalytic machinery that is poised for activity on the α-1,6-glucosidic linkage of a branchpoint in glycogen (Figure 4D). Consistent with these predicted roles for Asp778 and Glu807 as key catalytic residues of SpuA, substitution of these two residues to create separate Asp778Ala and Glu807Ala variants of SpuA in both cases abolished enzyme activity (Figure S3).

A Unique Role for CBMs

The CBMs of SpuA are joined to the catalytic region via the linker domain with the CBMs oriented so that CBM41-2 is distal from the active site and facing outward (Figure 3). The presence of a maltooligosaccharide bound to this module reveals that it remains capable of recognizing an α-glucan in the context of the complete enzyme, yet because of its distance from the active

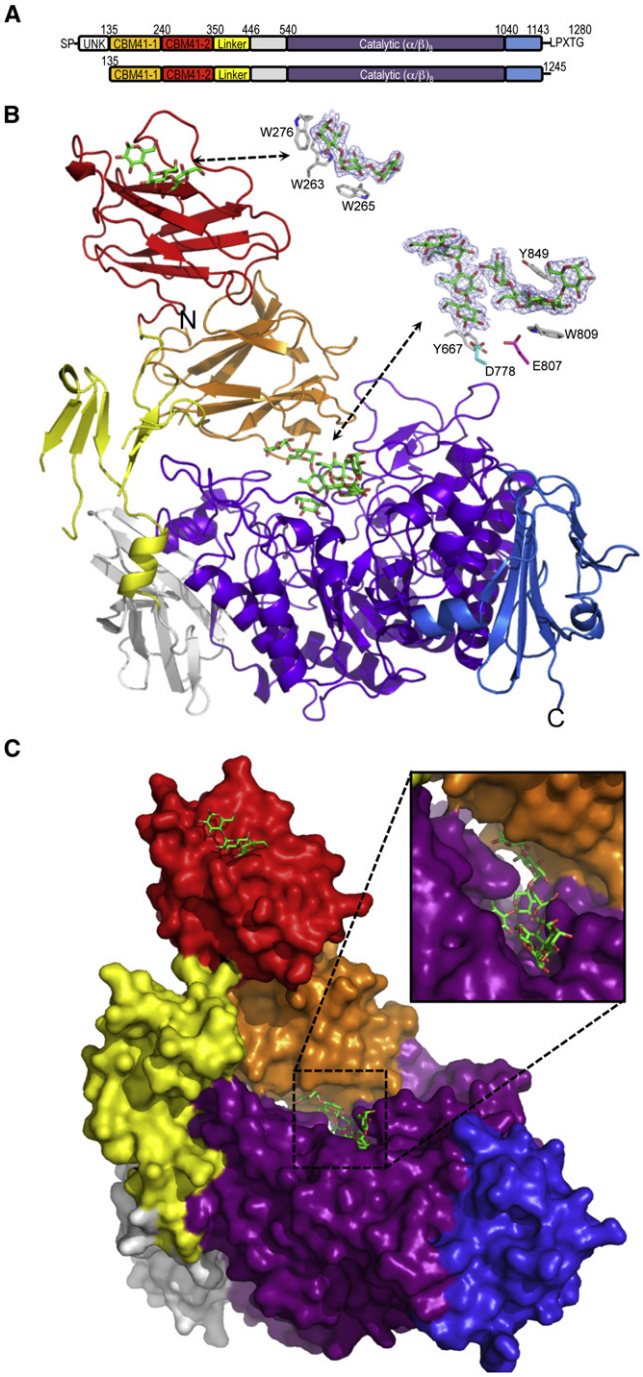

Figure 3. Structural Analysis of SpuA

(A) The modular organization of SpuA with the crystallized construct shown beneath.

(B) The 2.25 Å X-ray crystal structure of SpuA determined in complex with maltotetraose (modules are colored according to A). The sugars are shown as green sticks. Electron density for sugars bound to CBM41-2 (red module) and the catalytic site (the purple module) are shown as blue mesh in the insets. The maps are maximum-likelihood/σ_a-weighted $2F_{obs}$-F_{calc} maps contoured at 1σ (0.25 electrons/Å^3). Relevant aromatic residues are shown as gray sticks. The residues acting as the acid/base and nucleophile in the catalytic site are shown as pink and blue sticks, respectively.

(C) The solvent-accessible surface of SpuA with the bound maltooligosaccharides shown as sticks. The inset shows an expansion of the active site. Modules are colored as in (A) and (B).

site, likely has no direct effect on catalysis (Figure 3). In contrast the CBM41-1 module is positioned close to the catalytic center and contributes to forming a portion of the enzyme active site (Figures 3 and 4E). The three tryptophan residues that comprise the curved surface of the CBM41-1 binding site appear to make an integral part of substrate-binding subsites −3 and −4, suggesting that the CBM plays a role in the catalytic function of SpuA (Figure 4E). Such an intimate association of a CBM with a catalytic center is unprecedented among carbohydrate-active enzymes. This observation prompted us to probe the functional relevance of this architecture by comparing the activity and inhibition of SpuA and the SpuAΔCBM construct, which lacks the CBMs, through the use of compounds that are predicted to occupy these −3 and −4 subsites and, thus, interact directly with CBM41-1.

SpuA catalyzed hydrolysis of *para*-nitrophenyl-α-D-maltopentaoside (pNP-M5, Figure 5A), a commonly used α-amylase substrate (Eckhart, 1982), to release the colored nitrophenolate group would necessitate its occupation of subsites +1 (by the pNP aglycon group) to −4 (by the maltopentaoside glycon), and perhaps a −5 subsite if present (Figure 5A). As an alternative probe of the active site, we synthesized a novel α-glucanase inhibitor comprising maltotetraose and an O-(D-glucopyranosylidene)amino N-phenylcarbamate (PUGLU) residue at the "reducing" terminus (referred to as PUGMAL, Figure 5B) (for details of the synthesis, see the Supplemental Experimental Procedures). The trigonal C-1 of the PUGLU moiety is thought to mimic the transition state structure stabilized by such enzymes. Accordingly, these inhibitors typically bind tightly and such that the trigonal C1 of the PUGLU group is in the −1 subsite, and the oxime group spans the position normally occupied by the scissile glycosidic bond. Consequently, PUGMAL is predicted to interact with the same subsites as pNP-M5 (Figure 5B).

The kinetic analysis of pNP-M5 hydrolysis by SpuA yielded K_M and k_{cat} values of 360 ± 60 μM and 270 ± 10 min^{-1}, respectively (Figure 5C). The corresponding values of K_M and k_{cat} for SpuAΔCBM were 2100 ± 300 μM and 280 ± 20 min^{-1}, respectively. The equivalent k_{cat} values for the two proteins show that SpuA and SpuAΔCBM catalyze hydrolysis of the glycosidic bond equally well at saturating concentrations of substrate. However, the 6-fold difference in K_M suggests that the deletion of the CBMs significantly compromises the ability of SpuA to bind substrate. Consistent with this interpretation, SpuA and SpuAΔCBM showed different binding to PUGMAL. When using pNP-M5 as a substrate, PUGMAL was found to be a competitive inhibitor of SpuA and gave a K_i of 18.8 ± 0.3 μM (Figure 5D). Isothermal titration calorimetry (ITC) gave a similar K_d at 24 ± 4 μM (Figure 5F). The limiting quantities and lower affinity of PUGMAL for SpuAΔCBM only allowed us to approximate the K_i at 170 μM (Figure 5E), though the K_d determined by ITC, 120 ± 10 μM (Figure 5G), was in fair agreement, giving us confidence in the inhibition constant. Thus, the truncation of the CBMs in SpuAΔCBM results in an approximately 5- to

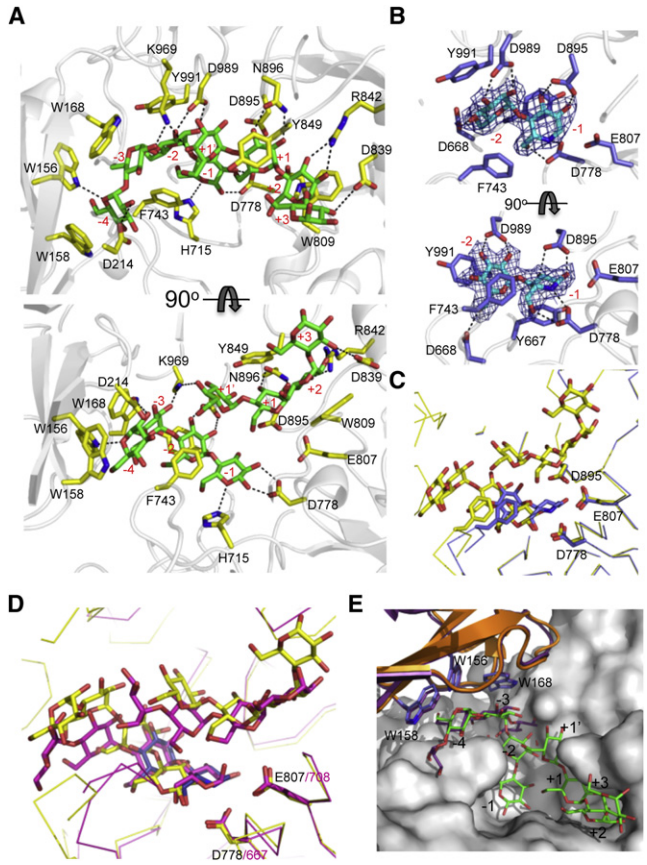

Figure 4. Architecture of the SpuA Active Site and Recognition of Sugars

(A) Two views of the SpuA active site with two maltotetraose molecules (green sticks) occupying eight different subsites (defined according to Davies et al., 1997). Key residues in the active site are shown as yellow sticks and potential hydrogen bonds as black dashed lines.

(B) Two views of the active site of SpuAΔCBM with a single-bound molecule of G1M (cyan sticks) occupying two subsites. Key residues in the active site are shown as dark-blue sticks and potential hydrogen bonds as black dashed lines. The blue mesh shows the electron density map for G1M as a maximum-likelihood/σ_a-weighted $2F_{obs}$-F_{calc} map contoured at 1σ (0.31 electrons/Å^3).

(C) An overlay of the SpuA structure in complex with maltotetraose (yellow) and the SpuAΔCBM-G1M complex (dark blue). The residues acting as the acid/base (Glu807) and nucleophile (Asp778) are shown as sticks for reference.

(D) An overlay of the SpuA structure in complex with maltotetraose (yellow), the SpuAΔCBM-G1M complex (dark blue; the protein is omitted for clarity), and the *Klebsiella pneumoniae* pullulanase in complex with maltotetraose (magenta). The residues acting as the acid/base and nucleophile are again shown as sticks for reference.

(E) The active site of SpuA in complex with maltotetraose with the components contributed by the catalytic module shown as a solvent-accessible surface. The CBM41-1 is shown as an orange ribbon with the tryptophans it contributes to the active site shown as dark-blue sticks. The sugar is shown as green sticks, and the subsites are labeled. The previously determined structure of CBM41-1 in complex with maltotetraose (van Bueren et al., 2007) is shown in purple and overlapped with CBM41-1 of the SpuA structure.

Supplemental data are found in Figure S2.

8-fold decrease in affinity for the inhibitor, which is consistent with the 6-fold change in K_M for pNP-M5. Taken together, these results reveal that deletion of the CBMs and, thus, loss of potential interactions at −3 and −4 subsites to some degree compromise the ability of the enzyme to recognize carbohydrates.

The activity of SpuA and SpuAΔCBM was further explored by semiquantitative HPAEC-PAD analysis of glycogen hydrolysis by these proteins. Both proteins produced the same series of oligosaccharides upon glycogen hydrolysis (data not shown), as was previously described (Abbott et al., 2010). However, analysis of the production of maltooligosaccharides from maltotriose (M3) to maltooctaose (M8) over the reaction time course showed that relative to SpuA, SpuAΔCBM was roughly half as effective at releasing these oligosaccharides from glycogen (Figure 6; Figure S4). Thus, in addition to playing a role in the recognition of smaller compounds, such as pNP-M5 and PUGMAL, the CBMs also have a role in promoting degradation of the highly polymerized substrate glycogen.

These results corroborate the proposed structural role of CBM41-1 in forming part of the active site and contributing to substrate recognition. However, the position of CBM41-1 in the crystal structure blocks off the active site immediately beyond the −4 subsite, which is made particularly clear through comparisons of the previously determined structure of CBM41-1 in complex with maltotetraose with the structure of CBM41-1 in the context of the SpuA active site (Figure 4E). Contradictory to this is the observed recognition of pNP-M5, PUGMAL, and glycogen, which necessitates that sugar chains extend beyond the −4 subsite. Consequently, it appeared to us that molecular motions in SpuA must enable movement of the CBMs relative to the catalytic domain so as to accommodate longer chains in the minus subsites and generally allow substrate to access the deep active site cleft. Unfortunately, uncomplexed SpuA was, despite extensive attempts, completely recalcitrant to crystallization, so we approached this hypothesis through the analysis of the solution structure properties of SpuA with SAXS.

SpuA Undergoes Conformational Changes

Three different concentrations of SpuA were analyzed by solution SAXS (Table 1). For each concentration of protein, 20 ab initio models were generated with DAMMIF, which gave average χ values of 1.4–2.1 and normalized spatial discrepancies of 0.6–0.7. The averaged ab initio form of SpuA in solution reveals a triangular shape that qualitatively matches the general shape of the complexed crystal structure of SpuA (Figure 7A). However, the best χ_{crysol} value generated using the complexed crystal structure as a model for the SAXS data was obtained using the data collected on 2.95 mg/ml SpuA, yet the value was relatively poor at 3.4, indicating that the closed form of SpuA trapped in the crystal structure is not an optimal model of SpuA in solution. To generate more suitable pseudoatomic models of uncomplexed SpuA in solution, two approaches were taken. Both approaches required that the coordinates of SpuA obtained from the X-ray crystallographic analysis be separated into two rigid bodies: the CBMs and linker domain as one rigid body (referred to as the N-fragment); and the catalytic domain as a second rigid body (referred to as the C-fragment). In the first approach the N- and C-fragments were manually placed into the ab initio envelope using the positions of the N and C termini that would join the two fragments as a constraint. The relative positions of the fragments were optimized based on calculated χ_{crysol} values. The final

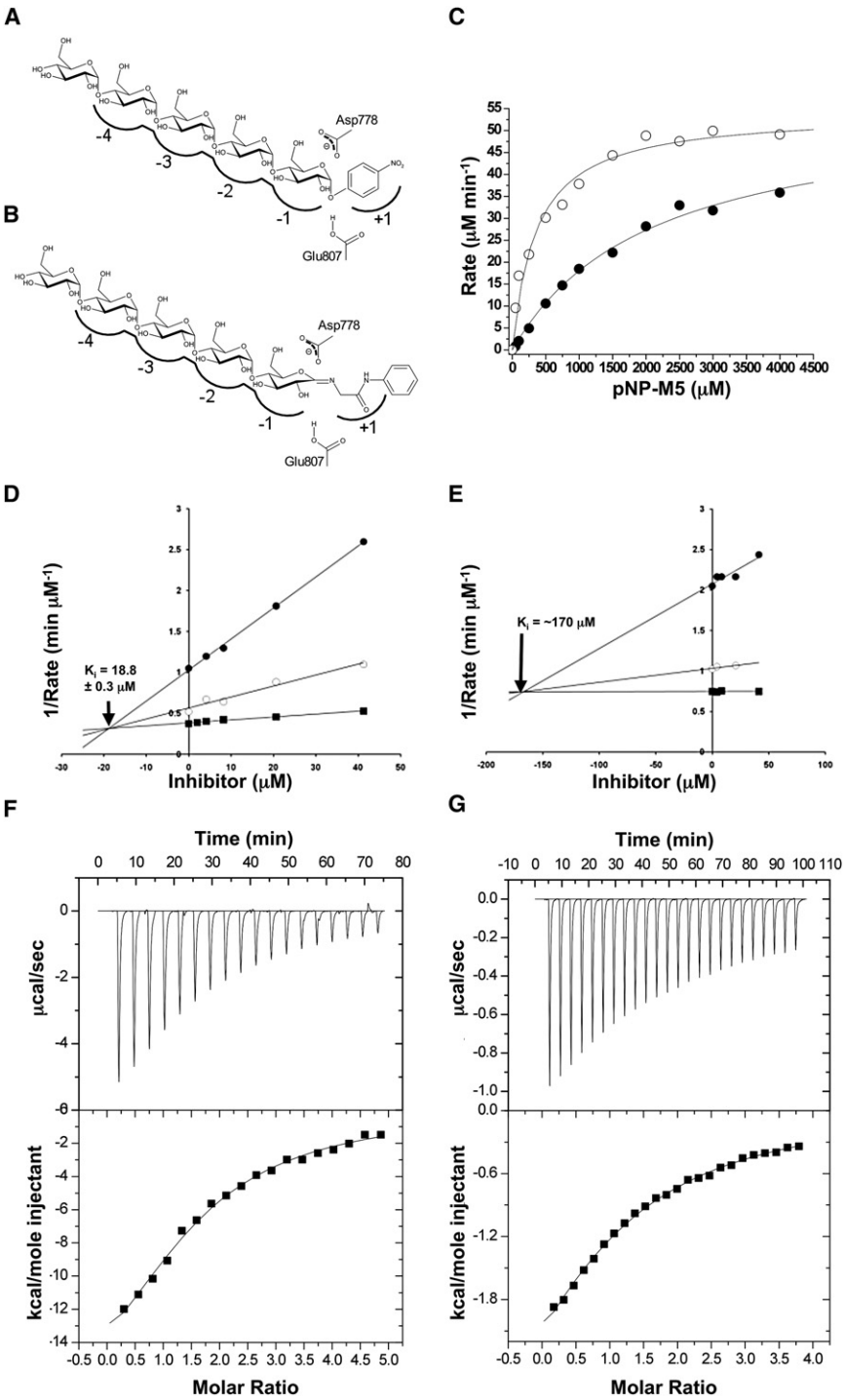

Figure 5. Activity and Inhibition of SpuA and SpuAΔCBM

(A) Schematic showing the structure of pNP-α-maltopentaoside (pNP-M5) and its predicted occupation of subsites. The catalytic residues are shown for reference.

(B) Schematic showing the structure of PUGMAL and its predicted occupation of subsites.

(C) Kinetic curves for the activity of SpuA (open circles) and SpuAΔCBM (closed circles) on pNP-M5. The results of nonlinear fitting to determine the Michaelis-Menton parameters are shown as solid lines.

(D and E) Dixon plots of (D) SpuA and (E) SpuAΔCBM inhibition by PUGMAL are shown. K_i values determined at the intersection points of the lines are shown.

(F and G) ITC analysis of (F) SpuA and (G) SpuAΔCBM binding to PUGMAL is shown. The raw titration data are shown in the top panes, and the integrated heats are shown as closed squares in the bottom panes. The results of data fitting as outlined in the Experimental Procedures are shown as solid lines.

model obtained, referred to as model 1, had a χ_{crysol} value of 1.9 (Figure 7B). Owing to the potentially flexible nature of SpuA in solution and the possible presence of a wide variety of conformations not accounted for in the manual generation of a SpuA model, the second approach used the ensemble optimization method (EOM) (Bernado et al., 2007) to model the SAXS data by treating the polypeptide region between the N- and C-fragments as a flexible hinge. The resulting ensemble of 20 SpuA models yielded an overall χ_{EOM} value of 2.1, with a range of χ_{crysol} values calculated for the individual models between 1.8 and 2.5; the best model with a χ_{crysol} of 1.8 is referred to as model 2 (Figure 7C). Notably, the manually generated model was a relatively close match to several of the individual EOM models. A comparison of the SAXS-based models, both manually generated and the EOM generated, with the complexed crystal structure of SpuA show the possible solution structures of SpuA to be in quite different conformations than in the crystal structure (Figure 7D). The general properties of all of the solution models reveal that the active site of the uncomplexed form is substantially more open, which is achieved through the retraction of the CBMs from the active site by ~20–40 Å. SAXS analysis of SpuA in solution in the presence of excess maltotetraose was also performed, and although the analysis of the data was complicated by the high background scattering of the oligosaccharide, the results suggested a generally closed conformation consistent with the X-ray crystal structure (data not shown).

DISCUSSION

The identification of SpuA as a virulence factor and its activity on glycogen present in a type II alveolar cell line link the glycogen-degrading ability of this protein to the pathogenesis of the

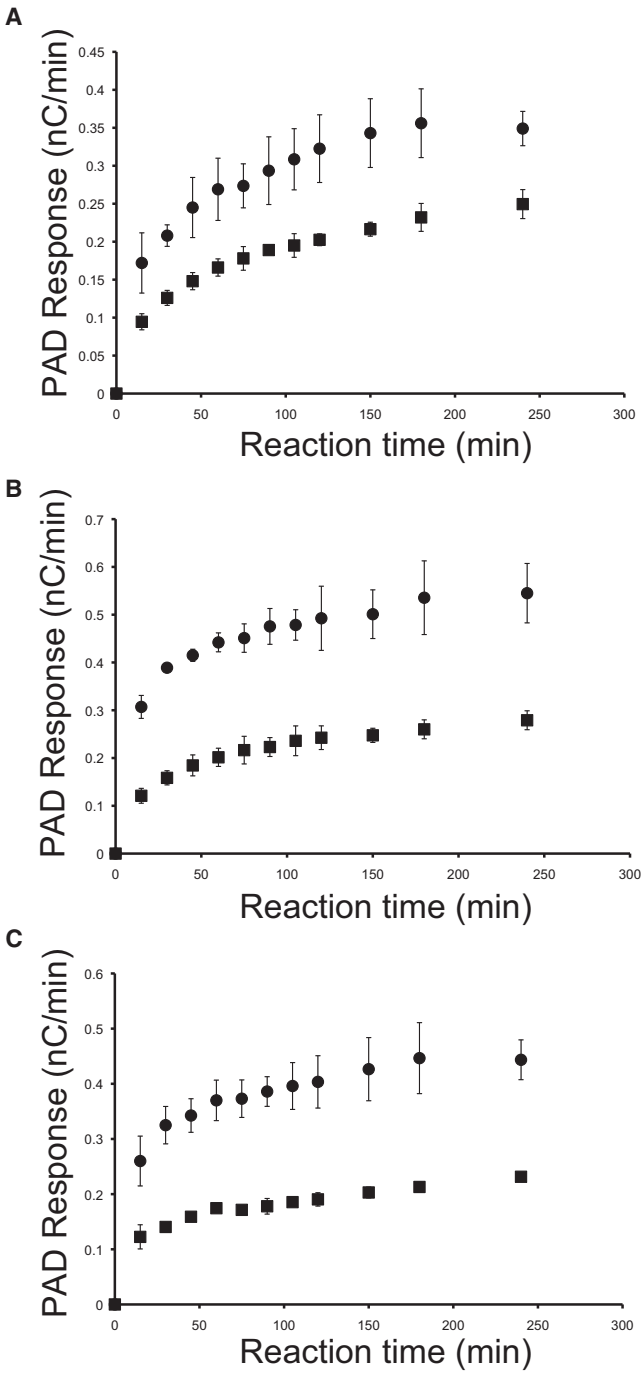

Figure 6. SpuA Catalyzed Degradation of Glycogen

SpuA (circles) and SpuAΔCBM (squares) catalyzed release of maltotetraose (A), maltohexaose (B), and maltooctaose (C) from glycogen over time as measured by HPAEC-PAD. The PAD response refers to the integrated peak area of the oligosaccharide being monitored in nanocoulombs per minute (nC/min). This value is proportional to the concentration of that oligosaccharide; thus, a comparison of the SpuA and SpuAΔCBM curves for release of a given oligosaccharide gives a relative measure of activity. Error bars represent the standard deviation of triplicate measurements. Supplemental data describing the standards there were utilized and additional maltooligosaccharides that were monitored are found in Figure S4.

bacterium. The overall dimensions of SpuA, provided by its structure, suggest that it can extend ~11–12 nm from the peptidoglycan to which it is attached. The capsule layer of the pneumococcus is estimated to extend >100 nm from the cell wall, though this is dependent upon the serotype (Hyams et al., 2010). Therefore, the capsule would present a substantial physical barrier to the access of a large polysaccharide like glycogen to the SpuA active site. Consistent with this proposal, the pneumococcus grows poorly on glycogen in culture, but growth similar to that on glucose is obtained if cultures are supplemented with exogenous recombinant SpuA to depolymerize the glycogen into smaller soluble fragments that could more easily diffuse through the capsule layer (Abbott et al., 2010). The implication is that the effect of SpuA is exerted depending on the capsular status of the bacterium while it is in the host. This further suggests that the optimum efficiency of SpuA is achieved when the pneumococcus is unencapsulated, such as after the bacterium has invaded a host cell (Hammerschmidt et al., 2005), or in a less-encapsulated "transparent" phase, which predominates at mucosal surfaces. Considering that glycogen is an intracellular energy store that is abundant in type II alveolar cells, we favor a model whereby SpuA degrades intracellular glycogen following invasion of this host cell type when the bacterium is unencapsulated, though the degradation of glycogen released by lysed host cells is also plausible (Abbott et al., 2010; van Bueren et al., 2007).

Although structural changes in proteins upon carbohydrate recognition are not uncommon (Quiocho et al., 1997; Nishimasu et al., 2007), SpuA presents a noteworthy case that results mainly from its large size and multimodularity. The glycoside hydrolase family 13 catalytic domain imparts the key hydrolytic activity to the enzyme; however, one of the most remarkable features revealed by the X-ray crystal structure of a nearly full-length fragment of SpuA is the position of its CBMs. The N-terminal CBM, CBM41-1, folds back to cover the active site, which created −3 and −4 subsites and, overall, a complex active site with at least eight subsites through which SpuA recognizes and hydrolyzes the α-1,6-branchpoints in glycogen. Such an architecture has not yet been observed in a carbohydrate-active enzyme, and the consequence is a very deep active site with minus subsites that nearly completely sequester the "arm" of the glycogen branchpoint that is being removed. This indicates complexities in the recognition of glycogen, including, at the extreme, an inability to hydrolyze the branchpoints of granular glycogen where the aggregation of the α-glucan chains and concomitant steric effects would limit accessibility to the deep active site of the enzyme. Consistent with this, SpuA appeared to be inactive on intracellular glycogen granules. Instead, the enzyme was selective for a more dispersed form of cellular glycogen, suggesting that SpuA requires the more accessible branchpoints in less aggregated or soluble glycogen. However, the crystal structure of SpuA shows the active site in this trapped form of the protein to be blocked off beyond the −4 subsites, thus predicting that only branches having four or less sugar residues could be removed from glycogen by SpuA. This is at odds with the observed *endo*-glycogenase activity of SpuA, where the products of glycogen hydrolysis range from glucose to maltooligosaccharides with a very high degree of polymerization (Abbott et al., 2010). Thus, though highly informative, the

Table 1. SAXS Parameters for SpuA

Concentration (mg/ml)	Calculated MW (kDa)[a]	$Rg_{(Guinier)}$ (Å)	$Rg_{(Gnom)}$ (Å)	D_{max} (Å)	$\chi_{(dammif)}$	Normalized Spatial Discrepancy
11.4	115.1	36.1	35.7	115	1.7 ± 0.1	0.7 ± 0.1
7.74	115.6	33.9	35.4	110	2.1 ± 0.1	0.6 ± 0.1
2.95	126.3	37.7	37.7	120	1.4 ± 0.1	0.7 ± 0.1

[a] The expected molecular weight for SpuA is 130 kDa for this 1137 amino acid polypeptide. The molecular weight was determined from the SAXS data by the method of Fischer et al. (2010).

crystal structure of SpuA likely represents only one possible trapped form of the protein because the observed recognition and hydrolysis of larger substrates must require a degree of malleability in the spatial organization of the protein's composite modules to accommodate the substrate.

The presence of conformational flexibility in SpuA is supported by our analysis of the solution structure of SpuA, which reveals a conformation with the dyad of CBMs significantly retracted from the active site, thus allowing access to the substrate. A comparison of the solution structure of SpuA with the complexed crystal structure revealed that carbohydrate recognition by this protein requires a conformational trajectory that ranges from a somewhat open active site in the absence of substrate to the closed form observed in the crystal structure. This motion markedly resembles a hammer descending upon on an anvil. The CBM-dyad forms the head of the hammer and the linker domain the handle of the hammer, whereas the catalytic domain comprises the anvil. CBM41-2 makes up one side of the hammerhead, which remains distal from the active site and poised to interact with glycogen in a noncatalytic, adherent role. During the recognition of substrate, the hammer closes in on the anvil, allowing CBM41-1 to pin a substrate, which may be already bound in the active site. Alternatively, it is also plausible that CBM41-1 may play a role in first binding then delivering the substrate to the catalytic machinery as the hammer closes on the anvil. In either case, conformational flexibility, which appears to be a feature of the enzyme's multimodular nature, imparts the plasticity necessary for SpuA to recognize and hydrolyze a structurally complex substrate such as glycogen.

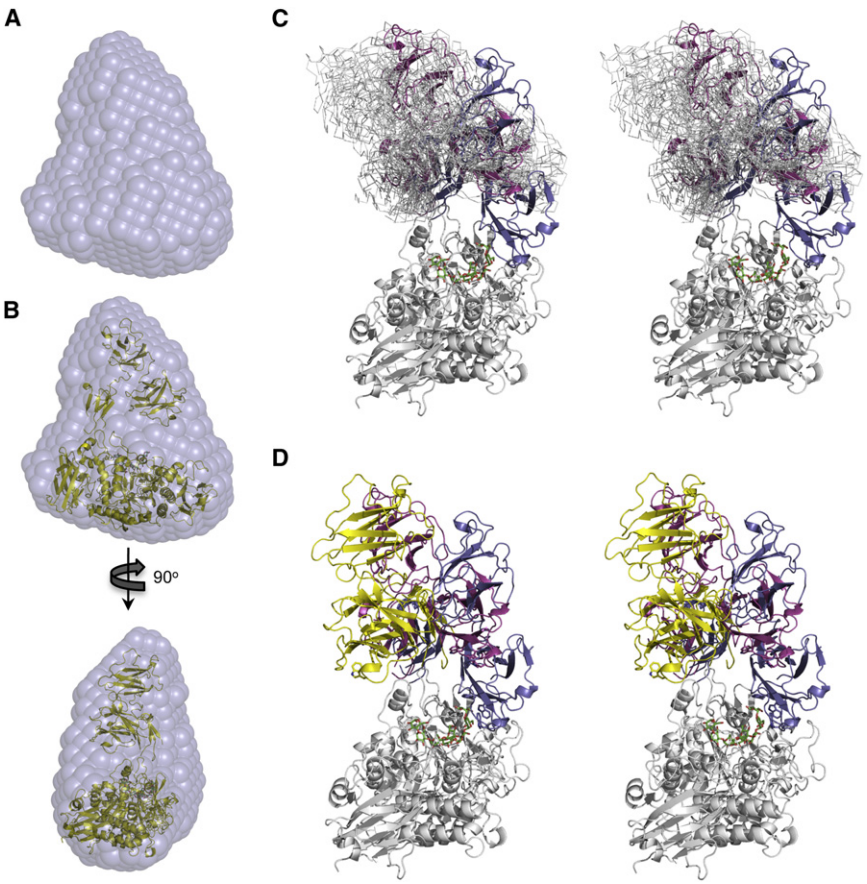

Figure 7. Solution Structure of SpuA Determined by SAXS Analysis

(A) Ab initio model of SpuA created from the average of 20 models generated by DAMMIF.

(B) Model 1 of SpuA generated by manually placing the N-terminal fragment of SpuA (CBMs and linker) and the C-terminal fragment (Ig-domains and $(\alpha/\beta)_8$-barrel) into the ab initio form.

(C) The ensemble of SpuA structures obtained from application of the EOM is shown in divergent stereo. The 20 models of the ensemble and the crystal structure were overlapped via the C-terminal catalytic domain (shown as a gray cartoon). The EOM models are shown as gray ribbons. The individual EOM-derived model giving the best individual χ_{crysol} of 1.8, referred to as model 2, is shown in a purple cartoon representation. The CBMs from the crystal structure are shown in blue cartoon representation.

(D) A comparison of the manually derived model 1 (yellow) with the best EOM model 2 (purple) and the X-ray crystal structure (blue). The models were overlapped via their catalytic modules, which are shown in gray. The residues on CBM41-1 that participate in forming the −3 and −4 subsites are shown as sticks.

In both (C) and (D), the sugars bound in the active site of SpuA in the crystal structure are shown in stick representation for reference.

The generally accepted paradigm of CBM function is one where CBMs are proposed to contribute to enzyme catalysis mainly by targeting the complete enzyme to a portion of the substrate that is near (on the order of the dimensions of the protein), but not immediately adjacent to the glycosidic bond being cleaved (Bolam et al., 1998; Boraston et al., 2004). This concept was established primarily on numerous observations that the truncation of CBMs from carbohydrate-active enzymes results in an enzyme with lowered activity on polysaccharides, most frequently insoluble polysaccharides. The best-known deviation from this paradigm is found among the glycoside hydrolase family 9 endoglucanases. Some of these enzymes have family 3c CBMs that are found at the immediate C terminus of the catalytic domain; the structures of these proteins reveal the CBM to be rigidly appended to the catalytic domain (Mandelman et al., 2003; Sakon et al., 1997). Consistent with the paradigm, deletion of the CBM reduces the activity of the enzyme, providing evidence of its importance in substrate recognition (Irwin et al., 1998). However, the unique feature of these cellulases is that the binding site of the CBM aligns with catalytic site, suggesting its role in extending the substrate-accommodating subsites of the active site (Mandelman et al., 2003; Sakon et al., 1997). SpuA also displays the expected paradigm "phenotype" of reduced activity when its CBMs are truncated. However, the structural analysis presented here indicates a substantial departure from the paradigm because the mode of substrate recognition by SpuA not only utilizes a CBM that is intimately associated with the active site, but substrate recognition also appears to be accompanied by conformational changes. These conformational changes involve a large portion of the enzyme's modular architecture and ultimately allow the CBM to retract from the active site and close in on the active site, likely to accommodate the somewhat heterogeneous glycogen substrate. These examples highlight the need to appreciate the overall three-dimensional structures of complete multimodular carbohydrate-active enzymes and the modular cooperation that this can reveal.

EXPERIMENTAL PROCEDURES

Construction of *S. pneumoniae* spuA Mutant

A defined, nonpolar deletion replacement mutant of *spuA* in *S. pneumoniae* D39 (serotype 2) was constructed by overlap-extension PCR essentially as described previously (Stroeher et al., 2007) (see also Supplemental Experimental Procedures).

Virulence Studies

D39 and D39Δ*spuA* were inoculated from an overnight blood agar plate into serum broth (SB) and grown to a culture density of approximately 1 × 10^8 cfu/ml. The bacteria were then diluted to the appropriate dose in sterile SB. Two groups of mice were injected i.n. with approximately 10^7 cfu total of a 1:1 mixture of D39wt:D39ΔspuA in 50 μl of SB. Before i.n. challenge, mice were anesthetized by intraperitoneal injection of Nembutal (pentobarbitone sodium; Rhone-Merieux) at a dose of 66 μg/g of body weight. The actual dose of bacteria administered was confirmed retrospectively by plating serial dilutions of the challenge inocula after administration to the mice. At 24 and 48 hr postinfection, bacteria were harvested from the nasopharynx, lungs, and blood of mice as described previously (Stroeher et al., 2007). Samples were serially diluted and plated on blood agar with or without erythromycin to enumerate wild-type and mutant pneumococci. Competitive indices between wild-type and mutant were analyzed using Student's one-sample t test (two-tailed).

Cloning, Recombinant Production, and Purification of Recombinant Proteins

A nearly full-length construct of SpuA, referred to as SpuA throughout, encoding a polypeptide lacking C-terminal LPXTG motif and the first 134 amino acids (omission of this N-terminal fragment resulted in a more stable polypeptide) was cloned, produced, and purified by immobilized metal affinity chromatography as described previously (Abbott et al., 2010). A more truncated DNA fragment encoding only the catalytic domain (amino acid residues 446–1245; SpuAΔCBM) and mutants of SpuA were generated and produced, purified, and quantified by standard molecular biology procedures (see Supplemental Experimental Procedures for details).

Immunofluorescence Microscopy

Cells (A549; ATCC CCL-185) were cultured in DMEM:F12, containing 10% FBS, 2 mM L-Glutamine, and 1 mM sodium pyruvate, and passaged into 8-well slide chambers. Cells were rinsed in PBS and fixed in 4% paraformaldehyde in PBS (15 min), rinsed, and blocked in 5% lamb serum in PBS Tween 20 (15 min, room temperature). Cells were treated with 2.5 mg/ml enzyme (SpuA or SpuAD778A), diluted in 20 mM Tris-HCl (pH 8.0), overnight at room temperature. Cells were rinsed three times in PBS before application of 1 μg/ml FITC-labeled CBM41 from *T. maritima*. FITC labeling of TmCBM41 was performed according to the manufacturer's guidelines (Sigma). Slides were mounted with coverslips and examined with an Olympus FluoView FV1000 laser-scanning confocal microscope. Single optical sections (1 μm) of fields of cells were collected, and fluorescence intensity per unit area was determined within a 5 × 5 μm sample area placed over the cytoplasm of cells chosen at random in monochrome images using ImageJ (version 1.43 m). Overall contrast and brightness were adjusted, and images were cropped and assembled using Photoshop (CS2). Statistical analysis (one-way ANOVA, Bonferroni's multiple comparison test) was done with GraphPad Prism (5.03).

Inhibitor Synthesis

The synthesis of maltopentaose-phenyl carbamate (PUGMAL) was performed as outlined in the Supplemental Experimental Procedures.

Enzyme Kinetics

All steady-state kinetic studies were carried out at 25°C in a Cary/Varian 300 Bio UV-Visible Spectrophotometer. Standard reaction mixtures for the determination of kinetic constants were done in PBS in 1 ml volumes containing 200 nM SpuA or SpuAΔCBM and 0–4 mM pNP-M5. The 4-nitrophenolate production was monitored at 400 nm; rates were calculated using an extinction coefficient, ε_{400nm}, of 7790 $cm^{-1}M^{-1}$. All experiments were performed in triplicate. Michaelis-Menten parameters were determined by nonlinear curve fitting using Origin7. The K_i value for SpuA or SpuAΔCBM inhibition by PUGMAL was determined using pNP-M5 as a substrate and by linear regression analysis of Dixon plots. The analysis of glycogen hydrolysis was performed as described in the Supplemental Experimental Procedures.

Binding Analysis

ITC was performed as described previously using a VP-ITC (MicroCal, Northampton, MA, USA) (van Bueren et al., 2007). Briefly, SpuA and SpuAΔCBM were extensively dialyzed against buffer (50 mM Tris HCl [pH 7.5]) then concentrated in a stirred ultrafiltration cell as above. PUGMAL solutions were prepared by mass in buffer saved from the ultrafiltration step. Both protein and sugar solutions were filtered and degassed immediately prior to use. The protein concentrations used were ~50–200 μM. Titrations were performed at 25°C. The SpuAΔCBM data were fit with a single binding site model to determine the dissociation constant. The SpuA data were more complicated and could not be adequately fit with a single binding site model. It was reasoned, on the basis of the structure of the protein, that this was due to binding in the active site as well as to the distal CBM41-2 module. We had previously determined the thermodynamics of maltotetraose recognition by CBM41-2 (van Bueren et al., 2007); thus, we fit the SpuA data to a two-site binding model, setting one class of binding site as a constant with parameters equal to those determined for CBM41-2. Due to limiting amounts of reagents, only single titrations could be performed; reported errors are those determined by model fitting.

Table 2. X-ray Data Collection and Structure Refinement Statistics

	SpuAΔCBM	SpuAΔCBM with G1M	SpuA with Maltotetraose
Data Collection			
Space group	$P2_1$	$P2_1$	$P2_12_12_1$
Cell dimensions			
a, b, c (Å)	57.57, 75.20, 87.04 (β = 97.31°)	57.78, 75.80, 88.09 (β = 97.72°)	80.14, 86.43, 193.22
Resolution (Å)	20.00–1.85 (1.92–1.85)[a]	20.00–2.37 (2.48–2.37)	20.00–2.25 (2.33–2.25)
R_{merge}	0.057 (0.403)	0.081 (0.250)	0.086 (0.396)
$<I/\sigma I>$	10.2 (2.3)	5.8 (2.6)	9.4 (2.8)
Completeness (%)	96.7 (94.1)	90.0 (91.1)	92.6 (94.8)
Redundancy	3.73 (3.64)	2.23 (2.14)	4.25 (4.18)
Refinement			
Number of reflections	57,697	25,638	56,573
R_{work}/R_{free}	0.17/0.23	0.21/0.26	0.21/0.27
Number of atoms			
Protein	5,708	5,636	7,880
Ligand/ion	1 (Na)/24 (GOL)/1 (Ca)	22 (G1M)/1 (Ca)/1 (Na)	124 (GLC)/4 (Na)/5 (SO_4)
Water	839	206	715
B factors			
Protein	26.2	26.5	34.6
Ligand/ion	15.4 (Na)/34.0 (GOL)/27.5 (Ca)	18.2 (G1M)/23.4 (Ca)/13.0 (Na)	34.2 (GLC)/40.8 (Na)/33.7 (SO_4)
Water	38.1	24.6	38.5
Rms deviation			
Bond lengths (Å)	0.024	0.014	0.011
Bond angles (°)	1.927	1.440	1.348
Ramachandran (%)			
Favored	97.1	96.3	95.4
Allowed	2.8	3.5	4.6
Outliers	0.1	0.1	0.0

[a] Numbers in parenthesis indicate values for the highest resolution shell.

Crystallization

Crystallization was performed using the vapor diffusion hanging drop method. Crystals of SpuAΔCBM (5 mg/ml) were obtained in 17.5% PEG 3350, 0.2 M $MgCl_2$, 0.1 M sodium acetate (pH 5.5), and 3% glycerol. SpuAΔCBM was complexed with 4-O-α-D-glucopyranosylmoranoline (Toronto Research Chemicals) by soaking a crystal in crystallization solution containing 2 mM of the compound. Crystals of SpuA (15 mg/ml) that was precomplexed with maltotetraose (at 5 mM) grew in 2 M $(NH_4)_2SO_4$, 0.2 M Na/K Tartrate, 0.1 M tri-sodium citrate (pH 5.6), supplemented with 3% glycerol. Crystals were flash cooled in liquid nitrogen using a cryoprotectant comprising crystallization solution supplemented with 20% glycerol for SpuAΔCBM and 15% glycerol for SpuA.

X-ray diffraction data sets were collected on a Rigaku R-AXIS 4++ area detector coupled to a MM-002 X-ray generator with Osmic "blue" optics and an Oxford Cryostream 700. Data were processed with CrystalClear/d*trek (Pflugrath, 1999). Data collection and processing statistics are given in Table 2. The structure of SpuAΔCBM was solved by molecular replacement using MOLREP (Vagin and Teplyakov, 2000) with the *Klebsiella pneumoniae* pullulanase catalytic module coordinates as a search model (PDB code 2FHF) (Mikami et al., 2006). The initial model was corrected and completed manually by successive rounds of building using Coot (Emsley and Cowtan, 2004) and refinement with REFMAC (Murshudov et al., 1997). SpuA was solved by molecular replacement with MOLREP using the coordinates of SpuAΔCBM and the CBM dyad (PDB code 2J44) as search models. The initial model of SpuA was manually completed, corrected, and refined as above.

In all cases, water molecules were added using COOT:FINDWATERS and inspected visually prior to deposition. Of the observations 5% were flagged as "free" and used to monitor refinement procedures (Brunger, 1992). Model validation was performed with SFCHECK (Vaguine et al., 1999) and PROCHECK (Laskowski et al., 1993). All data collection and model statistics are given in Table 2.

SAXS

Synchrotron X-ray scattering data from solutions of SpuA from 2.95 to 11.4 mg/ml were collected at the X33 beamline of the EMBL (DESY, Hamburg) using a MAR345 image plate detector. A 4.2 mg/ml solution of BSA was measured as a reference and for calibration. The scattering patterns were measured with an exposure time of 2 min at 288 K. The wavelength was 1.5 Å. The sample-to-detector distance was set at 2.4 m, leading to scattering vectors Q ranging from 0.06 to 0.5 $Å^{-1}$. Data processing and determination of Rg, and Dmax were performed as described previously (Ficko-Blean et al., 2009).

The ab initio low-resolution envelopes of SpuA were generated with DAMMIF (Franke and Svergun, 2009) using 20 independent runs with no shape constraints. The ab initio reconstructions were aligned, averaged, and filtered using the DAMAVER (Volkov and Svergun, 2003) suite of programs. SpuA was considered as a two-domain protein, with the first domain the two CBMs and the linker region (amino acids 130–436), and the second domain the catalytic module (amino acids 437–1143); these two domains were manually fit into the filtered envelope. The goodness of fit for all atomic models with the experimental data was determined using the discrepancy χ, defined according to Konarev et al. (2003). The solution scattering for the models, the fit to the experimental curve, and the goodness of fit was evaluated using CRYSOL (Svergun et al., 1995).

EOM (Bernado et al., 2007) was applied to the SAXS data as follows. The polypeptide region (~aa 435–438) between the linker domain and the N-domain of the catalytic module was treated as a "hinge," and a random pool of 10,000 conformers was generated around the flexible hinge using RANCH from which the genetic algorithm GAJOE selected the ensemble fitting the experimental data. EOM modeling of native SpuA was done using data collected on the 2.95 mg/ml concentration of SpuA. Both data sets were analyzed over a Q range of ~0.06–0.5 $Å^{-1}$. The number of ensembles to compare to the experimental data was set at 50, and the number of generations the genetic algorithm used to optimize the ensemble was 1000, with 20 crossings per generation and 50 repetitions. χ values were determined for the ensembles during EOM modeling and separately for the individual ensemble models using the program CRYSOL (Svergun et al., 1995).

ACCESSION NUMBERS

Coordinates and structure factors have been deposited with the PDB codes of 2ya1 for SpuA with maltotetraose, 2ya0 for SpuAΔCBM, and 2ya2 for SpuAΔCBM in complex with G1M.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at doi:10.1016/j.str.2011.03.001.

ACKNOWLEDGMENTS

We acknowledge Alexey Kikhney and all other staff members for help on the SAXS beamline X33 at the DORIS-III storage ring at DESY (Deutsches Elektronen-Synchrotron). This work was supported by grants from the British Columbia Lung Association (to A.B.B.), the Canada Institutes of Health Research (to A.B.B. and D.J.V.), the Natural Sciences and Engineering Research Council of Canada (NSERC; to R.D.B.), and the National Health and Medical Research Council of Australia (NHMRC; to J.C.P.). A.LvB. and M.A.H. are supported by doctoral fellowships from NSERC and the Michael Smith Foundation for Health Research (MSFHR). B.P. is supported by a MSFHR Post-Doctoral Fellowship. J.-H.H. was supported by a Marie-Curie doctoral Fellowship. J.C.P. is a NHMRC Australia Fellow. D.J.V. is a Canada Research Chair in Chemical Glycobiology and a MSFHR Career Scholar. A.B.B. is a Canada Research Chair in Molecular Interactions and a MSFHR Career Scholar.

Received: January 5, 2011
Revised: February 22, 2011
Accepted: March 3, 2011
Published: May 10, 2011

REFERENCES

Abbott, D.W., Higgins, M.A., Hyrnuik, S., Pluvinage, B., Lammerts van Bueren, A., and Boraston, A.B. (2010). The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae*. Mol. Microbiol. *77*, 183–199.

Abe, A., Tonozuka, T., Sakano, Y., and Kamitori, S. (2004). Complex structures of *Thermoactinomyces vulgaris* R-47 alpha-amylase 1 with malto-oligosaccharides demonstrate the role of domain N acting as a starch-binding domain. J. Mol. Biol. *335*, 811–822.

Bayer, E.A., Chanzy, H., Lamed, R., and Shoham, Y. (1998). Cellulose, cellulases and cellulosomes. Curr. Opin. Struct. Biol. *8*, 548–557.

Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. J. Am. Chem. Soc. *129*, 5656–5664.

Bolam, D.N., Ciruela, A., Mcqueen-Mason, S., Simpson, P., Williamson, M.P., Rixon, J.E., Boraston, A., Hazlewood, G.P., and Gilbert, H.J. (1998). Pseudomonas cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. Biochem. J. *331*, 775–781.

Bongaerts, R.J., Heinz, H.P., Hadding, U., and Zysk, G. (2000). Antigenicity, expression, and molecular characterization of surface-located pullulanase of *Streptococcus pneumoniae*. Infect. Immun. *68*, 7141–7143.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004). Carbohydrate-binding modules: fine tuning polysaccharide recognition. Biochem. J. *382*, 769–782.

Brunger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–475.

Davies, G.J., Wilson, K.S., and Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochem. J. *321*, 557–559.

Eckhart, L.J. (1982). Kinetic measurement of alpha-amylase by a P-nitrophenyl maltopentaoside method using the American Monitor Kda. Clin. Chem. *28*, 1603–1603.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Ficko-Blean, E., Gregg, K.J., Adams, J.J., Hehemann, J.H., Czjzek, M., Smith, S.P., and Boraston, A.B. (2009). Portrait of an enzyme, a complete structural analysis of a multimodular β-N-acetylglucosaminidase from *Clostridium perfringens*. J. Biol. Chem. *284*, 9876–9884.

Fischer, H., Neto, M.D., Napolitano, H.B., Polikarpov, I., and Craievich, A.F. (2010). Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. J. Appl. Crystallogr. *43*, 101–109.

Franke, D., and Svergun, D.I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Crystallogr. *42*, 342–346.

Gloster, T.M., and Davies, G.J. (2010). Glycosidase inhibition: assessing mimicry of the transition state. Org. Biomol. Chem. *8*, 305–320.

Gourlay, L.J., Santi, I., Pezzicoli, A., Grandi, G., Soriani, M., and Bolognesi, M. (2009). Group B streptococcus pullulanase crystal structures in the context of a novel strategy for vaccine development. J. Bacteriol. *191*, 3544–3552.

Hammel, M., Fierobe, H.P., Czjzek, M., Finet, S., and Receveur-Brechot, V. (2004). Structural insights into the mechanism of formation of cellulosomes probed by small angle X-ray scattering. J. Biol. Chem. *279*, 55985–55994.

Hammerschmidt, S., Wolff, S., Hocke, A., Rosseau, S., Muller, E., and Rohde, M. (2005). Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. Infect. Immun. *73*, 4653–4667.

Hava, D.L., and Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. *45*, 1389–1406.

Higgins, M.A., Whitworth, G.E., El Warry, N., Randriantsoa, M., Samain, E., Burke, R.D., Vocadlo, D.J., and Boraston, A.B. (2009). Differential recognition and hydrolysis of host carbohydrate antigens by *Streptococcus pneumoniae* family 98 glycoside hydrolases. J. Biol. Chem. *284*, 26161–26173.

Hyams, C., Camberlein, E., Cohen, J.M., Bax, K., and Brown, J.S. (2010). The *Streptococcus pneumoniae* capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. Infect. Immun. *78*, 704–715.

Irwin, D., Shin, D.H., Zhang, S., Barr, B.K., Sakon, J., Karplus, P.A., and Wilson, D.B. (1998). Roles of the catalytic domain and two cellulose binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. J. Bacteriol. *180*, 1709–1714.

King, S.J. (2010). Pneumococcal modification of host sugars: a major contributor to colonization of the human airway? Mol. Oral Microbiol. *25*, 15–24.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. *36*, 1277–1282.

Koropatkin, N.M., and Smith, T.J. (2010). SusG: a unique cell-membrane-associated alpha-amylase from a prominent human gut symbiont targets complex starch molecules. Structure *18*, 200–215.

Lammerts van Bueren, A., Finn, R., Ausio, J., and Boraston, A.B. (2004). Alpha-glucan recognition by a new family of carbohydrate-binding modules found primarily in bacterial pathogens. Biochemistry *43*, 15633–15642.

Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

MacGregor, E.A., Janecek, S., and Svensson, B. (2001). Relationship of sequence and structure to specificity in the alpha-amylase family of enzymes. Biochim. Biophys. Acta *1546*, 1–20.

Mandelman, D., Belaich, A., Belaich, J.P., Aghajari, N., Driguez, H., and Haser, R. (2003). X-ray crystal structure of the multidomain endoglucanase Cel9G from *Clostridium cellulolyticum* complexed with natural and synthetic cello-oligosaccharides. J. Bacteriol. *185*, 4127–4135.

Maruo, S., Kyotani, Y., Yamamoto, H., Miyazaki, K., Ogawa, H., Sakai, T., Kojima, M., and Ezure, Y. (1993). Effects of moranoline, 4-O-alpha-D-glucopyranosylmoranoline and their N-substituted derivatives on thermostability of cyclodextrin glycosyltransferase, glucoamylase, and beta-amylase. Biosci. Biotechnol. Biochem. *57*, 1294–1298.

Mikami, B., Iwamoto, H., Malle, D., Yoon, H.J., Demirkan-Sarikaya, E., Mezaki, Y., and Katsuya, Y. (2006). Crystal structure of pullulanase: evidence for parallel binding of oligosaccharides in the active site. J. Mol. Biol. *359*, 690–707.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Nishimasu, H., Fushinobu, S., Shoun, H., and Wakagi, T. (2007). Crystal structures of an ATP-dependent hexokinase with broad substrate specificity from the hyperthermophilic archaeon *Sulfolobus tokodaii*. J. Biol. Chem. *282*, 9923–9931.

Obert, C., Sublett, J., Kaushal, D., Hinojosa, E., Barton, T., Tuomanen, E.I., and Orihuela, C.J. (2006). Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. Infect. Immun. *74*, 4766–4777.

Pflugrath, J.W. (1999). The finer things in X-ray diffraction data collection. Acta Crystallogr. D Biol. Crystallogr. *55*, 1718–1725.

Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q. Rev. Biophys. *40*, 191–285.

Quiocho, F.A., Spurlino, J.C., and Rodseth, L.E. (1997). Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. Structure *5*, 997–1015.

Sakon, J., Irwin, D., Wilson, D.B., and Karplus, P.A. (1997). Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. Nat. Struct. Biol. *4*, 810–818.

Santi, I., Pezzicoli, A., Bosello, M., Berti, F., Mariani, M., Telford, J.L., Grandi, G., and Soriani, M. (2008). Functional characterization of a newly identified group B Streptococcus pullulanase eliciting antibodies able to prevent alpha-glucans degradation. PLoS ONE *3*, e3787.

Stam, M.R., Danchin, E.G., Rancurel, C., Coutinho, P.M., and Henrissat, B. (2006). Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. Protein Eng. Des. Sel. *19*, 555–562.

Stroeher, U.H., Kidd, S.P., Stafford, S.L., Jennings, M.P., Paton, J.C., and McEwan, A.G. (2007). A pneumococcal MerR-like regulator and S-nitrosoglutathione reductase are required for systemic virulence. J. Infect. Dis. *196*, 1820–1826.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRYSOL—a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Crystallogr. *28*, 768–773.

Turkenburg, J.P., Brzozowski, A.M., Svendsen, A., Borchert, T.V., Davies, G.J., and Wilson, K.S. (2009). Structure of a pullulanase from *Bacillus acidopullulyticus*. Proteins *76*, 516–519.

Vagin, A., and Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. Acta Crystallogr. D Biol. Crystallogr. *56*, 1622–1624.

Vaguine, A.A., Richelle, J., and Wodak, S.J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallogr. D Biol. Crystallogr. *55*, 191–205.

van Bueren, A.L., Higgins, M., Wang, D., Burke, R.D., and Boraston, A.B. (2007). Identification and structural basis of binding to host lung glycogen by streptococcal virulence factors. Nat. Struct. Mol. Biol. *14*, 76–84.

Volkov, V.V., and Svergun, D.I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. J. Appl. Crystallogr. *36*, 860–864.

Zechel, D.L., and Withers, S.G. (2001). Dissection of nucleophilic and acid-base catalysis in glycosidases. Curr. Opin. Chem. Biol. *5*, 643–649.