The Structural Basis for T-antigen Hydrolysis by Streptococcus pneumoniae

A TARGET FOR STRUCTURE-BASED VACCINE DESIGN

Matthew E. C. Caines¹,‡, Haizhong Zhu¹, Marija Vuckovic¹,¶, Lisa M. Willis⁵, Stephen G. Withers¹⁰, Warren W. Wackarchuk¹, and Natalie C. J. Strynadka¹⁹¹

From the Department of Biochemistry and Molecular Biology, Centre for Blood Research and Department of Chemistry, University of British Columbia, Vancouver, British Columbia, V6T 1Z3 and the National Research Council of Canada, Institute for Biological Sciences, Ottawa, Ontario, K1A 0R6, Canada

Streptococcus pneumoniae endo-α-N-acetylgalactosaminidase is a cell surface-anchored glycoside hydrolase from family GH101 involved in the breakdown of mucin type O-linked glycans. The 189-kDa mature enzyme specifically hydrolyzes the T-antigen disaccharide from extracellular host glycoproteins and is representative of a broadly important class of virulence factors that have remained structurally uncharacterized due to their large size and highly modular nature. Here we report a 2.9 Å resolution crystal structure that remarkably captures the multidomain architecture and characterizes a catalytic center unexpectedly resembling that of α-amylases. Our analysis presents a complete model of glycoprotein recognition and provides a basis for the structure-based design of novel Streptococcus vaccines and therapeutics.

Streptococcus pneumoniae is a Gram-positive human commensal that typically colonizes the respiratory epithelium. Carriage of the bacterium, although potentially asymptomatic, often progresses to a pathogenic state causing otitis media, pneumonia, septicaemia, and/or meningitis. Worldwide, S. pneumoniae is a major cause of morbidity and mortality; pneumococcal infections have been estimated to result in over one million pneumonia-related deaths per year (1). Polysaccharide-based vaccines are only relatively effective; poor efficacy in young and immunodeficient patients, the large number (>90) of S. pneumoniae serotypes, and the increased prevalence of non-vaccine serotypes complicates vaccination strategies. Furthermore, the emergence of increased antibiotic resistance in S. pneumoniae raises the specter of an increasing public health threat (2, 3).

Streptococcus sp. produce numerous glycoside hydrolases that act in concert to degrade host glycans. The ability to degrade extracellular glycoproteins may aid colonization in at least two ways. Firstly, the destruction of the mucin barrier may aid persistence, and secondly, the utilization of liberated glycans as a fermentable carbon source may improve the growth conditions for this pathogen (4). Endo-α-N-acetylgalactosaminidase, a glycoside hydrolase from family GH101 (see the Carbohydrate-Active Enzymes (CAZY) data base), is one such virulence factor. This enzyme, SpGH101,2 hydrolyzes the α-linkage between Galβ(1,3)GalNAc moieties and serine or threonine residues (the T-antigen or Core 1-Type O-glycan) of host glycoproteins (5). SpGH101 is exported to the extracellular environment by virtue of a YSIRK-G/S motif containing secretion signal. Although the enzyme has been observed as a soluble protein in the culture medium after growth in vitro (6), it has also been shown to be localized at the S. pneumoniae cell surface by cell surface staining and opsonophagocytic killing assays (7). The cell surface localization is presumably a consequence of the C-terminal LPXTG sortase signal. This motif, followed by a predicted transmembrane domain and a charged C terminus, is indicative of covalent anchoring of the enzyme to the peptidoglycan cell wall (8). It is possible that this covalent anchor is unstable under the culture conditions; similar behavior has been observed with both S. pneumoniae neuraminidase A and Group B Streptococcus immunogenic bacterial adhesin A (9, 10).

Extensive characterization of a related family GH101 enzyme from Bifidobacterium longum (40% sequence identity overall, 65% sequence identity over the catalytic domain) has revealed that hydrolysis proceeds with retention of configuration (5), suggestive of a double displacement mechanism (11). Further, the identity of catalytically important amino acids was inferred by mutagenesis of conserved aspartic and glutamic acid residues; mutation of Asp-682, Asp-789, and Glu-822 (corresponding to Asp-658, Asp-764, and Glu-826 in SpGH101) induced a significant reduction in enzymatic activity. Here we report the x-ray crystal structure of a ~170-kDa fragment of endo-α-N-acetylgalactosaminidase from S. pneumoniae strain R6 and discuss the structural, mechanistic, and therapeutic implications arising.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—The spgh101 gene was PCR-amplified from S. pneumoniae strain R6 (ATCC BAA-255) genomic DNA and subcloned as an Ndel/Sall fragment into the pCWori+ vector (12) to generate an N-terminal His₆-tagged fusion protein, corresponding to residues 40–1567. Primers used were: forward, 5’-GCTGGAGCTG-
ACATATGCACACCACACCACATCACCATTTTTAACCACCCGC-GTGGATCCTAGGCCTGACCGTCTGGTCCACGGCGAAA-TCTAAC-3’; reverse, 5’-GCTGGAAGCTGGAGCTGACCTTAT- TATGGCACAAAGTCCCACATCAGCGAATTACACTTAC-ATTGGAC-3’. Following amplification, the integrity of spgh101 was confirmed by DNA sequencing.

Selenomethionine (SeMet)-substituted SpGH101-(40–1567) was expressed using a metabolic inhibition protocol. *Escherichia coli* BL21(DE3) cells containing the spgh101/pCWori + construct were grown in M9 medium supplemented with 50 mg dm⁻³ l-selenomethionine and containing 100 µg ml⁻¹ ampicillin.

SpGH101-(40–1567) was purified by sequential use of Ni²⁺ affinity, anion exchange, hydrophobic interaction, and size exclusion chromatography to >95% purity, as determined by SDS-PAGE analysis. The His₆ tag was removed by cleavage of the fusion protein with thrombin prior to the anion exchange step. The activity of SpGH101-(40–1567) was confirmed by using a UV spectrophotometric assay to monitor the release of 2,4-dinitrophenol from Galβ(1,3)GalNacαDPN.

**Crystallization**—Crystals of SeMet-substituted SpGH101-(40–1567) were grown by using the sitting drop vapor diffusion method. Drops consisted of 1 µl of protein (10 mg ml⁻¹) and 1 µl of reservoir solution (29% w/v) polyethylene glycol monomethyl ether 2000, 0.2 M lithium citrate, and 0.1 M ammonium acetate) and were equilibrated at 21 °C.

**Data Collection and Structure Determination**—Cryoprotection was accomplished by sequential transfer into solutions of 30% (w/v) polyethylene glycol 3350, 0.2 M lithium citrate, and 0.1 M ammonium acetate, pH 6.5, containing 5% (v/v) glycerol and 10% (v/v) glycerol. Crystals were cryocooled by plunging into liquid nitrogen and x-ray data were collected at 100 K using a nitrogen stream.

A single anomalous dispersion data set was collected at the peak absorption wavelength (λ peak) at Beamline 8.2.2 of the Advanced Light Source (Berkeley, CA) using an ADSC Q315 CCD detector. The data were processed using MOSFLM (13) and the CCP4 suite of programs (14). The crystals belonged to the space group P2₁2₁2₁. The crystallographic asymmetric unit contains two SpGH101-(40–1567) molecules (labeled A and B).

Initial selenium site locations were located using SHELXD (15). Additional selenium site location, refinement, and phasing was performed using autoSHARP (16). Density modification was carried out with Solomon (17).

**Model Building and Refinement**—An initial model was manually built using Coot (18). Refinement was commenced using PHENIX (19), implementing simulated annealing. Further refinement was carried out using REFMAC5 (20), implementing positional minimization, TLS, and individual atomic displacement parameter refinement. Iterative manual model rebuilding was performed using Coot. Refinement was carried out using the maximum likelihood with Hendrickson-Lattman coefficients target function. Non-crystallographic symmetry restraints were used throughout refinement. 5% of the reflections were excluded from refinement for the calculation of R free. Data collection, phasing, and refinement statistics are shown in Table 1.

The final model contains residues 119–307, 319–1439, and 1451–1476 (chain A) and 122–222, 226–307, and 318–1481 (chain B). The electron density for domains 1–6 is, in general, of better quality in chain A. Notably, the electron density for domain 7 is of better quality in chain B, although this domain exhibits electron density of lower quality than the remainder of the structure. With the exception of domain 7, our analysis is based on the structure of chain A. Four features in the electron density were modeled as Ca²⁺ ions based upon the structure of chain A. Four features in the electron density were modeled as Ca²⁺ ions based upon coordination geometry and distance (21). Additional features were modeled as Na⁺ ions, based upon the same criteria, and glycerol molecules, although we note the electron density in these regions is not definitive, and these assignments should be regarded as provisional. Both chain A and chain B exhibited an unidentified peak of electron density at the periphery of the protein, distant from the active site, near Ser-1127, that was unmodeled. We also note that short sections of electron density in the 308–318 linker region and areas of domain 7, although visible, were ambiguous and therefore were not modeled.

Model validation was performed using MolProbity (22). Ramachandran analysis indicates the final model contains 96.2, 3.7, and 0.1% of residues in favored, allowed, and disallowed regions, respectively. The figures were prepared using PyMOL (23), and electrostatic potentials were calculated using APBS (24).

**Substrate Modeling**—An energy-minimized model of Galβ(1,3)GalNacαThr and its refinement restraints were calculated using the PRODRG server (25). This model was manually positioned in the active site of SpGH101-(40–1567) using the superposed *Thermoactinomyces vulgaris* R-47 α-amylase in complex with a substrate (2DH (26)) as a template. The Galβ(1,3)GalNacαThr coordinates were then subjected to energy minimization using CNS (27).
RESULTS AND DISCUSSION

The structure of SpGH101-(40–1567) can be traced between residues 119 and 1481 with a break in a linker region from residues 308 to 317. The overall architecture of SpGH101-(40–1567) (Fig. 1 and supplemental movie 1) comprises seven distinct domains. The N-terminal domain, residues 124–306, is a twisted β-sandwich composed of two sheets of six and seven antiparallel β-strands. The domain is structurally similar to the legume lectin-type fold (Dali Z-scores: 2RHP (29), 13.3; 2EXJ (30), 11.9; 1VIW (31), 11.7; 2EIG (32), 11.5), although notably, the extended metal and carbohydrate-binding loops appear to be missing. The absence of these loop regions exposes a deeper channel similar to the active site of the alginate lyases of polysaccharide lyase family PL7 (Dali Z-scores: 2ZAA, 11.1; 1VAV (33), 10.9). At the base of the cleft, four phenylalanine residues are exposed to solvent and may play a role in carbohydrate recognition. In the crystallized structure, domain 1 is positioned directly over the active site of the glycoside hydrolase domain (domain 3). The extended linker (residues 307–329) between domains 1 and 2 and the poor ordering of domain 1 in chain B implies that this orientation is most likely a crystallographic artifact. Based upon these observations, it is possible that domain 1 may function as a non-catalytic carbohydrate-binding module. Via the linker, this may anchor the main body of the enzyme in the mucin layer of the respiratory epithelium in a manner analogous to that proposed for the structurally similar legume lectin-type domains of the Vibrio cholerae neuraminidase (34).

Domain 2, residues 330–597, comprises a supersandwich fold of two antiparallel sheets. This exhibits the strongest structural similarity to the galactose mutarotase-like fold; however, comparison with a galactose mutarotase (Dali Z-score: 1SO0 (35), 12.7) reveals that the loops constituting the active site are absent and that the active site residues are not conserved. In place of the defined galactose mutarotase active site pocket, domain 2 exhibits a channel along the concave surface of the supersandwich fold similar to the proposed substrate-binding cleft in the periplasmic domain of YidC (Dali Z-score: 3B56 (36), 13.0). Although the function of this domain remains unclear in SpGH101, structurally similar domains are found in a number of glycoside hydrolases where they may confer stability to the multidomain architectures and/or be a relic of multifunctional ancestry.

The catalytic glycoside hydrolase domain (domain 3), residues 602–893, displays a distorted (β/α)β-barrel fold with three helices and strands replaced by a complex arrangement of loops. Overall, the domain displays the strongest structural similarity to the catalytic domains of the glycoside hydrolases of family GH13 (Dali Z-scores: 1GV1 (37), 15.3; 2D0H (26), 15.2; 1V9 (38), 15.1). Notably, the catalytic domain is most similar to the α-amylases, enzymes that also catalyze the hydrolysis of α-linked carbohydrates via a double displacement mechanism. The catalytic domain of SpGH101 possesses a subdomain, residues 699–733, in common with many α-amylases. This decora
dration, apparently lacking the typical structural Ca2+-binding site, defines one extremity of the active site cleft. Significant further insight can be gained by superposing the catalytic domain with the structure of the α-amylase from T. vulgaris R-47 in complex with a substrate (2D0H, supplemental Fig. S1a) (26). The α-amylase nucleophile and general acid/base catalytic residues, Asp-356 and Glu-396, respectively, superpose directly with residues Asp-764 and Glu-796 of SpGH101. The third residue implicated in SpGH101 catalysis, Asp-658, superposes with Tyr-223 of the α-amylase, a residue proposed to facilitate hydrophobic interactions with the carbohydrate bound in the −1 subsite. The active site (Fig. 2) appears negatively charged and possesses numerous polar residues capable of interacting with the substrate carbohydrate hydroxyl groups. The size of the well defined active site pocket in the enzyme surface, its extremities delimited on one side by Trp-724 and Trp-726, is consistent with the observed disaccharide substrate.
specificity. The opposite side of the cleft is bordered by a loop (Trp-867 and Gln-868) that appears less ordered, the flexibility of which may aid substrate binding.

Using the superposed α-amylase-substrate complex as a template, we have modeled the Galβ(1,3)GalNACα-Thr substructure in the active site of SpGH101 (Fig. 2 and supplemental Fig. S1). The postulated nucleophile and general acid/base residues, Asp-764 and Glu-796 respectively, are positioned well with respect to their catalytic roles. This model is consistent with the observed substrate specificity of the B. longum endo-α-N-acetylglactosaminidase (5), and the residues surrounding the active site show a high level of sequence identity (supplemental Fig. S2). Asp-658 is well placed to interact with the GalNAC 4-hydroxyl in this model, a binding contribution that may explain the significant loss of activity upon mutation of this residue and the observed intolerance of GlcNAc in the −1 subsite. Although our model provides insight into the carbohydrate-binding site, the seemingly fortuitous positioning of domain 1 provides a model of potential interaction with the substrate protein moiety.

The domain neighboring the catalytic domain (domain 4), residues 894–1043, constitutes an extensive antiparallel β-sheet structure that effectively shields one side of the catalytic domain from solvent. Analogous, albeit structurally dissimilar, β-sheet domains are also observed in comparable locations with respect to the catalytic domain in α-amylase structures. These domains have been proposed to stabilize the catalytic domain, in particular the secondary structural elements influencing the active site. It seems feasible that domain 4 of SpGH101 could serve a similar purpose.

Domains 5 and 6, residues 1058–1213 and 1214–1417, respectively, possess comparable β-sandwich jelly roll folds comprised of two sheets of five antiparallel β-strands. Although the core jelly roll fold of both domains is similar and both domains present a concave solvent exposed surface, domain 6 possesses more extensive loop regions interconnecting the β-strands. Domain 5 exhibits structural similarity to the carbohydrate-binding modules (for review, see Ref. 39) of families CBM16 (Dali Z-score: 2ZEY (40), 14.3) and CBM4 (Dali Z-scores: 1GUI (41), 11.5; 1K42 (42), 10.5). The shallow cleft of SpGH101 domain 5 most closely resembles that of the CBM16 protein. The extended loops that form the more prominent substrate-binding clefts of the CBM4 proteins are absent. The putative binding site in domain 5 is lined predominantly by hydrophilic residues and is devoid of the aromatic side chains that are commonly associated with carbohydrate recognition. Domain 5 also displays structural similarity to the carbohydrate-binding domain of the F-box protein of the E3 ubiquitin ligase complex (Fbs1, Dali Z-score: 2E32 (43), 11.1). The Fbs1 carbohydrate-binding domain displays a significantly different mode of substrate binding to the proteins of families CBM16 and CBM4; solvent-exposed aromatic residues in loop regions distant to the CBM16 and CBM4 clefts provide a shelf for carbohydrate recognition. This mode of substrate binding appears unlikely in domain 5 of SpGH101; however, as such aromatic residues are absent. Domain 6 is structurally similar to the carbohydrate-binding modules of families CBM22 and CBM16 (Dali Z-scores: 1H6X (44), 10.2; 1DYO (45), 10.0; 2ZEY (40), 9.6). It is, however, unclear as to whether the extended loops in domain 6 could form a similar substrate-binding site, as has been found with these structurally similar carbohydrate-binding modules. Indeed, the formation of a defined surface cleft would require flexibility of at least one loop, and such a movement would expose a number of hydrophobic residues to solvent. Notably, the calcium-binding site observed in these CBMs is conserved in SpGH101 domain 6. Domains 5 and 6 may therefore act as carbohydrate recognition domains, but in the absence of binding data, we note that it is also possible that these β-sandwich domains have no role in carbohydrate binding as has been previously observed (46). It is possible that these domains serve entirely structural functions to stabilize the overall protein architecture.

The final domain in the structure, residues 1421–1481, is a helical bundle comprising three α-helices that forms an interface between symmetry-related molecules. In addition to the seven domains described by our structural studies, a further domain (residues 1497–1622) is predicted between domain 7 and the cell wall anchor that, from primary sequence analysis, is predicted to be a member of carbohydrate-binding module family CBM32. Overall therefore, SpGH101 possesses four potential carbohydrate-binding modules. These domains may serve to either position the enzyme on the cell wall or within the Streptococcus polysaccharide capsule or enable multivalent binding of the enzyme, and thus bacterium, to the epithelial mucin layer.

The vaccines currently clinically available against S. pneumoniae offer only partial protection against pneumococcal diseases across both developed and developing countries. The 23-valent capsular polysaccharide-based vaccine is limited by its low efficacy in children under 2 years of age and in immunodeficient patients, whereas the 7-valent conjugate vaccine, although showing greater efficacy in young children, is limited by its narrow serotype coverage (47). Vaccines based on non-polysaccharide antigens present promising alternatives for the development of future vaccine candidates. To identify novel vaccine antigens, a recent approach analyzed a peptide fragment library describing the proteome of S. pneumoniae strain TIGR4 (7). The library was screened using serum antibodies from both exposed, but not infected, individuals and patients recovering from pneumococcal disease. Subsequent in vitro validation and animal testing confirmed SP0368, a homologue of SpGH101 (98% sequence identity), as providing “significant cross-protection in a stringent mouse model of lethal sepsis” (7). Although SP0368 was discounted in the final stage of trials due to sequence diversity, it is possible to utilize the structure of SpGH101 presented here to design improved epitopes that specifically target structurally conserved and surface-exposed regions of the enzyme (3 of the 5 immunogenic regions are mapped onto the structure of SpGH101 in supplemental Fig. S3).

The structure of SpGH101 could be further utilized to inform the design of novel therapeutics against the growing population of antibiotic resistant strains. The availability of genome sequence information has facilitated widespread genomic panning for genes essential to pathogenic organisms. SpGH101 was one of a number of essential genes inferred by such studies on S. pneumoniae (48) and represents a potential
therapeutic target. This result is in keeping with other pathogenic bacterial species where large, modular glycoside hydrolases have been shown to be essential virulence factors. The work presented here provides the opportunity to probe, at the molecular level, potential strategies for inhibition.

Sequence analysis suggests that glycoside hydrolases often utilize multidomain architectures. Unfortunately, their large size and modular nature has made them recalcitrant to structural analysis in the intact form; dissection of individual domains has been required, thus precluding the observation of structural details within the context of the assembled enzyme. The structure presented here, the first of an enzyme from family GH101, provides important insight into glycoside hydrolase multidomain assembly and will guide studies toward elucidating the biological function of this *Streptococcus* virulence factor.

Acknowledgments—We thank the staff at Beamline 8.2.2 of the Advanced Light Source (Berkeley, CA) for expert assistance.

REFERENCES

1. Bogaert, D., de Groot, R., and Hermans, P. W. M. (2004) *Lancet Infect. Dis.* 4, 144–154
2. Adam, D. (2002) *J. Antimicrob. Chemother.* 50, S1–S5
3. Levy, S. B., and Marshall, B. (2004) *Nat. Med.* 10, S122–S129
4. Burnaugh, A. M., Frantz, L. J., and King, S. J. (2008) *J. Bacteriol.* 190, 221–230
5. Fujita, K., Oura, F., Nagamine, N., Katayama, T., Hiratake, J., Sakata, K., Kumagai, H., and Yamamoto, K. (2005) *J. Biol. Chem.* 280, 37415–37422
6. Holm, L., and Sander, C. (1996) *Science* 273, 595–603
7. Desvaux, M., Meinke, A. L., Hanner, M., Henics, T., Bui, M. D., Gelbmann, D., Lundberg, U., Senn, B. M., Schunn, M., Habel, A., Henriques-Norback, C., Poerio, E., and Payan, F. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 55, 360–362
8. Desvaux, M., Dumas, E., Chafsey, I., and Hebraud, M. (2006) *FEMS Microbiol. Lett.* 256, 1–5
9. Santi, I., Scarselli, M., Mariani, M., Pezzicoli, A., Masignani, V., Taddei, A., Grandi, G., Telford, J. L., and Soriani, M. (2007) *Curr. Biol.* 17, 754–767
10. Zechel, D., and Withers, S. (2000) *Acc. Chem. Res.* 33, 11–18
11. Lee, H.-S., Kim, M.-S., Cho, H.-S., Kim, J.-I., Choi, J.-H., Park, C., Lee, H.-S., Oh, B.-H., and Park, K.-H. (2002) *J. Biol. Chem.* 277, 21891–21897
12. Lee, H.-S., Kim, M.-S., Cho, H.-S., Kim, J.-I., Choi, J.-H., Park, C., Lee, H.-S., Oh, B.-H., and Park, K.-H. (2002) *J. Biol. Chem.* 277, 21891–21897
13. Afonine, P. V., Grosse-Kunstleve, R. W., and Adams, P. D. (2005) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 62, 678–682
14. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. I., Wang, X., Murray, L. W., Arendall, W. B., Snoeinka, J., Richardson, J. S., and Richardson, D. C. (2007) *Nucleic Acids Res.* 35, W375–383
15. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, Version 1, DeLano Scientific, Palo Alto, CA
16. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) *Proc. Natl. Acad. Sci. U. S. A.* 98, 10037–10041
17. Schutteloepf, A. W., and van Aalten, D. M. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60, 1355–1363
18. Abe, A., Yoshida, H., Tonozuka, T., Sakano, Y., and Kamitori, S. (2005) *FEBS J.* 272, 6145–6153
19. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gross, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszeleski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 54, 905–921
20. Holm, L., and Sander, C. (1996) *Science* 273, 595–603
21. Carlson, C. B., Liu, Y., Keck, J. L., and Mosher, D. F. (2008) *J. Biol. Chem.* 283, 20069–20076
22. Brux, C., Ben-David, A., Shalim-Shezifi, D., Leon, M., Niefind, K., Shoham, G., Shoham, Y., and Schomburg, D. (2006) *J. Mol. Biol.* 359, 97–109
23. Hasegawa, Y., Suzuki, A., Kumamoto, K., Taddei, A., Grandi, G., Telford, J. L., and Soriani, M. (2007) *Curr. Biol.* 17, 754–767
24. Suzuki, A., Ichimura, H., Matsumoto, T., Yamauchi, T., and Shimizu, T. (1999) *FEMS Microbiol. Lett.* 179, 267–274
25. Zechel, D., and Withers, S. (2000) *Acc. Chem. Res.* 33, 11–18
26. Mariani, M., Pezzicoli, A., Masignani, V., Taddei, A., Grandi, G., Telford, J. L., and Soriani, M. (2007) *Curr. Biol.* 17, 754–767
27. Lee, H.-S., Kim, M.-S., Cho, H.-S., Kim, J.-I., Choi, J.-H., Park, C., Lee, H.-S., Oh, B.-H., and Park, K.-H. (2002) *J. Biol. Chem.* 277, 21891–21897
28. Lee, H.-S., Kim, M.-S., Cho, H.-S., Kim, J.-I., Kim, T.-J., Choi, J.-H., Park, C., Lee, H.-S., Oh, B.-H., and Park, K.-H. (2002) *J. Biol. Chem.* 277, 21891–21897
29. Mizuno, M., Tonozuka, T., Uechi, A., Ohtaki, A., Ichikawa, K., Kamitori, S., Nishikawa, A., and Sakano, Y. (2004) *Eur. J. Biochem.* 271, 2530–2538
30. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) *Biochem. J.* 382, 769–781
31. Bae, B., Ohene-Adjei, S., Kocherginskaya, S., Mackie, R. I., Spies, M. A., Cachat, K. E., and Nair, S. K. (2008) *J. Biol. Chem.* 283, 12415–12425
32. Boraston, A. B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D. R., Kilburn, D. G., and Davies, G. J. (2002) *J. Mol. Biol.* 319, 1143–1156
33. Simpson, P. J., Jamieson, S. J., Abu-Hachem, M., Nordberg Karlsson, E., Gilbert, H. J., Holst, O., and Williamson, M. P. (2002) *Biochemistry* 41, 5712–5719
34. Mizushima, T., Yoshida, Y., Kumanomidou, T., Hasegawa, Y., Suzuki, A., Yamane, T., and Tanaka, K. (2007) *Proc. Natl. Acad. Sci. U. S. A.* 104, 5777–5781
35. Xie, H., Gilbert, H. J., Charnock, S. J., Davies, G. J., Williamson, M. P., Simpson, P. J., Raghothama, S., Fontes, C. M. G. A., Dias, F. M. V., Ferreira, L. M. A., and Bolam, D. N. (2001) *Biochemistry* 40, 9167–9176
36. Charnock, S. J., Bolam, D. N., Shoham, G., Gilbert, H. J., and Ferreira, L. M. A., Davies, G. J., and Fontes, C. M. G. A. (2000) *Biochemistry* 39, 5013–5021
37. Sulzenbacher, G., Bignon, C., Nishimura, T., Tarling, C. A., Withers, S. G., Henrissat, B., and Bourne, Y. (2004) *J. Biol. Chem.* 279, 13119–13128
38. Song, J. H., Ko, K. S., Lee, J. Y., Baek, J. Y., Oh, W. S., Yoon, H. S., Jeong, J. Y., and Chun, J. (2005) *Mol. Cells* 19, 365–374

ACCELERATED PUBLICATION: Structural Analyses of SpGH101