BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

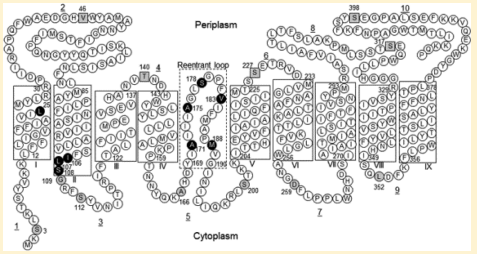
Topological Investigation of Glucosyltransferase V in *Shigella flexneri* using the Substituted Cysteine Accessibility Method

Anthony D. Rusden,[†] David P. Stephenson,[†] and Naresh K. Verma*

Division of Biomedical Science and Biochemistry, Research School of Biology, The Australian National University, Canberra, Australian Capital Territory 0200, Australia

Supporting Information

ABSTRACT: Modification of the lipopolysaccharide O-antigen of *Shigella* converts the serotype, which is significant as acquired immune responses are serotype specific. Glucosyltransferases (Gtrs) modify the O-antigen by the addition of glucosyl-groups; however the precise mechanism of O-antigen modification is not fully understood. This study aims to substantiate inferences made on the GtrV topological structure using the substituted cysteine accessibility method (SCAM). Twenty-one amino acid residues were tested to clarify three features of GtrV: the extramembrane regions, a proposed reentrant loop, and a membrane border region. Overall, the results agreed with a previous topology proposed for GtrV. The topology of GtrV consists of 11 extramembrane regions with a cytoplasmic N-terminus, periplasmic C-terminus and 9 transmembrane (TM) helices. The existence of a reentrant loop between TM helices IV and V was verified, and the cytoplasmic membrane border region of TM helix II was examined in depth.



Shigella is a Gram-negative bacterial pathogen and the causal agent of bacterial dysentery, also known as shigellosis, causing an estimated ~165 million cases and ~1.1 million deaths annually.[1] There are four serogroups of *Shigella* which are further divided into distinct serotypes: *Shigella dysenteriae* (serogroup A; 15 serotypes), *Shigella flexneri* (serogroup B; 16 serotypes), *Shigella boydii* (serogroup C; 20 serotypes), and *Shigella sonnei* (serogroup D; 1 serotype).[2,3] *S. flexneri* is hyperendemic in the developing world and is consequently the most common cause of *Shigella* infection worldwide.[4] Greater than 50% of all shigellosis episodes in South-East Asia, Latin America, Middle East, and Sub-Saharan Africa are due to *S. flexneri* infection.[4]

Lipopolysaccharides (LPS) are major bacterial outer membrane components playing key roles as virulence factors and protective structures for Gram-negative bacteria.[5,6] These molecules anchor in the outer membrane surface and interact with the extracellular environment. LPS are comprised of three structural components: lipid A, core polysaccharide and, O-antigen. The O-antigen of *S. flexneri* is comprised of a repeating tetrasaccharide subunit of *N*-acetylglucosamine − rhamnose I − rhamnose II − rhamnose III, which can be modified by the addition of phosphoethanolamine, glucosyl-, and/or O-acetyl-groups.[7,8] The modifications result in the conversion of the *Shigella* serotype, having important repercussions to host acquired immune responses as the immune response to *Shigella* is serotype specific. The basic unmodified *S. flexneri* O-antigen is known as serotype Y. The serotype can be converted to 16 altered serotypes based on the unique glucosyl-, phosphoethanolamine, and/or O-acetyl-modifications to the O-antigen.[7,8]

Glucosyltransferases (Gtrs) are a family of proteins responsible for O-antigen modification by the addition of glucosyl-groups. Currently, six *S. flexneri* Gtrs have been identified (GtrI,[9] GtrIc,[10] GtrII,[11] GtrIV,[12] GtrV,[13] and GtrX[14]), all of which are integral membrane proteins consisting of 8−10 transmembrane helices. GtrI, GtrII, GtrV, and GtrX have cytoplasmic N-termini and periplasmic C-termini,[15,16] while GtrIc and GtrIV have N- and C-termini located within the cytoplasm.[17,18] The active sites within these proteins are located in the large periplasmic loops located at the N-terminal and C-terminal ends of the protein.[16,17,19] Within the Carbohydrate Active Enzymes database (http://wwww.cazy.org/),[20] the *S. flexneri* Gtr proteins are "non classified" indicating they have weak similarity to other known glycosyltransferase families. Based upon of the known characteristics of the *S. flexneri* Gtrs, they are predicted to be members of the GT-C glycosyltransferase superfamily which consists of large hydrophobic proteins with 8−13 transmembrane domains, an active site located on a long-loop region, and utilize a lipid phosphate-activated donor sugar substrate.[21,22]

The *S. flexneri* Gtrs are encoded in a three-gene cluster consisting of *gtrA*, *gtrB*, and $gtr_{[type]}$.[23] The *gtrA* and *gtrB* genes are highly conserved, while the $gtr_{[type]}$ are unique to each serotype.[23] The proposed mechanism of the Gtr cluster is that on the cytoplasmic side of the inner membrane the GtrB transfers glucose from UDP-glucose to the bactoprenol phosphate carrier (UndP), creating UndP-glucose, which is then transferred across the inner membrane, from the

Received: February 10, 2013
Revised: March 26, 2013
Published: March 27, 2013

   dx.doi.org/10.1021/bi400168h | *Biochemistry* 2013, 52, 2655−2661

cytoplasm to the periplasm, via the flippase action of GtrA, and then the UndP-glucose binds with $Gtr_{[type]}$, which then attaches the glucosyl residue to the appropriate rhamnose residue of the O-antigen unit.[23,24] Glucosyltransferase V (GtrV) has one of the most extensively studied topologies and is responsible for converting the unaltered serotype-Y to serotype 5a by attaching a glucosyl-residue to rhamnose II of the O-antigen subunit via an α1,3 linkage.[7]

Understanding of the functional mechanisms of integral membrane proteins can begin by determining the topology of the protein within the membrane.[25,26] In *S. flexneri* serotype conversion, the mechanism behind $Gtr_{[type]}$ O-antigen modification is not fully understood and is currently under investigation. In order to predict mechanisms of function, it is necessary to elucidate structural elements of the protein such as orientation in the membrane and protein loop lengths.[15,16,25] The substituted cysteine accessibility method (SCAM) has been implemented as a means of determining membrane protein topology by probing the location of single amino acid residues. This method minimizes structural alteration to the protein, which is a notable advantage over other methods for topological investigation[26,27] such as a dual reporter system which has previously been performed on GtrV.[15]

EXPERIMENTAL PROCEDURES

Bacterial Strains, Plasmids, and Growth Conditions. The bacterial strains and plasmids used in this study are listed in Table S1. *Escherichia coli* and *Shigella flexneri* strains were grown in Luria−Bertani (LB) medium[28] at 37 and 30 °C, respectively, and supplemented when required with ampicillin (Amp) and/or kanamycin (Kan) at a concentration of 100 μg mL^{-1} and 50 μg mL^{-1}, respectively. Broth culture densities were monitored at 600 nm with an Ultrospec 10 Cell Density Meter spectrophotometer (GE Healthcare).

DNA Methods. PCR amplification was performed using the *PfuUltra II* Fusion HS DNA Polymerase (Stratagene) according to the manufacturer's directions using custom oligonucleotides (Sigma-Aldrich) listed in Table S2. PCR products were purified using the Wizard SV Gel and PCR Clean Up System (Promega). Plasmid isolation was performed using the Axyprep Plasmid Miniprep kit (Axygen Biosciences). Site-directed mutagenesis was performed using the Stratagene QuikChange Site-directed Mutagenesis procedure with the *PfuUltra II* Fusion HS DNA Polymerase (Stratagene) and DpnI restriction endonuclease (New England Biolabs). Mutated plasmids were transformed into *E. coli* XL1-Blue by electroporation and verified for the correct mutation by Sanger DNA sequencing (ACRF Biomolecular Resource Facility, John Curtin School of Medical Research, The Australian National University).

Creation of GtrV Cys-less Expression Vector. The plasmid pNV1077[15] was used as template DNA to create a Cys-less variant of by sequential substitution of the native Cys residues (C175, C215, C243, C271, C278, and C369) to Ala residues using site-directed mutagenesis (Korres, H. and Verma, N. K., unpublished data). The Cys-less variant plasmid (pNV1645) was used as a template and amplified using the primers pNV1077 GtrV SacI Fwd and pNV1077 GtrV *Hin*dIII (Table S2). The PCR products were purified and digested with SacI and *Hin*dIII restriction endonucleases (New England Biolabs) and ligated using T4 DNA ligase (Promega) into purified SacI and *Hin*dIII digested pBAD/Myc-His A (Invitrogen). The ligation mix was electroporated into *E. coli* JM109, and transformants were verified by DNA sequencing with the pBAD Forward and pBAD Reverse primers (Invitrogen). The resultant Amp^R plasmid, pNV1976, contained the *gtrV* gene under the control of the *araBAD* arabinose inducible promoter and encoded the GtrV protein with a MDPSSLE N-terminal and a RSKLGPEQKLISEEDLNSAV-DHHHHHH C-terminal extension. The pNV1976 plasmid was used as template DNA for the creation of 21 single Cys variants via site-directed mutagenesis (Table S2). These mutagenized plasmids were electroporated into *E. coli* TOP10 for expression studies.

Functionality Testing of Single-Cys GtrV. The single Cys variant plasmids were electroporated into the *S. flexneri* serotype Y strain SFL1616 in order to examine the functionality of the mutants. SFL1616 contains the plasmid pNV1241 which harbors the Kan^R gene and the GtrA and GtrB genes and complementation of this plasmid with the plasmid containing the single Cys GtrV completes the three-gene cluster required for serotype conversion. These transformants were grown on $LB_{Amp100;Kan50}$ agar overnight at 30 °C. Single colonies were picked from plates and added to a drop of *S. flexneri* type V antisera (Denka Seikan) on a glass slide. Functionality was indicated by agglutination of the *S. flexneri* cells with the antisera. Only agglutination which occurred within 1 min was recorded as positive. Agglutination levels were compared against that observed for the Cys-less variant of GtrV (SFL2376) and recorded as either high (+ + +), medium (+ +), low (+), or no agglutination (−) (Table 1).

Table 1. Serum Slide Agglutination Assay Using Type V Antisera against *Shigella* Strains Expressing Single-Cys Variants of GtrV

strain	residue mutation in Cys-less GtrV	agglutination level[a]	functional
SFL1616	no GtrV	−	N/A
SFL2376	none (Cys-less)	+ + +	yes
SFL2417	Ser3 → Cys	+ + +	yes
SFL2418	Leu25 → Cys	+ + +	yes
SFL2419	Val46 → Cys	+ + +	yes
SFL2420	Ser112 → Cys	+ + +	yes
SFL2421	Thr140 → Cys	+	yes
SFL2429	Ala166 → Cys	+ + +	yes
SFL2422	Ser178 → Cys	+	yes
SFL2423	Ser200 → Cys	+ + +	yes
SFL2424	Ser227 → Cys	+ + +	yes
SFL2425	Ser317 → Cys	+ + +	yes
SFL2426	Ser398 → Cys	+ + +	yes
SFL2427	Asp259 → Cys	+ + +	yes
SFL2428	Leu352 → Cys	+ + +	yes
SFL2430	Iso106 → Cys	+ +	yes
SFL2431	Leu107 → Cys	+ + +	yes
SFL2432	Ser108 → Cys	+ + +	yes
SFL2433	Gly109 → Cys	+ +	yes
SFL2434	Ala171 → Cys	+	yes
SFL2435	Ala175 → Cys	+ + +	yes
SFL2436	Val183 → Cys	+ + +	yes
SFL2437	Met188 → Cys	+ + +	yes

[a]Levels of agglutination were observed as being either high (+ + +), medium (+ +), low (+), or no agglutination (−). The level of agglutination is an indication of the level of functionality of GtrV. No agglutination indicates GtrV is not functional or not present.

dx.doi.org/10.1021/bi400168h | *Biochemistry* 2013, 52, 2655−2661

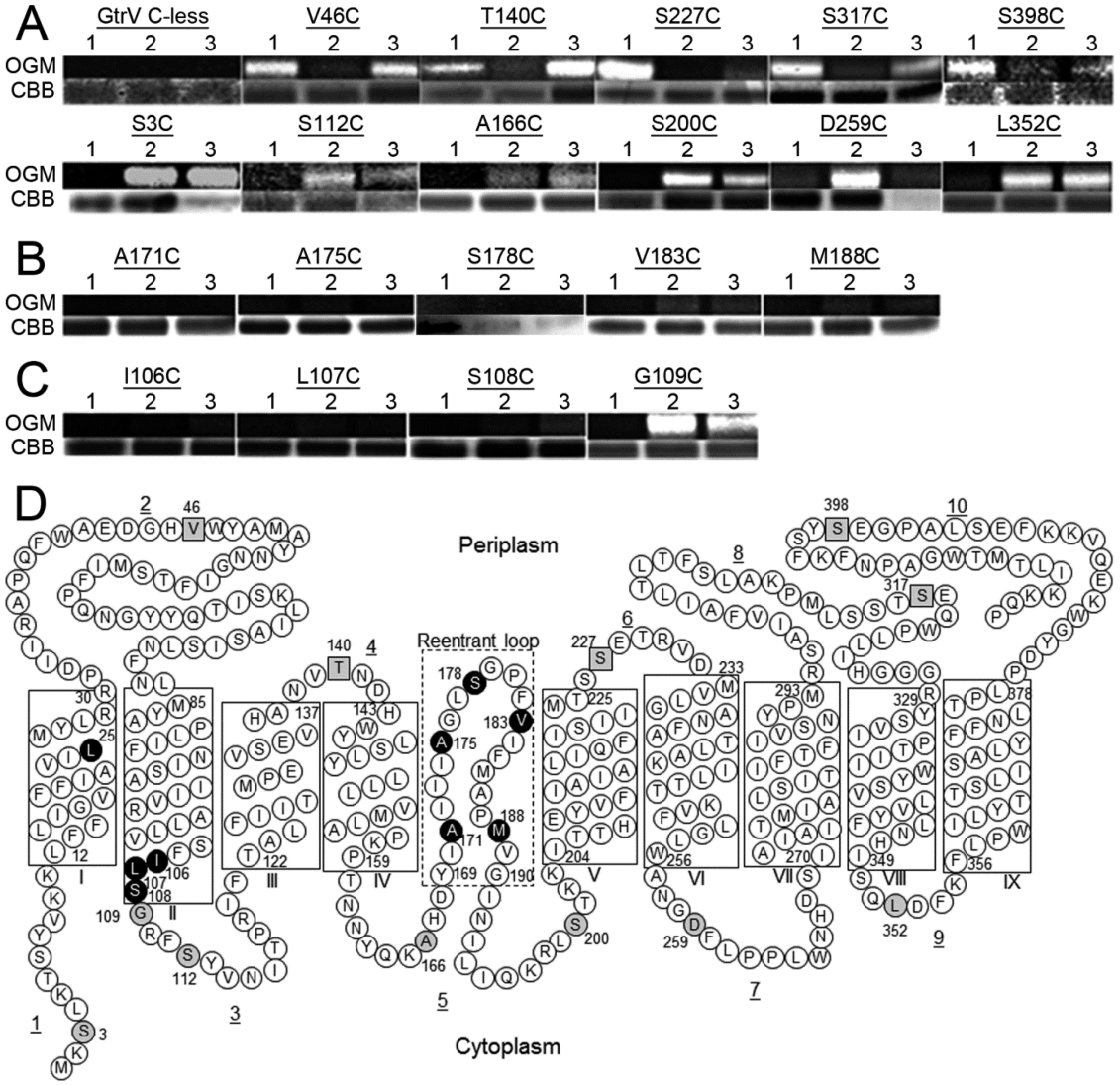


Figure 1. SCAM labeling of single-Cys GtrV variants. (A) Analysis of putative extramembrane located residues. (B) Analysis of putative reentrant loop residues. (C) Analysis of the cytoplasmic membrane border region of TM helix II. Labeling of lanes in A–C as 1, 2, and 3 refers to OGM treatments labeling cysteine's located in the periplasm, cytoplasm, or both, respectively. OGM indicated gels which were analyzed for OGM labeling of cysteine residues, while CBB indicates the same gel which was subsequently stained with Coomassie Brilliant Blue. (D) Revised topology map of GtrV. The extramembrane loops are labeled with underlined numerals, while TM helicies boxed and are numbered with Roman numerals. Predicted membrane border residues are numbered within the transmembrane boxes. The reentrant loop is indicated with a dashed box. The residue number of single cysteine mutants are indicated, while residues which were present in the periplasm are indicated by gray squares, residues located in the cytoplasm are indicated by gray circles, and residues which were not labeled by OGM representing within membrane locations are indicated by black circles.

GtrV Expression. Single colonies were used to inoculate LB_{Amp100} broth and were incubated overnight at 37 °C at 200 rpm. Three milliliters of the overnight starter culture was used to inoculate 200 mL (1/66 dilution) of prewarmed LB_{Amp100} broth which was incubated at 37 °C at 200 rpm for ~3 h until the culture reached an optical density (OD_{600}) of 0.6. GtrV expression was induced with L-arabinose at a final concentration of 0.2% (w/v). Cultures were grown for 4 h at 37 °C at 200 rpm. The cells were harvested by centrifugation (10000*g*; 15 min; 4 °C), the supernatant was discarded, and the cell pellet was washed twice with 10 mL of 30 mM Tris-HCl (pH 8.0) and pelleted by centrifugation (3000*g*; 10 min; 4 °C). The cell pellets were stored at −80 °C.

SCAM Labeling of Single Cysteines. Labeling GtrV with Oregon Green 488 maleimide carboxylic acid (OGM) was adapted from the protocol described by Larue et al.[29] Cell pellets were thawed from −80 °C on ice and were resuspended in 9 mL of buffer A (50 mM Tris-Cl, pH 8.0, containing 100 mM NaCl) and divided into three 3 mL aliquots. All aliquots were protected from light throughout the entire process by covering in aluminum foil. OGM dissolved in dimethylformamide (Sigma) was added to aliquot 1 (40 μM), labeling cysteine residues accessible to the periplasm. The sample was incubated for 15 min at room temperature and quenched with β-mercaptoethanol (6 mM). In aliquot 2, the periplasmic cysteine residues were blocked by the addition of methanethiosulfonate ethyltrimethylammonium (MTSET) (2 mM) for 15 min at room temperature. Aliquots 1 and 2 were pelleted by centrifugation (3000*g*; 10 min; 4 °C) and washed twice with 10 mL of buffer A. All three aliquots were pelleted by centrifugation and resuspended in 2 mL of buffer A containing 5 mM EDTA, 100 mg mL^{-1} lysozyme, and 20% (w/v) sucrose

dx.doi.org/10.1021/bi400168h | *Biochemistry* 2013, 52, 2655–2661

and incubated at room temperature for 20 min. Ten milliliters of cold water was added to each aliquot, and they were lysed by a single pass through an Emulsiflex-B15 high pressure homogenizer (Avestin). Aliquots 2 and 3 were incubated with OGM (10 μM) for 15 min at room temperature, labeling any available and unblocked cytoplasmic cysteine residues in aliquot 2 and labeling all of the accessible cysteine residues in aliquot 3. The labeling reaction was quenched in aliquots 2 and 3 with β-mercaptoethanol (6 mM). Cellular debris was removed by centrifugation (15000*g*; 15 min; 4 °C), and the cell-free supernatant was centrifuged (100000*g*; 30 min; 4 °C) to collect the membrane proteins.

GtrV Purification. The membrane protein pellet was resuspended in 1 mL of buffer A containing 0.5% (w/v) *n*-dodecyl-beta-D-maltoside (DDM) and homogenized in a glass tissue homogenizer. The homogenized membrane proteins were protected from light and left to solubilize overnight at 4 °C with gentle agitation. Solubilized protein was mixed with 100 μL of nickel affinity resin (Novagen) for 1 h at room temperature with agitation. The solubilized membrane resin mixture was loaded into 1 mL microcentrifuge snap-cap spin columns within 2 mL microcentrifuge tubes (Pierce). Five-hundred microliters of wash buffer 1 (buffer A containing 20 mM imidazole, 0.025% [v/v] DDM) was added to the spin column and centrifuged at 800*g* for 1 min. The flowthrough was discarded and the wash was repeated three times. The resin was washed with 500 μL of wash buffer 2 (buffer A containing 50 mM imidazole, 0.025% (v/v) DDM) and centrifuging at 800*g* for 1 min, discarding the flowthrough afterward. Proteins were eluted with 200 μL of elution buffer (buffer A containing 500 mM imidazole, 1% (w/v) SDS).

Protein Analysis. Eighty microliters of protein sample was mixed with 20 μL of 4× SDS loading buffer (200 mM Tris-Cl (pH6.8), 8% (w/v) SDS, 0.4% (w/v) bromophenol blue, 40% glycerol, 200 mM β-mercaptoethanol (added immediately prior to use)) and warmed for 1 h at 40 °C. Sixty microliters of sample was loaded per well in a 1.5 mm 12% (w/v) polyacrylamide gel, and 30 μL of sample was loaded into the well of a 0.75 mm 12% (w/v) polyacrylamide gel. The gels were electrophoresed at 200 V for approximately 50 min, or until the dye front had run off the gel. Visualization of the SCAM analysis was performed following SDS-PAGE by placing the gels in a Gel Doc XR+ System (BioRad) with the Xcita Blue conversion screen and subjecting the gels to UV trans-illumination, causing the OGM bound to protein to emit a signal which was used to detect fluorescence profiles of proteins. The 1.5 mm SDS-PAGE gel was stained with Coomassie Blue allowing visualization and quantification of GtrV protein bands.

Western Blot Analysis. Proteins from the 0.75 mm SDS-PAGE gel were transferred to PVDF membrane (Millipore) using the Mini Trans-Blot Cell (BioRad) at 40 V for 2 h at 4 °C. The membrane was wet in methanol for 15 s and removed to dry for 10 min. Once completely dry, the membrane was rewet in 100% methanol and transferred to a SNAP-ID blot holder (Millipore) prewet with phosphate buffered saline containing 0.05% (v/v) Tween-20 (PBST). The PVDF membrane was blocked with blocking solution (PBST containing 0.5% w/v skim milk powder) which was drawn through membrane using a vacuum pump. The membrane was incubated for 10 min with Anti-His/Hrp conjugate antibody (R&D Systems) diluted 1/3000 (v/v) in PBST. The antibody was then pulled through membrane using a vacuum pump, and the membrane was washed with three cycles of 30 mL of PBST. Antibody binding was visualized using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) and visualized using a Fusion SL chemiluminescence system (Vilber Lourmat).

RESULTS

Functionality of Single-Cys GtrV Variants. Functionality of GtrV was determined using antisera slide agglutination testing. Agglutination of SFL1616 (serotype Y) expressing a single-Cys GtrV variant with type V antisera indicated the strain had undergone serotype conversion from type Y to type 5a and the GtrV variant was functional. Agglutination was observed for each *S. flexneri* strain expressing a single-Cys variant of GtrV, indicating that each variant strain had been converted from serotype Y to serotype 5a by a functional GtrV (Table 1). Agglutination levels of the single-Cys variants were compared to the Cys-less variant and most strains demonstrated high levels of agglutination indicating fully functional GtrV. Lower levels of agglutination were observed when T140, A171, and S178 were mutated to cysteine, indicating reduced ability of GtrV to modify O-antigen. Even though these mutants showed reduced agglutination, they were functional, implying structural integrity, and therefore were also included in the SCAM analysis.

Investigating the Proposed Topology of GtrV. The 11 extramembrane regions previously predicted for GtrV (five located periplasmically and six cytoplasmically) were examined using OGM labeling (Figure 1A). The cysteine-less GtrV variant was tested as a negative control and showed no fluorescence in any of the three treatments. The L25 residue was tested as a within membrane control, and like the Cys-less variant the L25C variant showed no fluorescence (data not shown). Absence of fluorescence in all treatments is typical of cysteines located within the membrane where they cannot be accessed by OGM. Therefore this result confirms the within membrane location of the L25C residue. The single cysteine GtrV mutants of V46, T140, S227, S317, and S398 were used to analyze the periplasmic predicted regions, while mutants of S3, S112, A166, S200, D259, and L352 were used to analyze the cytoplasmic predicted regions. Fluorescence results for these residues demonstrate all residues putatively located in the periplasm had fluorescence in treatments 1 and 3 (periplasmic OGM labeling), while putative cytoplasmic amino acids had fluorescence in treatments 2 and 3 (Figure 1A). These results indicate that each residue is located in their predicted location. The third labeling treatment labels both cytoplasmic and periplasmic residues and should fluoresce whenever an extramembrane cysteine is present. Mutants S227C and D259C were exceptions where the fluorescence in treatment 3 was markedly reduced compared with the fluorescence in treatment 1 and 2, respectively. The amount of protein analyzed in treatment 3 of S227C appears to be consistent with treatments 1 and 2, indicating some experimental error with the OGM labeling of treatment 3. The treatment 3 sample of D259C contained a low concentration of GtrV protein, as observed in the Coomassie stained gel image, explaining why very little OGM signal was detected. The SCAM results support the topology map of GtrV proposed by Korres and Verma[15] and indicate that GtrV has 11 extramembrane regions, 5 periplasmic and 6 cytoplasmic, including a cytoplasmic N-terminus and a periplasmic C-terminus.

Investigation into the Proposed Reentrant Loop. Five residues (A171, A175, S178, V183, and M188) distributed within the proposed reentrant loop region were substituted with cysteine to create single-Cys variants of GtrV. The variants A171C, A175C, and S178C all produced within membrane labeling profiles, while V183C and M188C produced very faint labeling in treatments 2 and 3 (Figure 1B). This suggests that V183C and M188C are likely to be located within the membrane but are close to the cytoplasmic border, and therefore a small amount of labeling has occurred. These results in addition to cytoplasmic location of the residues A166C and S200C suggest that A171C, A175C, S178C, V183C, M188C exist within the membrane as part of a reentrant loop between TM helices IV and V.

Defining the Cytoplasmic Border of Transmembrane Helix II. Four consecutive residues predicted to be at the cytoplasmic border of TM helix II were individually mutated to cysteine and analyzed as a proof of concept to determine the resolution at which OGM could be used to define borders in SCAM. Residues I106C and L107C show profiles typical of within membrane cysteines, while S108C shows slight fluorescence in treatments 2 and 3 (Figure 1C), suggesting inefficient labeling by OGM which may be due to the residue being close to the cytoplasmic space; however these results do not suggest S108C is fully present in the cytoplasm. G109C clearly identifies with a cytoplasmic fluorescence profile (Figure 1C). These results indicate that it is possible to clearly define the membrane-protein borders of GtrV using OGM in SCAM analysis.

Refinement of the Existing GtrV Topology Map. Using the experimentally defined locations within this study, a proposed topology map of GtrV has been presented (Figure 1D). The topology map illustrates the location of all the single cysteine variants generated in this study and revised the cytoplasmic border region of TM helix II.

DISCUSSION

Investigation into TM protein structure is often used to elucidate mechanisms behind the protein functionality. Determination of membrane protein topology can be achieved using a number of techniques and is useful in identifying the orientation, number, and size of extramembrane and TM regions. This gives insight into regions which are involved in catalytic activity and thus the grounds for further investigation to identify potential functionality. The mechanism of $Gtr_{[type]}$ in O-antigen modification is yet to be confirmed. The process of O-antigen modification causes *Shigella* serotype conversion, which is important in vaccine development.

The SCAM analysis used in this study was able to identify the location of each cysteine in 21 single-Cys variants of GtrV. Probing the locations of residues S3C and S398C confirmed the topological locations of the N- and C-termini as cytoplasmic and periplasmic, respectively. A cytoplasmic N-terminal is one of the few structures conserved among all $Gtr_{[type]}$[16] and may suggest a possible conserved role. The identification of nine other residues located in extramembrane regions additionally confirmed the existence of 11 extramembrane regions (6 cytoplasmic and 5 periplasmic regions), agreeing with the topology map generated using the Pho-Lac reporter system.[15] The reason for the reduced fluorescence in treatment 3 of S227C appears to be poor labeling of the GtrV protein, potentially through experimental error, while the treatment 3 sample of D259C contained very little protein as observed in the Coomassie stained gel image (Figure 1A). Fortunately, residue topology can be accurately determined using treatments 1 and 2 only;[30−32] thus the topology of these residues was still able to be confirmed. Furthermore, all of the residues that were selected for investigation were found to be in the same topological space predicted by the Pho-Lac analysis by Korres and Verma.[15] These results suggest that the Pho-Lac reporter system was an appropriate method for defining an approximate topology map; however SCAM analysis should be considered a superior alternative as there is little question about the structural integrity of the protein. Furthermore the ability of SCAM to define membrane borders and the within membrane amino acids allows for refinement of protein structures not achievable with the Pho-Lac reporter system.

OGM labeling of the residues at the cytoplasmic membrane border of TM helix II was very successful in defining the border at the residue level. Further analysis of the other membrane borders is now possible using this technique and would be of benefit in providing accurate data regarding how the protein is inserted into the membrane which may potentially assist in defining how the protein functions. SCAM provides a very accurate analysis of the residues in and around the membrane border in an *in vivo* context. In the future, this analysis could be paired with structural predictions or additional structural analysis, such X-ray crystallography, to provide an accurate *in vivo* topological map of GtrV.

The GtrV topology model presented by Korres and Verma[15] proposed the existence of a reentrant loop even though this model contradicted the *in silico* models. The region proposed to be the reentrant loop was predicted by the *in silico* algorithms to be a transmembrane helix; yet the Pho-Lac reporter data suggested that residues Q164 and R198, on either side of this predicted transmembrane helix, were both located cytoplasmically. In order to match the predicted topology model with the experimental data, using the Pho-Lac reporter, it was proposed that the protein consisted of nine transmembrane helices, and the hydrophobic region between residues Q164-R198 was a reentrant loop which dipped into and out of the membrane on the cytoplasmic side. Interestingly, recent topology predictions using OCTOPUS[33] and MEMSAT-SVM,[34] both of which are designed to identify reentrant loops, also predicted a transmembrane helix rather than a reentrant loop at that location (data not shown). To confirm that the region did contain a reentrant loop, the cytoplasmic regions surrounding the loop and residues within the loop were subjected to SCAM analysis. The cytoplasmic location of two residues in close proximity to Q164 and R198 (A166 and S200, respectively) was confirmed using SCAM, verifying the results previously obtained for these residues by Korres and Verma.[15] Additionally, five single Cys substituted residues within the 164−198 amino acid region (A171C, A175C, S178C, V183C, and M188C) could not be efficiently labeled by OGM suggesting they are located within the membrane as part of a reentrant loop structure. Low fluorescence in treatments 2 and 3 of V183C and M188C suggested some labeling with OGM had occurred, which could be explained if V183 and M188 exist close to the membrane border. On the basis of these observations and the analysis by Korres and Verma,[15] it is predicted that the reentrant loop exists between Y169 and G190. This region has a composition typical of reentrant loops, containing amino acids with high and low levels of hydrophobicity, and no hydrophilic amino acids.[35] As mentioned above, this predicted region could be further refined using SCAM labeling of sequential residues to

dx.doi.org/10.1021/bi400168h | *Biochemistry* 2013, 52, 2655−2661

accurately define the borders of the reentrant loop. Detailed knowledge of this region may assist with determining how the reentrant loop moves within the membrane and in defining the role the reentrant loop has in GtrV function.

Extramembrane loops 2 and 10 have been investigated previously for functional roles based on their relative large size and periplasmic location.[19,36] Extramembrane loop 8 has previously been overlooked for functional investigation despite being as large as loops 2 and 10 and located periplasmically, the location where GtrV is proposed to function. Given the size and location of loop 8, it too may potentially be involved in the function of the GtrV protein. The existence of a large periplasmic loop between the reentrant loop and the C-terminal has also been proposed in GtrIc, GtrII, and GtrX, however not in GtrIV.[15−18] The large size and location, coupled with conservation of a similar structure in other $Gtr_{[type]}$, suggest extramembrane loop 8 has a functional role. As has been performed previously with loops 2 and 10 to infer function, deletion and/or replacement mutations could be performed on residues in loop 8 to begin analyzing possible functional elements of the loop. Critical regions could then be subjected to single amino acid substitutions such as those used in previous studies to identify critical residues and motifs in membrane proteins.[19,30,37] Such techniques could be employed to identify any critical residues/motifs in loop 8 and therefore ascertain whether this loop is functionally important.

Evidence suggesting that flexibility in the O-mannosyltransferase ScPmt1p is required for catalytic activity[38] has led to the hypothesis that GtrV's reentrant loop may play a similar role, allowing catalytic domains, such as those previously identified in loops 2 and 10, to interact.[19] Two single-Cys mutants probing the reentrant loop presented reduced functionality. These mutations may have impacted critical residues within the reentrant loop or disrupted its native structure such that it reduced the loop's ability to permit the flexibility required for interaction of loops 2 and 10, thus implying the reentrant loop is critical in GtrV function and implying a role in other $Gtr_{[type]}$. Given the previous analysis indicating the reentrant loops requirement for function and the reduced function observed in this study, the role and function of the reentrant loop required further investigation.

In conclusion, the previously proposed topology of GtrV was verified in this study using SCAM, confirming the presence of 11 extramembrane domains, 9 transmembrane helices, and the existence of a reentrant loop. Additionally, the cytosolic/membrane border of TM helix II was accurately defined using the SCAM technique. Greater knowledge in the structure of this protein enhances the understanding of O-antigen modification mechanisms of GtrV and subsequently other $Gtr_{[type]}$ which could potentially be exploited in attempts to develop a vaccine against multiple serotypes of *S. flexneri.*

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Supporting tables containing the bacterial strains, plasmids, and primers used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Mailing address: Research School of Biology, Bldg. 134 Linnaeus Way, The Australian National University, Canberra, ACT, Australia 0200. Phone: 61-2-6125-2666. Fax: 61-2-6125-0313. E-mail: naresh.verma@anu.edu.au.

Author Contributions

†(A.D.R. and D.P.S) These two authors contributed equally to this work and are considered cofirst authors

Funding

This research was supported by the National Health and Medical Research Council of Australia, Project Number 1004846.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We wish to thank Haralambos Korres for the production and provision of the Cys-less GtrV mutant.

■ ABBREVIATIONS

Amp, ampicillin; CBB, Coomassie Brilliant Blue; Cys, Cysteine; Cys-less, Cysteine-less; DDM, *n*-dodecyl-beta-D-maltoside; Gtr, glucosyltransferase; Kan, kanamycin; LB, Luria−Bertani; LPS, lipopolysaccharide; MTSET, methanethiosulfonate ethyltrimethylammonium; OGM, Oregon green maleimide 488 carboxylic acid; SCAM, Substituted cysteine accessibility method; TM, transmembrane

■ REFERENCES

(1) Camacho, A. I., Irache, J. M., and Gamazo, C. (2013) Recent progress towards development of a Shigella vaccine. *Expert Rev. Vaccines 12*, 43−55.

(2) Niyogi, S. K. (2005) Shigellosis. *J. Microbiol. 43*, 133−143.

(3) Sun, Q., Lan, R., Wang, Y., Wang, J., Luo, X., Zhang, S., Li, P., Ye, C., Jing, H., and Xu, J. (2011) Genesis of a novel *Shigella flexneri* serotype by sequential infection of serotype-converting bacteriophages SfX and SfI. *BMC Microbiol. 11*, 269.

(4) Kotloff, K. L., Winickoff, J. P., Ivanoff, B., Clemens, J. D., Swerdlow, D. L., Sansonetti, P. J., Adak, G. K., and Levine, M. M. (1999) Global burden of Shigella infections: Implications for vaccine development and implementation of control strategies. *Bull. World Health Organ. 77*, 651−666.

(5) Zhong, Q. P. (1999) Pathogenic effects of O-polysaccharide from *Shigella flexneri* strain. *World J. Gastroenterol. 5*, 245−248.

(6) Tran, A. X., Dong, C., and Whitfield, C. (2010) Structure and functional analysis of LptC, a conserved membrane protein involved in the lipopolysaccharide export pathway in *Escherichia coli*. *J. Biol. Chem. 285*, 33529−33539.

(7) Allison, G. E., and Verma, N. K. (2000) Serotype-converting bacteriophages and O-antigen modification in Shigella flexneri. *Trends Microbiol. 8*, 17−23.

(8) Sun, Q., Knirel, Y. A., Lan, R., Wang, J., Senchenkova, S. N., Jin, D., Shashkov, A. S., Xia, S., Perepelov, A. V., Chen, Q., Wang, Y., Wang, H., and Xu, J. (2012) A Novel Plasmid-Encoded Serotype Conversion Mechanism through Addition of Phosphoethanolamine to the O-Antigen of *Shigella flexneri*. *PLoS ONE 7*, No. e46095.

(9) Bastin, D. A., Lord, A., and Verma, N. K. (1997) Cloning and analysis of the glucosyl transferase gene encoding type I antigen in Shigella flexneri. *FEMS Microbiol. Lett. 156*, 133−139.

(10) Stagg, R. M., Tang, S. S., Carlin, N. I. A., Talukder, K. A., Cam, P. D., and Verma, N. K. (2009) A novel glucosyltransferase involved in O-antigen modification of *Shigella flexneri* serotype 1c. *J. Bacteriol. 191*, 6612−6617.

(11) Mavris, M., Manning, P. A., and Morona, R. (1997) Mechanism of bacteriophage SfII-mediated serotype conversion in Shigella flexneri. *Mol. Microbiol. 26*, 939−950.

(12) Adams, M. M., Allison, G. E., and Verma, N. K. (2001) Type IV O antigen modification genes in the genome of *Shigella flexneri* NCTC 8296. *Microbiology 147*, 851−860.
(13) Huan, P. T., Whittle, B. L., Bastin, D. A., Lindberg, A. A., and Verma, N. K. (1997) *Shigella flexneri* type-specific antigen V: Cloning, sequencing and characterization of the glucosyl transferase gene of temperate bacteriophage SfV. *Gene 195*, 207−216.
(14) Verma, N. K., Verma, D. J., Huan, P. T., and Lindberg, A. A. (1993) Cloning and sequencing of the glucosyl transferase-encoding gene from converting bacteriophage X (SFX) of *Shigella flexneri*. *Gene 129*, 99−101.
(15) Korres, H., and Verma, N. K. (2004) Topological analysis of glucosyltransferase GtrV of *Shigella flexneri* by a dual reporter system and identification of a unique reentrant loop. *J. Biol. Chem. 279*, 22469−22476.
(16) Lehane, A. M., Korres, H., and Verma, N. K. (2005) Bacteriophage-encoded glucosyltransferase GtrII of *Shigella flexneri*: Membrane topology and identification of critical residues. *Biochem. J. 389*, 137−143.
(17) Ramiscal, R. R., Tang, S. S., Korres, H., and Verma, N. K. (2010) Structural and functional divergence of the newly identified GtrIc from its Gtr family of conserved *Shigella flexneri* serotype-converting glucosyltransferases. *Mol. Membr. Biol. 27*, 114−122.
(18) Nair, A., Korres, H., and Verma, N. K. (2011) Topological characterisation and identification of critical domains within glucosyltransferase IV (GtrIV) of *Shigella flexner*. *BMC Biochem. 12*, 67.
(19) Korres, H., and Verma, N. K. (2006) Identification of essential loops and residues of glucosyltransferase V (GtrV) of *Shigella flexneri*. *Mol. Membr. Biol. 23*, 407−419.
(20) Cantarel, B. I., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res. 37*, D233−D238.
(21) Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: Structures, functions, and mechanisms. *Annu. Rev. Biochem. 77*, 521−555.
(22) Liu, J., and Mushegian, A. (2003) Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci. 12*, 1418−1431.
(23) Korres, H., Mavris, M., Morona, R., Manning, P. A., and Verma, N. K. (2005) Topological analysis of GtrA and GtrB proteins encoded by the serotype-converting cassette of *Shigella flexneri*. *Biochem. Biophys. Res. Commun. 328*, 1252−1260.
(24) Guan, S., Bastin, D. A., and Verma, N. K. (1999) Functional analysis of the O antigen glucosylation gene cluster of *Shigella flexneri* bacteriophage SfX. *Microbiology 145*, 1263−1273.
(25) Bogdanov, M., Zhang, W., Xie, J., and Dowhan, W. (2005) Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesi. *Methods 36*, 148−171.
(26) Zhu, Q., and Casey, J. R. (2007) Topology of transmembrane proteins by scanning cysteine accessibility mutagenesis methodology. *Methods 41*, 439−450.
(27) Alexeyev, M. F., and Winkler, H. H. (1999) Membrane topology of the *Rickettsia prowazekii* ATP/ADP translocase revealed by novel dual pho-lac reporters. *J. Mol. Biol. 285*, 1503−1513.
(28) Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
(29) Larue, K., Ford, R. C., Willis, L. M., and Whitfield, C. (2011) Functional and structural characterization of polysaccharide co-polymerase proteins required for polymer export in ATP-binding cassette transporter-dependent capsule biosynthesis pathways. *J. Biol. Chem. 286*, 16658−16668.
(30) Nanatani, K., Maloney, P. C., and Abe, K. (2009) Structural and functional importance of transmembrane domain 3 (TM3) in the aspartate:alanine antiporter AspT: Topology and function of the residues of TM3 and oligomerization of AspT. *J. Bacteriol. 191*, 2122−2132.
(31) Nanatani, K., Fujiki, T., Kanou, K., Takeda-Shitaka, M., Umeyama, H., Ye, L., Wang, X., Nakajima, T., Uchida, T., Maloney, P. C., and Abe, K. (2007) Topology of AspT, the Aspartate:Alanine Antiporter of Tetragenococcus halophilus, Determined by Site-Directed Fluorescence Labeling. *J. Bacteriol. 189*, 7089−7097.
(32) Kimura, T., Ohnuma, M., Sawai, T., and Yamaguchi, A. (1997) Membrane topology of the transposon 10-encoded metal-tetracycline/H+ antiporter as studied by site-directed chemical labeling. *J. Biol. Chem. 272*, 580−585.
(33) Viklund, H., and Elofsson, A. (2008) OCTOPUS: Improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics 24*, 1662−1668.
(34) Nugent, T., and Jones, D. T. (2009) Transmembrane protein topology prediction using support vector machines. *BMC Bioinformatics 10*, 159.
(35) Yan, C., and Luo, J. (2010) An analysis of reentrant loops. *Protein J. 29*, 350−354.
(36) Moscoso, J. A., Korres, H., George, D. T., and Verma, N. K. (2010) Identification of active site residues in the *Shigella flexneri* glucosyltransferase GtrV. *Mol. Membr. Biol. 27*, 104−113.
(37) Marolda, C. L., Li, B., Lung, M., Yang, M., Hanuszkiewicz, A., Rosales, A. R., and Valvano, M. A. (2010) Membrane topology and identification of critical amino acid residues in the Wzx O-antigen translocase from *Escherichia coli* O157:H4. *J. Bacteriol. 192*, 6160−6171.
(38) Girrbach, V., Zeller, T., Priesmeier, M., and Strahl-Bolsinger, S. (2000) Structure-function analysis of the dolichyl phosphate-mannose: Protein O-mannosyltransferase ScPmt1p. *J. Biol. Chem. 275*, 19288−19296.