JB
Journals.ASM.org

Membrane Topology of the *Bacillus anthracis* GerH Germinant Receptor Proteins

Mary J. Wilson, Paul E. Carlson, Brian K. Janes, and Philip C. Hanna

Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan, USA

Bacillus anthracis **spores are the etiologic agent of anthrax. Nutrient germinant receptors (nGRs) packaged within the inner membrane of the spore sense the presence of specific stimuli in the environment and trigger the process of germination, quickly returning the bacterium to the metabolically active, vegetative bacillus. This ability to sense the host environment and initiate germination is a required step in the infectious cycle. The nGRs are comprised of three subunits: the A-, B-, and C-type proteins. To date there are limited structural data for the A- and B-type nGR subunits. Here the transmembrane topologies of the *B. anthracis* $GerH_A$, $GerH_B$, and $GerH_C$ proteins are presented. C-terminal green fluorescent protein (GFP) fusions to various lengths of the GerH proteins were overexpressed in vegetative bacteria, and the subcellular locations of these GFP fusion sites were analyzed by flow cytometry and protease sensitivity. GFP fusion to full-length $GerH_C$ confirmed that the C terminus of this protein is extracellular, as predicted. $GerH_A$ and $GerH_B$ were both predicted to be integral membrane proteins by topology modeling. Analysis of C-terminal GFP fusions to full-length $GerH_B$ and nine truncated $GerH_B$ proteins supports either an 8- or 10-transmembrane-domain topology. For $GerH_A$, C-terminal GFP fusions to full-length $GerH_A$ and six truncated $GerH_A$ proteins were consistent with a four-transmembrane-domain topology. Understanding the membrane topology of these proteins is an important step in determining potential ligand binding and protein-protein interaction domains, as well as providing new information for interpreting previous genetic work.**

Spore-forming bacteria exist in two distinct morphotypes: a metabolically dormant spore and the actively growing vegetative bacillus. Bacterial spores are formed in response to nutrient starvation and are resistant to many environmental stressors, allowing them to remain viable in the environment for years (33, 39). When favorable conditions are encountered, spores are able to sense the environment and quickly return to the vegetative state through the process of germination (8). For *Bacillus anthracis* and many other pathogenic sporeformers, the ability to sense the host environment and initiate germination is vital for the infectious cycle (8). While spores are the infectious particles, germination associated with alveolar phagocytes is required for disease pathology during anthrax infection (33, 39). Genetic studies of many *Bacillus* species, including *B. anthracis*, have identified the proteins responsible for this environmental sensing, which are called the nutrient germinant receptors (nGRs) (4, 17, 28, 37, 46).

Each nGR is made up of three proteins, the A-, B-, and C-type proteins, which are encoded on a family of tricistronic operons. Many species express multiple nGR operons that are believed to have been acquired by processes of gene duplication and divergence (27). *B. anthracis* has five functional nGRs, annotated as GerH, GerK, GerL, GerS, and GerX, and these nGR proteins are homologous, with 30 to 60% similarity between receptor subunits (12, 17, 40, 46). The known ligands for four of these (GerH, GerK, GerL, and GerS) include combinations of L-amino acids and purine nucleosides; the cognate ligand(s) for GerX has yet to be determined (4, 12, 17, 46). Although each nGR recognizes a specific set of germinants, the nGRs (with the exception of GerX) are functionally redundant, as any one of these nGRs is sufficient for wild-type levels of germination (4).

During sporulation, nGR proteins are expressed within the forespore compartment under regulation by the σ^G transcription factor and are inserted into the inner membrane of the developing spore, where they are later able to interact with ligand and initiate germination (15, 16, 36). All three proteins (A, B, and C) are required for activity (27); however, the mechanism of nGR action remains to be elucidated. In order to study the mechanism of action of these receptor proteins, it is important to have a basic understanding of their structure. Although the crystal structure of *B. subtilis* $GerB_C$ is available, the structures of the A- and B-type proteins remain unsolved, likely because they are integral membrane proteins (26). One method to obtain a basic level of structural information about a membrane protein is by determining its topology. Topological analysis gives a simple but powerful understanding of the structure of proteins by identifying intracellular and extracellular domains. This structural analysis is crucial for further investigation of potential ligand binding domains and protein-protein interactions. In this work, topology studies were performed to determine the membrane topology of the *B. anthracis* $GerH_A$, $GerH_B$, and $GerH_C$ proteins in order to provide an initial model of nGR protein structure.

MATERIALS AND METHODS

Bacterial strains and growth conditions. All the work in this study was performed with the *B. anthracis* Sterne $34F_2$ strain ($pXO1^+$ $pXO2^-$). The plasmids used are listed in Table 1. Liquid cultures were grown in brain heart infusion medium (Difco) with chloramphenicol supplemented at 10 μg/ml where necessary. All strains were stored as spore stocks that were prepared as follows. Strains were grown in modified G medium supple-

Received 16 November 2011 **Accepted** 6 December 2011
Published ahead of print 16 December 2011
Address correspondence to Philip C. Hanna, pchanna@umich.edu.
Supplemental material for this article may be found at http://jb.asm.org/.

doi:10.1128/JB.06538-11

TABLE 1 Plasmids used in this work

Plasmid	Relevant characteristics	Source or reference
pAD123	Vector carrying GFPmut3A gene; Cm	10
pMJ01	Promoterless GFP fusion vector derived from pAD123	This work
pGFP	Vector expressing GFPmut3a gene under the control of the *B. subtilis veg* promoter	This work
pFpuA::GFP	pGFP with full-length *fpuA* fused to GFPmut3a gene	This work
pGerH$_C$::GFP	pGFP with full-length *gerH*$_C$ fused to GFPmut3a gene	This work
pGerH$_A$-492::GFP	pGFP with GFPmut3a gene fused to codon 492 of *gerH*$_A$	This work
pGerH$_A$-526::GFP	pGFP with GFPmut3a gene fused to codon 526 of *gerH*$_A$	This work
pGerH$_A$-576::GFP	pGFP with GFPmut3a gene fused to codon 576 of *gerH*$_A$	This work
pGerH$_A$-609::GFP	pGFP with GFPmut3a gene fused to codon 609 of *gerH*$_A$	This work
pGerH$_A$-630::GFP	pGFP with GFPmut3a gene fused to codon 630 of *gerH*$_A$	This work
pGerH$_A$-665::GFP	pGFP with GFPmut3a gene fused to codon 665 of *gerH*$_A$	This work
pGerH$_A$-747::GFP	pGFP with GFPmut3a gene fused to codon 747 of *gerH*$_A$	This work
pGerH$_B$-41::GFP	pGFP with GFPmut3a gene fused to codon 41 of *gerH*$_B$	This work
pGerH$_B$-77::GFP	pGFP with GFPmut3a gene fused to codon 77 of *gerH*$_B$	This work
pGerH$_B$-108::GFP	pGFP with GFPmut3a gene fused to codon 108 of *gerH*$_B$	This work
pGerH$_B$-139::GFP	pGFP with GFPmut3a gene fused to codon 139 of *gerH*$_B$	This work
pGerH$_B$-174::GFP	pGFP with GFPmut3a gene fused to codon 174 of *gerH*$_B$	This work
pGerH$_B$-218::GFP	pGFP with GFPmut3a gene fused to codon 218 of *gerH*$_B$	This work
pGerH$_B$-258::GFP	pGFP with GFPmut3a gene fused to codon 258 of *gerH*$_B$	This work
pGerH$_B$-300::GFP	pGFP with GFPmut3a gene fused to codon 300 of *gerH*$_B$	This work
pGerH$_B$-333::GFP	pGFP with GFPmut3a gene fused to codon 333 of *gerH*$_B$	This work
pGerH$_B$-364::GFP	pGFP with GFPmut3a gene fused to codon 364 of *gerH*$_B$	This work

mented with chloramphenicol at 10 μg/ml for 3 days at 37°C with shaking (22). Spores were prepared as previously described and stored at room temperature in sterile water (38).

Plasmid construction. Plasmid pMJ01, which carries the gene for GFPmut3A (where GFP is green fluorescent protein) with no promoter, was created by modifying the *B. cereus* transcriptional fusion vector pAD123 (10) such that 20 base pairs were removed between the XbaI site and the initiation codon of the GFPmut3A gene and a BglII site was added behind the initiation codon of the GFPmut3A gene. pGFP, which constitutively expresses GFPmut3A during logarithmic growth, was constructed by adding a promoter to pMJ01 to control expression of the GFPmut3A gene (Fig. 1A). This promoter region was constructed in a two-step PCR process. First, the p43 promoter from *B. subtilis* strain PV79 (P_{veg}) was PCR amplified with primers MJWtopo3 and MJWtopo4 (see Table S1 in the supplemental material). MJWtopo3 added a BamHI site upstream of the native P_{veg} sequence. In the second PCR step, primers MJWtopo3 and MJWtopo5 (see Table S1 in the supplemental material) were used to add a Shine-Dalgarno sequence (Fig. 1A, SD) to P_{veg}. The Shine-Dalgarno box was designed to contain an XbaI site directly upstream of the initiation codon. This PCR fragment was then cloned into the BamHI and XbaI sites of pMJ01 to complete pGFP (Fig. 1A). Cloning into the XbaI and BglII sites of this plasmid allows for C-terminal translational fusion of GFP to a target protein (Fig. 1A).

GFP translational fusion vectors were made using the following methods. Gene fragments of interest were PCR amplified with Phusion DNA polymerase (NEB) as per the manufacturer's instructions. Primers were designed such that an XbaI site was introduced before the initiation codon and a BglII site was introduced at the desired site of fusion (see Table S1 in the supplemental material for primer sequences). PCR products were cloned into the corresponding sites of pGFP using standard methods. Cloned fragments were confirmed by sequencing at the University of Michigan DNA Sequencing Core and then passaged through *Escherichia coli* SCS110 cells (*dam dcm*) and introduced into *B. anthracis* using electroporation as described previously (38).

Membrane topology modeling and protein sequence alignment. The amino acid sequences of *B. anthracis* GerH$_A$ (GBAA4984), GerH$_B$ (GBAA4985), and GerH$_C$ (GBAA4986) were downloaded from the JCVI Comprehensive Microbial Resource page (http://cmr.jcvi.org/tigr-scripts/CMR/shared/Genomes.cgi). The membrane topologies of GerH$_A$ and GerH$_B$ were modeled using the most current web versions of 11 different topology methods: SOSUI (http://bp.nuap.nagoya-u.ac.jp/sosui/) (13), SPLIT (http://split.pmfst.hr/split/4/) (19), PHD (http://www.predictprotein.org/) (41), TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/) (23), HMMTOP (http://www.enzim.hu/~tusi/hmmtop/) (43), TMpred (http://www.ch.embnet.org/software/TMPRED_form.html) (14), TopPred (http://mobyle.pasteur.fr/cgi-bin/portal.py?-forms::toppred) (45), TMMOD (http://www.enzim.hu/~tusi/hmmtop/) (20), TSEG (http://www.genome.jp/SIT/tsegdir/) (21), MEMSAT3, and MEMSAT-SVM file://localhost/(http/::bioinf.cs.ucl.ac.uk:psipred:%3Fprogram=svmmemsat) (18, 34). All methods were used in single-protein mode, and all user-adjustable parameters were left at their default values. The amino acid sequence of *B. subtilis* GerA$_A$ (BSU33050) was also obtained from the JCVI Comprehensive Microbial Resource page. Global alignment of GerH$_A$ and GerA$_A$ protein sequences was done using Clustal Omega version 1.0.3 (42), and sequence similarity between these proteins was determined using EMBOSS Stretcher with the EBLOSUM62 protein weight matrix (32). Again, default user settings were used for both alignment servers.

Flow cytometry. *Bacillus anthracis* strains were grown to late log phase (optical density at 600 nm [OD_{600}] = 1.0), centrifuged, washed twice in phosphate-buffered saline (PBS), and fixed in 2% paraformaldehyde for 20 min at room temperature. After fixation, bacteria were washed once with PBS, and approximately 50,000 live cells were analyzed by flow cytometry (FACSCanto; Becton Dickinson). The live cell population was predetermined using a LIVE/DEAD BacLight bacterial viability kit (Molecular Probes). GFP fluorescence emission was collected through a 30-nm band pass filter centered at 530 nm. The fluorescent population (GFP^+) was defined as cells with a fluorescence intensity greater than 95% of that of control bacteria (Fig. 1B, top left panel). Three independent cultures were examined for each strain, and data were analyzed using FlowJo software version 9.2 (Tree Star, Inc.).

Protoplast production. Room temperature overnight cultures were diluted 1:500 in fresh medium and incubated at 37°C to an OD_{600} of 0.1. Bacteria were pelleted, washed twice with PBS–10% sucrose (PBS-S), and resuspended in an equal volume of PBS-S with 2.5 mg/ml lysozyme (Sigma). Samples were incubated at 37°C with gentle shaking and were monitored by phase-contrast microscopy until >90% protoplasts were observed (usually 2 to 4 h). Protoplasts were pelleted by centrifugation at

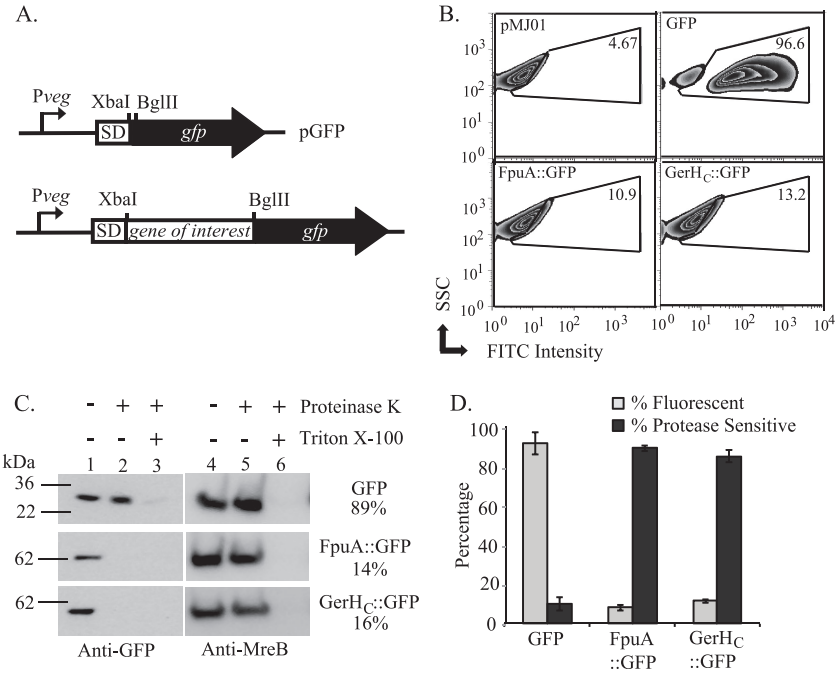

FIG 1 Discrimination between intracellular and extracellular protein domains in *B. anthracis*. (A) Representation of the gene fusion site in the translational fusion vector pGFP. SD, Shine-Dalgarno sequence. (B) Flow cytometric analysis of GFP expression. Approximately 50,000 vegetative bacteria containing the control vector (pMJ01) (upper left) or plasmids expressing GFP alone (upper right), FpuA::GFP (lower left), or $GerH_C$::GFP (lower right) were analyzed. The results are displayed as fluorescein isothiocyanate (FITC) intensity versus side scatter (SSC). The GFP^+ gate was drawn to include cells with fluorescence intensity greater than that of the vector control (upper left). The percentage of GFP^+ cells in each sample is displayed. (C) Protease sensitivity assays. Protoplasts expressing the indicated proteins were treated with 250 μg/ml PK for 5 min, with or without 2% Triton X-100. Immunoblots of samples to evaluate levels of full-length GFP fusion proteins (left) and the intracellular control MreB (right) are shown. Band intensities were quantified using AlphaEaseFC software, and GFP signals were normalized to MreB levels for each sample. Percent protease sensitivity as determined from normalized GFP band intensities of treated and untreated samples are indicated. (D) The means from three independent trials of flow cytometry (gray bars) and protease sensitivity assays (black bars) are shown. Error bars indicate standard errors of the means.

~2,000 × *g* for 5 min, washed once with PBS-S, and resuspended in PBS-S.

Protease sensitivity assays. Samples were split into three groups: untreated (PBS-S), protease treated (250 μg/ml proteinase K [PK] or 200 μg/ml trypsin, as indicated), and protease and Triton treated (the indicated protease with 2% Triton X-100). Samples were incubated at room temperature for 5 min, and phenylmethylsulfonyl fluoride (Sigma) was added to all samples at a final concentration of 1 mM to stop the reactions. A one-third volume of 4× SDS-PAGE buffer supplemented with 8 M urea and 0.4 M dithiothreitol (DTT) was added to each sample and boiled for 5 min. Samples were stored at −20°C.

Western blot analysis. Samples were separated on 4 to 12% bisacrylamide gels (Invitrogen) with MES (morpholineethanesulfonic acid) running buffer (Invitrogen) at 165 V for 60 min. Proteins were transferred to nitrocellulose membranes using semidry transfer at 30 V for 1 h. Membranes were blocked with 5% milk in PBS–0.2% Tween 20 (PBS-T) for 30 min. Primary antibodies were diluted in blocking solution and incubated with membranes overnight at 4°C. For GFP detection, mouse anti-GFP antibodies (Roche) were used at a 1:2,500 dilution in blocking solution. For MreB detection, polyclonal antibodies raised against the *Vibrio cholerae* MreB (a kind gift from M. Sandkvist, University of Michigan) were diluted 1:10,000 in blocking solution. Membranes were washed with PBS-T and incubated for 1 h at room temperature in the appropriate peroxidase-conjugated secondary antibody (Invitrogen) diluted 1:10,000 in blocking solution. Membranes were washed again with PBS-T and developed using enhanced chemiluminescence reagent solution (Amersham). Digital chemiluminescent images were obtained using an Alpha Innotech Fluorchem 8900 (Cell Biosciences). Band intensities were quantified using AlphaEaseFC software (AlphaInnotech).

RESULTS

***Bacillus anthracis* GerH overexpression in vegetative bacilli.** In order to evaluate the membrane topology of the GerH proteins, an expression system was developed in *B. anthracis* vegetative bacteria. Nutrient germinant receptor (nGR) proteins are natively expressed within the forespore compartment during sporulation and inserted into the inner membrane of the spore (16, 36). Since the natural abundance of the nGR proteins in the spore was found to be insufficient for detection in these studies (data not shown), expression of these proteins was uncoupled from sporulation, and the GerH proteins were expressed in the vegetative bacillus. Constitutive overexpression of nGR proteins in the vegetative cell was achieved by construction of an expression system under the control of a modified version of the well-characterized *B. subtilis* P43 promoter (P_{veg}) (Fig. 1A) (30). This version of P_{veg} has had one of its two native RNA polymerase binding sites removed and maintains a medium level of protein expression during logarithmic growth (24). Insertion of proteins into the bacterial membrane occurs by the same mechanisms in both the mother cell and the forespore (31). The membrane insertion machinery, including the Sec translocase and YidC homologs, SpoIIIJ and YqjG, are expressed within the forespore, the mother cell, and the vegetative bacillus (31). In terms of membrane topology of proteins expressed within this system, the cytoplasm can simulate the forespore compartment, while the extracellular space represents the intermembrane space of the developing spore.

GerH$_C$ topology. Information about the membrane topologies of the C-type proteins has not been directly confirmed. GerH$_C$ contains a putative N-terminal signal sequence, suggesting that it is transported to the intermembrane space of the developing forespore, where it is predicted to be anchored to the inner membrane by a lipid moiety (47). Due to these predicted posttranslational modifications, it was hypothesized that GerH$_C$ would be extracellular but membrane associated in the vegetative bacillus overexpression system. In order to confirm the extracellular location of GerH$_C$, translational fusion to green fluorescent protein (GFP) was used. GFP functions as an efficient reporter for analysis of protein topology, since it is fluorescent when located in the cytoplasm and nonfluorescent when transported outside the plasma membrane (11).

A fluorescence-based assay was adapted to evaluate subcellular localization of GFP fusion proteins expressed in vegetative *B. anthracis*. In this assay, vegetative *B. anthracis* bacteria expressing native GFP or various GFP fusion proteins were fixed, and the fluorescence intensities of these bacteria were measured by flow cytometry. By comparing the fluorescence of these samples to that of bacteria containing the promoterless GFP control plasmid, the population of bacteria in each sample exhibiting increased fluorescence over the control was determined (Fig. 1B, top left panel). This population was defined as the GFP$^+$ population. As an intracellular control, native GFP, a cytoplasmic protein, was expressed in vegetative *B. anthracis*. Flow cytometry analysis determined that greater than 90% of bacteria expressing native GFP were GFP$^+$ (Fig. 1B [top right panel] and D [gray bars]). As a control for extracellular localization, GFP was fused to the carboxyl terminus of FpuA (FpuA::GFP). FpuA functions as the extracellular receptor of the siderophore petrobactin and is believed to be located on the extracellular surface of the bacterial plasma membrane (3). In contrast to cells expressing native GFP, only 10% of bacteria were GFP$^+$ when FpuA::GFP was expressed in *B. anthracis* (Fig. 1B [bottom left panel] and D [gray bars]). These data show the ability to distinguish between intracellular and extracellular GFP by scoring via flow cytometry. Next, the assay was used to evaluate the membrane topology of full-length GerH$_C$. A C-terminal GFP fusion to full-length GerH$_C$ (GerH$_C$::GFP) was expressed in vegetative bacteria and evaluated by flow cytometry. Similar to the case for the extracellular control, flow cytometric analysis of bacteria expressing GerH$_C$::GFP found that fewer than 15% of these cells were GFP$^+$ (Fig. 1B [bottom right panel] and D [gray bars]), suggesting external localization.

While fluorescence of a given GFP fusion protein indicates an intracellular localization, the absence of fluorescence, alone, does not demonstrate extracellular localization. Fusion proteins simply could be degraded or misfolded. To confirm extracellular localization of GFP$^-$ fusion constructs, a second assay was used to evaluate the protease sensitivities of fusion proteins. When intact bacteria are exposed to proteases, most extracellular proteins are quickly degraded, while intracellular proteins are protected from degradation, as these enzymes typically do not cross the plasma membrane. The cell walls of vegetative cells overexpressing fusion proteins were removed by lysozyme treatment to allow full exposure of the plasma membrane proteins, and the resulting protoplasts were briefly treated with proteinase K (PK). PK-treated and untreated samples were subjected to immunoblotting with anti-GFP antibody to assess whether full-length GFP fusion proteins were sensitive to PK degradation. Immunoblots were performed with samples expressing GFP, FpuA::GFP, or GerH$_C$::GFP (Fig. 1C). Untreated protoplasts from all three samples contained full-length proteins (Fig. 1C, lane 1). After 5 min of incubation with PK, the intracellular control, GFP, remained protected from PK degradation, while degradation of the extracellular FpuA::GFP and GerH$_C$::GFP was apparent (Fig. 1C, lane 2). Lysed protoplasts treated with PK were analyzed to identify any protease-resistant protein species (Fig. 1C, lane 3). To quantify PK sensitivity, GFP band intensities of the treated (Fig. 1C, lane 2) and untreated (Fig. 1C, lane 1) samples were determined. These intensities were then normalized to levels of the known intracellular protein, MreB, to account for any unexpected protoplast lysis (Fig. 1C, lanes 4 and 5). The percentage of each GFP fusion protein that was sensitive to degradation by PK was calculated using these normalized values (Fig. 1C and D [black bars]). Collectively, these assays confirmed that the C terminus of GerH$_C$::GFP was extracellular, as predicted, and that subcellular GFP localization of proteins expressed in *B. anthracis* could effectively be distinguished by these assays.

TABLE 2 Putative transmembrane domains of the *B. anthracis* GerH$_B$ protein

Putative TM	No. of programs predicting TM	Length (amino acids)	Avg starting point	Avg ending point
1	11/11	22	12 (±4)	33 (±6)
2	11/11	22	44 (±3)	65 (±3)
3	11/11	24	84 (±3)	107 (±9)
4	10/11	22	116 (±5)	137 (±5)
5	11/11	23	143 (±5)	165 (±5)
6	9/11	19	190 (±6)	208 (±2)
7	11/11	21	220 (±4)	240 (±3)
8[a]	2/11	20	245 (±1)	264 (±3)
9	11/11	25	267 (±8)	291 (±2)
10	11/11	19	303 (±4)	321 (±4)
11	11/11	21	334 (±3)	354 (±2)

[a] This helix was excluded from GFP fusion analysis (see text).

GerH$_B$ topology. To develop a hypothetical model of the membrane topology of the putative integral membrane protein GerH$_B$, topology prediction algorithms were utilized. Previous topology analyses of proteins from *E. coli* have shown that the reliability of a predicted topology increases when multiple prediction methods agree (9). Here, 11 different prediction programs (see Materials and Methods) were used to map the membrane topology of GerH$_B$. The majority of prediction programs (8 of 11) determined a 10-transmembrane-domain (TM) topology for GerH$_B$ with both the N and C termini located intracellularly (see Table S2 in the supplemental material). Two programs predicted 9 TMs and a single program predicted 11 TMs for this protein (see Table S2 in the supplemental material). By comparing the results of these programs, 11 putative TMs were identified (Table 2). Eight of the 11 putative TMs were predicted by all prediction programs (Table 2). In contrast, TM4 was predicted by 10 prediction methods, TM6 by 9 prediction methods, and TM8 by only 2 prediction methods (Table 2). TM8 was excluded from analysis because of its low predictive value.

Based on these results, 10 locations within GerH$_B$ were selected to be analyzed by GFP translational fusion. C-terminal GFP translational fusions were constructed with full-length GerH$_B$, as well as nine truncated GerH$_B$ proteins. The locations of the GerH$_B$

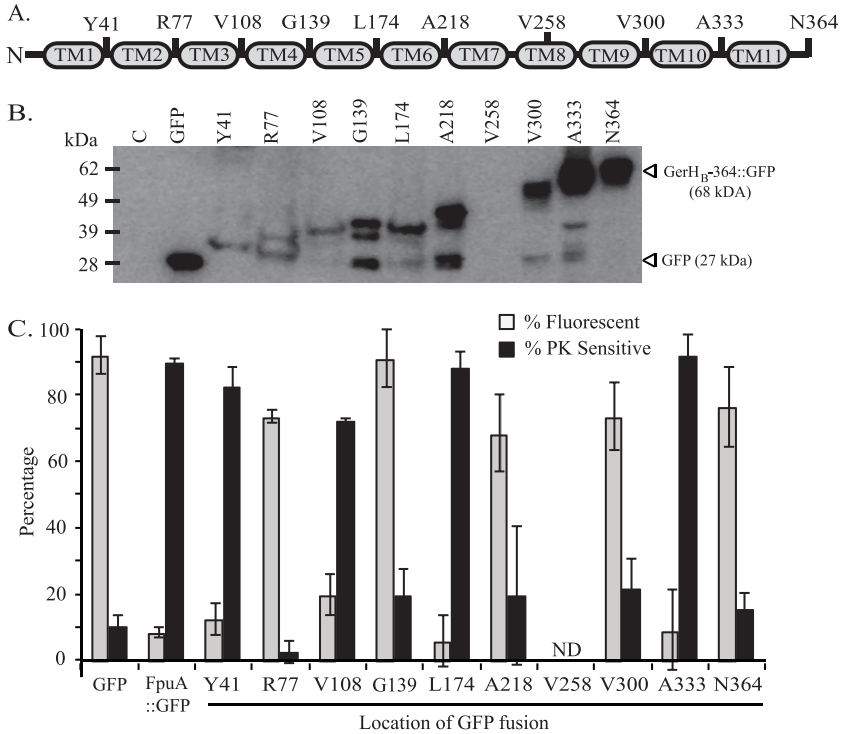

FIG 2 Analysis of $GerH_B$ topology. (A) Model of $GerH_B$ with 11 putative transmembrane domains as determined by topology prediction. To test these putative TMs, C-terminal GFP fusions were made to full-length $GerH_B$ (N364) and nine $GerH_B$ proteins truncated at the indicated locations. TM8 was excluded because of its low predictive value. (B) Expression of $GerH_B$ fusion proteins in *B. anthracis*. Whole-cell lysates from bacteria expressing a vector control (lane C), GFP, or $GerH_B$ fusion proteins were analyzed by immunoblotting with anti-GFP antibodies to assess protein expression and stability. The predicted sizes of GFP and full-length $GerH_B$-364::GFP are indicated. (C) Analysis of $GerH_B$ fusion proteins by flow cytometry and protease sensitivity assays. Flow cytometry was used to measure the fluorescence intensities of bacteria expressing control proteins (GFP or FpuA::GFP) or the $GerH_B$ fusion proteins. The percentage of GFP^+ cells in each sample was determined as described for Fig. 1B (gray bars). Protease sensitivity assays were performed with proteinase K. The percentage of each fusion protein that was protease sensitive was determined as described for Fig. 1C (black bars). Each bar represents the mean ± standard error of the mean. ND, not determined.

truncations and GFP fusion sites are depicted in Fig. 2A. Expression levels of these proteins are shown by Western blotting with anti-GFP antibodies (Fig. 2B). $GerH_B$-258::GFP, which was located between predicted TM7 and TM8, was not stable in this system and was excluded from further analysis (Fig. 2B). The nine remaining fusion proteins were stable when expressed in the vegetative bacillus, though overall protein stability was variable between these truncated protein species (Fig. 2B). There was some concern that less-stable GFP fusion proteins would exhibit minimal fluorescence. Bacteria expressing less-stable GFP fusion proteins exhibited decreased average fluorescence intensity when analyzed by flow cytometry, but the percentage of these bacteria with fluorescence intensity greater than that of the vector control was similar to that for bacteria expressing more stable fusions (see Fig. S1A in the supplemental material). Therefore, the nine stable $GerH_B$ fusion proteins were analyzed by both flow cytometry and protease sensitivity assay.

Analysis of the $GerH_B$ fusion proteins revealed two distinct groups. Strains expressing $GerH_B$-77::GFP, $GerH_B$-139::GFP, $GerH_B$-218::GFP, $GerH_B$-300::GFP, and $GerH_B$-364::GFP had increased fluorescence similar to that of the GFP control sample, with greater than 50% of cells being GFP^+ (Fig. 2C, gray bars; see Fig. S1A in the supplemental material). These five fusion proteins were also insensitive to PK degradation (Fig. 2C, black bars; see Fig. S1B in the supplemental material). Together, these results are consistent with an intracellular localization. The four remaining strains, expressing $GerH_B$-41::GFP, $GerH_B$-108::GFP, $GerH_B$-174::GFP, and $GerH_B$-333::GFP, were significantly less fluorescent ($P < 0.01$) than the intracellular GFP control sample (Fig. 2C, gray bars; see Fig. S1A in the supplemental material). These four fusion proteins were also significantly more sensitive to PK degradation ($P < 0.05$) (Fig. 2C, black bars; see Fig. S1B in the supplemental material). Decreased fluorescence and increased protease sensitivity are consistent with an extracellular location of these four GFP fusion sites. Based on these data, we can conclude that there are five intracellular loops and at least four extracellular loops of $GerH_B$.

These data also provided information about the putative TMs of the protein. If a putative TM was actually a membrane-spanning helix, GFP fusions immediately upstream and downstream of the predicted TM would have opposite GFP phenotypes, indicating that these fusion sites were located on opposite sides of the membrane. The data strongly suggested that at least 8 of the 10 putative TMs (Fig. 2A) were indeed membrane-spanning regions of the protein (TM1, TM2, TM3, TM4, TM5, TM6, TM9 and TM10), as GFP fusions before and after these regions were determined to be located on opposite sides of the membrane (Fig. 2C). The instability of $GerH_B$-258::GFP led to incomplete data about predicted TM7 and TM8 and did not allow us to effectively distinguish between 8- and 10-TM models.

$GerH_A$ topology. $GerH_A$ was also predicted to be an integral membrane protein. The same 11 topology prediction programs

TABLE 3 Putative transmembrane domains of the *B. anthracis* $GerH_A$ protein

Putative TM	No. of programs predicting TM	Length (amino acids)	Avg starting point	Avg ending point
1	4/11	19	496 (±1)	514 (±4)
2	11/11	22	534 (±4)	555 (±2)
3	4/11	20	578 (±4)	597 (±2)
4	8/11	20	606 (±4)	625 (±5)
5	11/11	22	628 (±4)	649 (±6)
6	11/11	25	661 (±5)	685 (±3)

were used to determine a hypothetical model of $GerH_A$ membrane topology (see Materials and Methods). Unlike $GerH_B$, in which TMs were predicted to occur throughout the length of the protein, $GerH_A$ was predicted to contain two distinct regions: a C-terminal membrane-spanning domain and a large hydrophilic N-terminal domain. The majority of the programs (8 of 11) predicted a four-TM topology for $GerH_A$; however, one program predicted five TMs and two predicted six TMs. The predicted locations of the N and C termini varied, although the majority of the programs agreed that the hydrophilic N terminus was cytoplasmic (see Table S3 in the supplemental material). Additionally, though eight programs predicted a four-TM topology, the locations of the predicted domains were variable (see Table S3 in the supplemental material). By combining the results of these prediction programs, six putative TMs were identified (Table 3).

Based on these predictions, seven locations within $GerH_A$ were selected to be analyzed by GFP translational fusion. C-terminal GFP fusions were constructed with full-length $GerH_A$ and six truncated $GerH_A$ proteins. The locations of these truncations and GFP fusion sites are depicted in Fig. 3A. Expression levels of these fusion proteins were evaluated by immunoblotting with anti-GFP antibodies (Fig. 3B). Despite differences in overall protein stability, all seven fusion proteins were deemed suitable for analysis (Fig. 3B).

Both flow cytometry and protease protection assays were used to analyze the subcellular localization of the seven $GerH_A$ fusion sites. Flow cytometry identified four strains, expressing $GerH_A$-492::GFP, $GerH_A$-526::GFP, $GerH_A$-609::GFP, and $GerH_A$-747::GFP, with greater than 60% of the bacteria GFP^+ (Fig. 3C, gray bars; see Fig. S2A in the supplemental material). As expected, these four fusion proteins were also insensitive to PK degradation, consistent with an intracellular location (Fig. 3C, black bars; see Fig. S2B in the supplemental material). Analysis of strains expressing fusion proteins $GerH_A$-630::GFP and $GerH_A$-665::GFP found that these strains had significantly lower fluorescence ($P < 0.01$) than the intracellular GFP control (Fig. 3C, gray bars; see Fig. S2A in the supplemental material). These two fusion proteins were also significantly more sensitive to PK degradation ($P < 0.5$) (Fig. 3C, black bars; see Fig. S2B in the supplemental material), consistent with an extracellular localization.

One fusion protein, $GerH_A$-576::GFP, gave inconsistent results. Bacteria expressing this protein exhibited decreased fluorescence when analyzed by flow cytometry (Fig. 3C, gray bars; see Fig. S2A in the supplemental material), consistent with an extracellular localization; however, this fusion protein was also resistant to PK treatment (Fig. 3C, black bars; see Fig. S2B in the supplemental material), consistent with an intracellular localization. In order to address these contradictory results, an additional analysis was performed to determine if $GerH_A$-576::GFP was resistant to PK for reasons besides intracellular localization. Instead of using PK, protoplasts containing $GerH_A$-576::GFP were treated with 200 μg/ml trypsin for 5 min. In contrast to the PK results, 70% of $GerH_A$-576::GFP was sensitive to trypsin degradation (Fig. 3C, hatched bar; see Fig. S2C in the supplemental material), while the intracellular MreB was protected (data not shown). This suggested that the GFP fusion to $GerH_A$ at Q576 was indeed extracellular but was resistant to PK for reasons other than intracellular

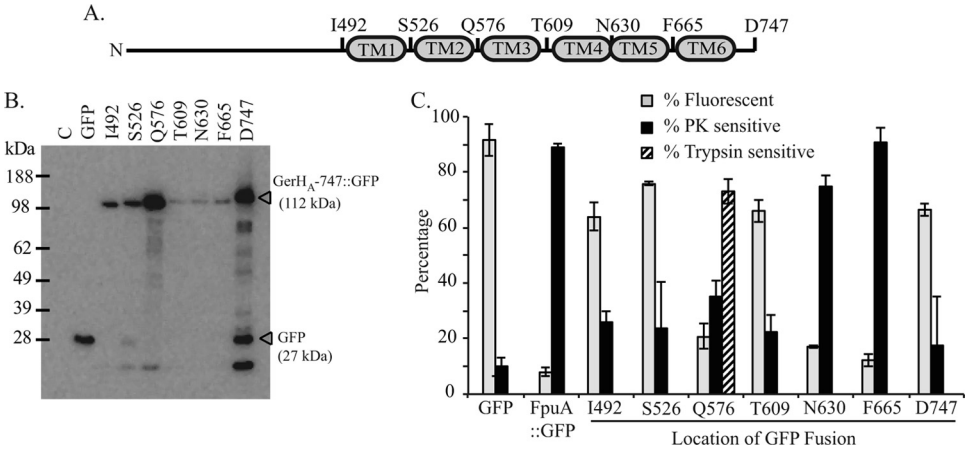

FIG 3 Analysis of $GerH_A$ topology. (A) Model of $GerH_A$ with six putative transmembrane domains as determined by topology prediction. To test these putative TMs, C-terminal GFP fusions were made to full-length $GerH_A$ (D747) and six $GerH_A$ proteins truncated at the indicated locations. (B) Expression of $GerH_A$ fusion proteins in *B. anthracis*. Whole-cell lysates from bacteria expressing a vector control (lane C), GFP, or $GerH_A$ fusion proteins were analyzed by immunoblotting with anti-GFP antibodies to assess protein expression and stability. The predicted sizes of GFP and full-length $GerH_A$-747::GFP are indicated. (C) Analysis of $GerH_A$ fusion proteins by flow cytometry and protease sensitivity assays. Flow cytometry was used to measure the fluorescence intensities of bacteria expressing control proteins (GFP or FpuA::GFP) or the $GerH_A$ fusion proteins. The percentage of GFP^+ cells in each sample was determined as described for Fig. 1B (gray bars). Protease sensitivity assays were performed with proteinase K (black bars) or trypsin (hatched bar). The percentage of each fusion protein that was protease sensitive was determined as described for Fig. 1C. Bars represent the means of three independent trials ± standard errors of the means.

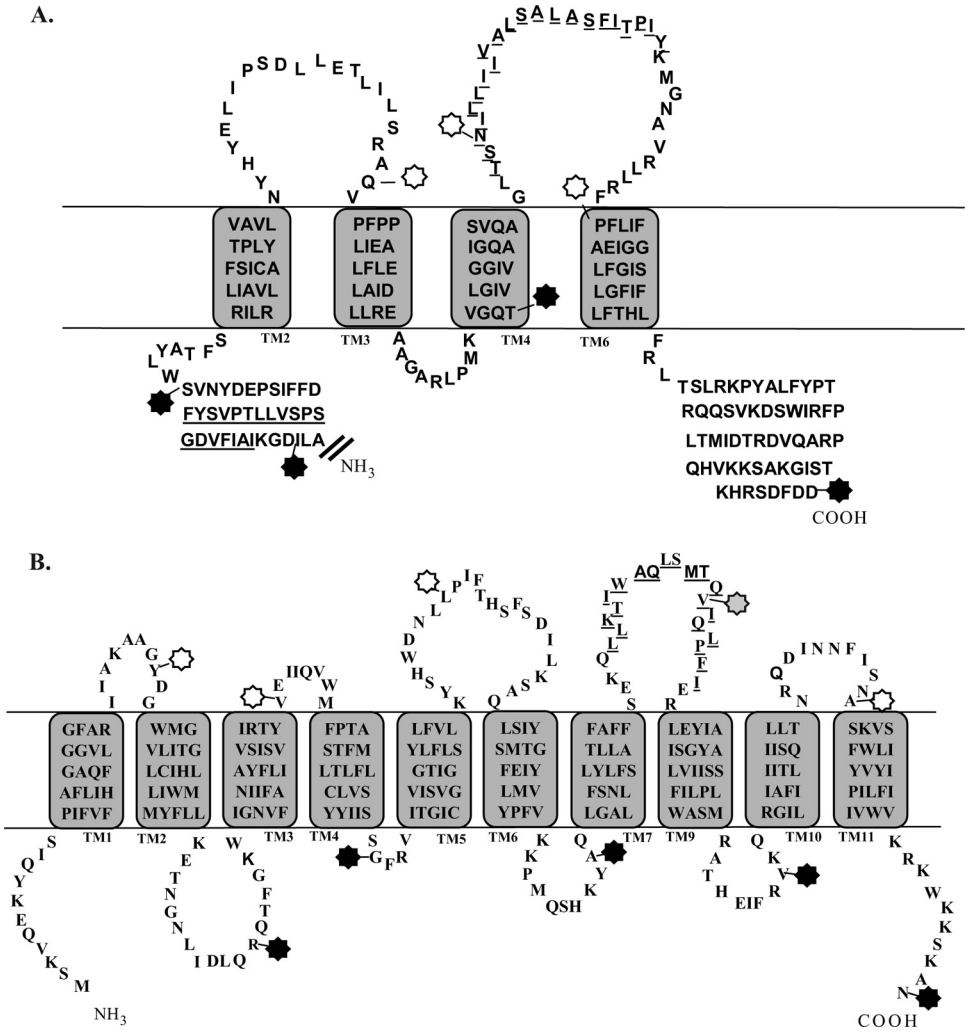

FIG 4 Topology models of *B. anthracis* $GerH_A$ and $GerH_B$. Topology models of the membrane-spanning regions (amino acids 490 to 747) of $GerH_A$ (A) and full-length $GerH_B$ (B) are based on a combination of experimental results and topology modeling. The locations of GFP fusions used in this study are marked by stars. Black stars represent fusion locations that were determined to be intracellular, while white stars represent extracellular fusion locations. The gray star marks the unstable $GerH_B$ fusion, which was unable to be analyzed. Membrane-spanning sequences were determined based on consensus sequences from 11 topology prediction algorithms. The actual membrane-spanning sequences of the proteins may vary from these predicted domains. Underlined amino acid residues indicate predicted transmembrane domains that were not supported by GFP fusion analysis.

localization. Collectively, these data demonstrated that four of the seven $GerH_A$ fusion sites were intracellular, while the remaining three fusion sites were found to be extracellular.

As before, because each fusion was located upstream or downstream of a putative TM (Fig. 3A), these data also provided information on which putative TMs were actual membrane-spanning regions of the protein. Four of the six predicted TMs of $GerH_A$ (TM2, TM3, TM4, and TM6) were confirmed to be actual membrane-spanning regions, as fusions before and after these domains were determined to be on opposite sides of the membrane (Fig. 3C).

DISCUSSION

The role of the *B. subtilis* GerA germinant receptor in triggering L-alanine-based germination was first realized over 30 years ago (28). Since that time, significant genetic analyses of GerA and its homologs have elucidated much information about how germinant receptors function as a whole, but many questions remain about the specific roles of the individual subunits in recognizing germinants and triggering germination. In order to proceed with in-depth functional analyses of these proteins, a basic understanding of protein structure is required. Previous attempts to overexpress nGR proteins of the GerA receptor within the context of the spore were unsuccessful, as overexpression of these proteins leads to arrest of spore formation at stage III of sporulation (2). In this work, nGR protein expression was uncoupled from sporulation, which led to the successful overexpression of the GerH proteins in the vegetative bacillus. Using this system, the membrane topologies of the *B. anthracis* $GerH_A$, $GerH_B$, and $GerH_C$ proteins were mapped, providing the first structural data on the A and B subunits of the nGRs. Although only the GerH protein topology was analyzed, the results are likely applicable to other nGR family members, as these proteins are predicted to have similar structures (4, 6, 26, 29).

Based on the results of this study, topology models of the integral membrane proteins $GerH_A$ (Fig. 4A) and $GerH_B$ (Fig. 4B) can

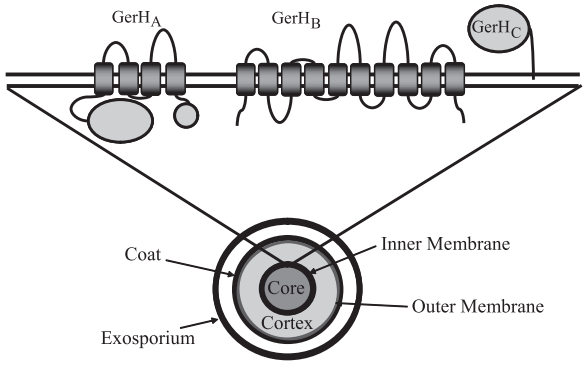

FIG 5 Current model of the nutrient germinant receptors in the inner membrane of the spore. The traditional model of nGR structure (27) was adapted to represent the data presented in this work. A representation of the *Bacillus anthracis* spore with the GerH nGR proteins modeled in the inner membrane is shown. This model is consistent with the traditional model of the B and C nGR subunits. However, the current model of GerH$_A$ differs somewhat from the traditional model of the A subunit. Here GerH$_A$ contains four transmembrane domains, while the traditional model contained five. This difference moves the C terminus of GerH$_A$ to the inside of the spore core. (Reprinted with kind permission from Springer Science+Business Media: Cell. Mol. Life Sci., Spore germination, vol. 59, 2002, p 403–409, A Moir, BM Corfe, J Behravan, Fig. 1 [27].)

be produced. GFP fusion analysis of GerH$_A$ found four of the six predicted TMs of this protein to be actual membrane-spanning regions. Our model of GerH$_A$ contains these four transmembrane domains with two extracellular loops and large intracellular domains at both termini (Fig. 4A). Interestingly, only one prediction program, MEMSAT-SVM, was consistent with the experimental GerH$_A$ data. The fusion protein data for GerH$_B$ were not able to distinguish between an 8- or 10-TM model, because the fusion designed between predicted domains TM7 and TM8 was not stable enough for analysis. Eight of the 11 topology prediction programs predicted a 10-TM topology for GerH$_B$, while none predicted an 8-TM topology. Here, GerH$_B$ is modeled containing 10 TMs with both N and C termini intracellular (Fig. 4B), though an eight-TM topology cannot be completely ruled out based on the data presented.

The results of this study also allow for the GerH nGR proteins to be modeled in the inner membrane of the spore (Fig. 5). The traditional model of the nGR proteins is based on topology prediction algorithms with *B. subtilis* GerA proteins (27). It is of note that the results presented here are mostly consistent with this model. Only GerH$_A$ was found to deviate from the traditional model of nGR topology. In this work, the C terminus of GerH$_C$ was determined to be located extracellularly, consistent with the hypothesis that this protein is transported to the intermembrane space and anchored to the inner membrane of the spore (Fig. 5). Interestingly, though the C subunit is entirely extracellular, it is the only nGR subunit that has never been implicated in germinant specificity (25). This protein may interact with other proteins located in the cortex to mediate signal transduction after germinant binding or play a role in localizing the nGR complex. Further studies are required to understand the role of this protein in germination. It is clear, however, that any potential interactions between the C-type subunits and the A- and B-type subunits must occur outside the spore core. GerH$_A$ and GerH$_B$ are modeled in the inner membrane with 4 and 10 TMs, respectively, and both termini located in the spore core (Fig. 5). Both A- and B-type nGR subunits have been associated with germinant recognition (5, 6, 29, 35). Site-directed mutagenesis of B-type nGR proteins in *Bacillus subtilis* and *Bacillus megaterium* has identified residues in the membrane-spanning regions of these proteins that are potentially involved in ligand binding (5, 6). Both GerH$_A$ and GerH$_B$ contain membrane-spanning regions that could be available for interactions with nutrient germinants. Additionally, though the A and B subunits are likely to interact, the interaction domains of these proteins have yet to be elucidated (15, 44). This model illuminates multiple domains within the A and B subunits that could potentially interact, either within the membrane or within the spore core (Fig. 5).

The current model of GerH$_A$ topology differs from the traditional five-transmembrane-domain model of the A-type proteins (27). This new understanding of the structure of the A-type proteins can allow for reevaluation of previous studies. For example, mutational analysis of *B. subtilis* GerA$_A$ identified two interesting regions of this protein: the highly conserved PFPP domain located at amino acids 323 to 326 (see Fig. S1, box 1, in the supplemental material) and amino acid residues L373, G398, L399, and S400 located in the extracellular loop between the third and fourth TMs (see Fig. S1, boxes 2 and 3, in the supplemental material) (29). Mutations in the PFPP domain led to a variety of phenotypes, including increased sensitivity to germinant and a sporulation defect that resulted in phase-dark spores. A similar mutation of *B. subtilis* GerB$_A$ at P326 resulted in the novel ability of GerB to respond to L-alanine or L-asparagine in the absence of GerK, which is normally required for GerB activity (1, 35). According to the model presented here, the PFPP domain lies within the beginning of the second transmembrane domain on the extracellular side of the membrane (see Fig. S1, box 1, in the supplemental material). Though a single proline is often found at the beginnings of alpha-helices, proline residues located within a transmembrane helix can lead to kinking or breaking of these helices (7). Since changes to this structure lead to multiple methods of dysregulation of germination, the presence of this conserved proline-dense region within the transmembrane domain may indicate an important regulatory domain of the A protein that requires the specific PFPP structure to properly respond to germinant. Point mutations within the second interesting region of GerA$_A$ at amino acids L373, G398, L399, and S400 resulted in germination deficiencies but did not affect GerA complex formation, as measured by the presence of GerA$_C$ in spore extracts (29). This region could potentially play a role in signal transduction or germinant binding (see Fig. S1, boxes 2 and 3, in the supplemental material).

The system presented in this work is the first that successfully expresses nGR proteins to levels high enough for biochemical analysis. The system can be exploited for further analyses that require large amounts of these proteins. This and our new understanding of the orientation of the A and B subunits in the inner membrane of the spore can allow for exploring functional domains of the nGR proteins. Protein-protein interactions between nGR subunits and the domains responsible for these interactions can be evaluated in this overexpression system. Additionally, interactions with other putative binding partners, such as GerD, can also be assessed outside the spore environment.

ACKNOWLEDGMENTS

This work was supported in part by grant AI092024 from the National Institutes of Health. This research was performed while Mary Wilson was under an appointment to the U.S. Department of Homeland Security (DHS) Scholarship and Fellowship Program. Paul Carlson is supported by grant UL1RR024986 from the National Center for Research Resources.

We thank Suzanne Lybarger and members of the Maria Sandkvist lab (University of Michigan) for their gift of the rabbit polyclonal antibody against the *Vibrio cholerae* MreB protein.

REFERENCES

1. **Atluri S, Ragkousi K, Cortezzo DE, Setlow P.** 2006. Cooperativity between different nutrient receptors in germination of spores of Bacillus subtilis and reduction of this cooperativity by alterations in the GerB receptor. J. Bacteriol. **188**:28–36.
2. **Cabrera-Martinez RM, Tovar-Rojo F, Vepachedu VR, Setlow P.** 2003. Effects of overexpression of nutrient receptors on germination of spores of Bacillus subtilis. J. Bacteriol. **185**:2457–2464.
3. **Carlson PE, Jr, et al.** 2010. Genetic analysis of petrobactin transport in Bacillus anthracis. Mol. Microbiol. **75**:900–909.
4. **Carr KA, Lybarger SR, Anderson EC, Janes BK, Hanna PC.** 2010. The role of Bacillus anthracis germinant receptors in germination and virulence. Mol. Microbiol. **75**:365–375.
5. **Christie G, Gotzke H, Lowe CR.** 2010. Identification of a receptor subunit and putative ligand-binding residues involved in the Bacillus megaterium QM B1551 spore germination response to glucose. J. Bacteriol. **192**:4317–4326.
6. **Cooper GR, Moir A.** 2011. Amino acid residues in the GerAB protein important in the function and assembly of the alanine spore germination receptor of Bacillus subtilis 168. J. Bacteriol. **193**:2261–2267.
7. **Cordes FS, Bright JN, Sansom MS.** 2002. Proline-induced distortions of transmembrane helices. J. Mol. Biol. **323**:951–960.
8. **Dixon TC, Meselson M, Guillemin J, Hanna PC.** 1999. Anthrax. N. Engl. J. Med. **341**:815–826.
9. **Drew D, et al.** 2002. Rapid topology mapping of Escherichia coli inner-membrane proteins by prediction and PhoA/GFP fusion analysis. Proc. Natl. Acad. Sci. U. S. A. **99**:2690–2695.
10. **Dunn AK, Handelsman J.** 1999. A vector for promoter trapping in Bacillus cereus. Gene **226**:297–305.
11. **Feilmeier BJ, Iseminger G, Schroeder D, Webber H, Phillips GJ.** 2000. Green fluorescent protein functions as a reporter for protein localization in Escherichia coli. J. Bacteriol. **182**:4068–4076.
12. **Fisher N, Hanna P.** 2005. Characterization of Bacillus anthracis germinant receptors in vitro. J. Bacteriol. **187**:8055–8062.
13. **Hirokawa T, Boon-Chieng S, Mitaku S.** 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics **14**:378–379.
14. **Hofmann K, Stoffel W.** 1993. TMbase—a database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler **374**:166.
15. **Hudson KD, et al.** 2001. Localization of GerAA and GerAC germination proteins in the Bacillus subtilis spore. J. Bacteriol. **183**:4317–4322.
16. **Igarashi T, Setlow P.** 2006. Transcription of the Bacillus subtilis gerK operon, which encodes a spore germinant receptor, and comparison with that of operons encoding other germinant receptors. J. Bacteriol. **188**:4131–4136.
17. **Ireland JA, Hanna PC.** 2002. Amino acid- and purine ribonucleoside-induced germination of Bacillus anthracis DeltaSterne endospores: gerS mediates responses to aromatic ring structures. J. Bacteriol. **184**:1296–1303.
18. **Jones DT.** 2007. Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics **23**:538–544.
19. **Juretic D, Zoranic L, Zucic D.** 2002. Basic charge clusters and predictions of membrane protein topology. J. Chem. Infect. Comput. Sci. **42**:620–632.
20. **Kahsay RY, Gao G, Liao L.** 2005. An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. Bioinformatics **21**:1853–1858.
21. **Kihara D, Shimizu T, Kanehisa M.** 1998. Prediction of membrane proteins based on classification of transmembrane segments. Protein Eng. **11**:961–970.
22. **Kim HU, Goepfert JM.** 1974. A sporulation medium for Bacillus anthracis. J. Appl. Bacteriol. **37**:265–267.
23. **Krogh A, Larsson B, von Heijne G, Sonnhammer EL.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. **305**:567–580.
24. **Le Grice SF, Sonenshein AL.** 1982. Interaction of Bacillus subtilis RNA polymerase with a chromosomal promoter. J. Mol. Biol. **162**:551–564.
25. **Li Y, et al.** 2011. Structure-based functional studies of the effects of amino acid substitutions in GerBC, the C subunit of the Bacillus subtilis GerB spore germinant receptor. J. Bacteriol. **193**:4143–4152.
26. **Li Y, Setlow B, Setlow P, Hao B.** 2010. Crystal structure of the GerBC component of a Bacillus subtilis spore germinant receptor. J. Mol. Biol. **402**:8–16.
27. **Moir A, Corfe BM, Behravan J.** 2002. Spore germination. Cell. Mol. Life Sci. **59**:403–409.
28. **Moir A, Lafferty E, Smith DA.** 1979. Genetics analysis of spore germination mutants of Bacillus subtilis 168: the correlation of phenotype with map location. J. Gen. Microbiol. **111**:165–180.
29. **Mongkolthanaruk W, Cooper GR, Mawer JS, Allan RN, Moir A.** 2011. Effect of amino acid substitutions in the GerAA protein on the function of the alanine-responsive germinant receptor of Bacillus subtilis spores. J. Bacteriol. **193**:2268–2275.
30. **Moran CP, Jr, et al.** 1982. Nucleotide sequences that signal the initiation of transcription and translation in Bacillus subtilis. Mol. Gen. Genet. **186**:339–346.
31. **Murakami T, Haga K, Takeuchi M, Sato T.** 2002. Analysis of the Bacillus subtilis spoIIIJ gene and its paralogue gene, yqjG. J. Bacteriol. **184**:1998–2004.
32. **Myers E, Miller W.** 1988. Optimal alignments in linear space. Comput. Appl. Biosci. (Camb.) **4**:11–17.
33. **Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P.** 2000. Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. Microbiol. Mol. Biol. Rev. **64**:548–572.
34. **Nugent T, Jones DT.** 2009. Transmembrane protein topology prediction using support vector machines. BMC Bioinformatics **10**:159.
35. **Paidhungat M, Setlow P.** 1999. Isolation and characterization of mutations in Bacillus subtilis that allow spore germination in the novel germinant D-alanine. J. Bacteriol. **181**:3341–3350.
36. **Paidhungat M, Setlow P.** 2001. Localization of a germinant receptor protein (GerBA) to the inner membrane of Bacillus subtilis spores. J. Bacteriol. **183**:3982–3990.
37. **Paidhungat M, Setlow P.** 2000. Role of Ger proteins in nutrient and nonnutrient triggering of spore germination in Bacillus subtilis. J. Bacteriol. **182**:2513–2519.
38. **Passalacqua KD, Bergman NH, Herring-Palmer A, Hanna P.** 2006. The superoxide dismutases of Bacillus anthracis do not cooperatively protect against endogenous superoxide stress. J. Bacteriol. **188**:3837–3848.
39. **Piggot PJ, Hilbert DW.** 2004. Sporulation of Bacillus subtilis. Curr. Opin. Microbiol. **7**:579–586.
40. **Read TD, et al.** 2003. The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria. Nature **423**:81–86.
41. **Rost B, Fariselli P, Casadio R.** 1996. Topology prediction for helical transmembrane proteins at 86% accuracy. Protein Sci. **5**:1704–1718.
42. **Sievers F, et al.** 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. **7**:539.
43. **Tusnady GE, Simon I.** 1998. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J. Mol. Biol. **283**:489–506.
44. **Vepachedu VR, Setlow P.** 2007. Analysis of interactions between nutrient germinant receptors and SpoVA proteins of Bacillus subtilis spores. FEMS Microbiol. Lett. **274**:42–47.
45. **von Heijne G.** 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J. Mol. Biol. **225**:487–494.
46. **Weiner MA, Read TD, Hanna PC.** 2003. Identification and characterization of the gerH operon of Bacillus anthracis endospores: a differential role for purine nucleosides in germination. J. Bacteriol. **185**:1462–1464.
47. **Zuberi AR, Moir A, Feavers IM.** 1987. The nucleotide sequence and gene organization of the gerA spore germination operon of Bacillus subtilis 168. Gene **51**:1–11.