Eubacterial SpoVG homologs constitute a new family of site-specific DNA-binding proteins

Brandon L. Jutras  
*University of Kentucky, blju222@uky.edu*

Alicia M. Chenail  
*University of Kentucky, alicia.chenail@uky.edu*

Christi L. Rowland  
*University of Kentucky, clrowland@uky.edu*

Dustin Carroll  
*University of Kentucky, dwhittakerc@uky.edu*

M Clarke Miller  
*University of Louisville*

*See next page for additional authors*

**Click here to let us know how access to this document benefits you.**

Follow this and additional works at:  [https://uknowledge.uky.edu/microbio_facpub](https://uknowledge.uky.edu/microbio_facpub)

Part of the [Medical Microbiology Commons](https://uknowledge.uky.edu/microbio_facpub)

**Repository Citation**

Jutras, Brandon L.; Chenail, Alicia M.; Rowland, Christi L.; Carroll, Dustin; Miller, M Clarke; Bykowski, Tomasz; and Stevenson, Brian, "Eubacterial SpoVG homologs constitute a new family of site-specific DNA-binding proteins" (2013). *Microbiology, Immunology, and Molecular Genetics Faculty Publications*. 5.  
[https://uknowledge.uky.edu/microbio_facpub/5](https://uknowledge.uky.edu/microbio_facpub/5)

This Article is brought to you for free and open access by the Microbiology, Immunology, and Molecular Genetics at UKnowledge. It has been accepted for inclusion in Microbiology, Immunology, and Molecular Genetics Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
Authors
Brandon L. Jutras, Alicia M. Chenail, Christi L. Rowland, Dustin Carroll, M Clarke Miller, Tomasz Bykowski, and Brian Stevenson

Eubacterial SpoVG homologs constitute a new family of site-specific DNA-binding proteins

Notes/Citation Information
Published in PLoS ONE, v. 8, no. 6, e66683.

© 2013 Jutras et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Digital Object Identifier (DOI)
http://dx.doi.org/10.1371/journal.pone.0066683

This article is available at UKnowledge: https://uknowledge.uky.edu/microbio_facpub/5
Eubacterial SpoVG Homologs Constitute a New Family of Site-Specific DNA-Binding Proteins

Brandon L. Jutras1, Alicia M. Chenail1, Christi L. Rowland2, Dustin Carroll3, M. Clarke Miller4, Tomasz Bykowski1*, Brian Stevenson1*

1 Department of Microbiology, Immunology, and Molecular Genetics, University of Kentucky College of Medicine, Lexington, Kentucky, United States of America, 2 Department of Agricultural Sciences, University of Kentucky College of Agriculture, Lexington, Kentucky, United States of America, 3 Graduate Center for Toxicology, University of Kentucky College of Medicine, Lexington, Kentucky, United States of America, 4 Brown Cancer Center, University of Louisville, Louisville, Kentucky, United States of America

Abstract

A site-specific DNA-binding protein was purified from Borrelia burgdorferi cytoplasmic extracts, and determined to be a member of the highly conserved SpoVG family. This is the first time a function has been attributed to any of these ubiquitous bacterial proteins. Further investigations into SpoVG orthologues indicated that the Staphylococcus aureus protein also binds DNA, but interacts preferentially with a distinct nucleic acid sequence. Site-directed mutagenesis and domain swapping between the S. aureus and B. burgdorferi proteins identified that a 6-residue stretch of the SpoVG α-helix contributes to DNA sequence specificity. Two additional, highly conserved amino acid residues on an adjacent β-sheet are essential for DNA-binding, apparently by contacts with the DNA phosphate backbone. Results of these studies thus identified a novel family of bacterial DNA-binding proteins, developed a model of SpoVG-DNA interactions, and provide direction for future functional studies on these wide-spread proteins.

Introduction

To be successful, single-celled organisms must efficiently and rapidly adapt to changing conditions. This is often accomplished through exquisite regulatory networks involving numerous, dynamic trans-acting factors. Prokaryotic proteins that bind to nucleic acids govern virtually every cellular process, including nucleoid organization, transcription, translation, and DNA replication, modification, repair, and recombination. Remarkably, most DNA-binding proteins are poorly characterized, and it is likely that many more await discovery.

In our studies of the VlsE antigenic variation system of Borrelia burgdorferi, the causative agent of Lyme disease [1,2], we discovered that these bacteria produce a cytoplasmic protein which specifically binds a DNA site within the vlsE open reading frame. Using a powerful, unbiased approach, we identified that protein to be the borrelial SpoVG. A broad range of Eubacteria, including many important human pathogens, encodes homologs of SpoVG. The name derives from observations that Bacillus spp. spoVG mutants are unable to complete stage five of sporulation [3–5]. Bacillus spp. mutants exhibit additional defects, such as abnormal cell cycle and division [4,6,7]. Staphylococcus aureus spoVG mutants are less virulent than are wild-type bacteria, and produce significantly lower levels of several pathogenesis-related factors [8–10]. With many organisms, production of spoVG is developmentally regulated and often utilizes alternative sigma factors [11–23]. The three dimensional structures have been determined for SpoVG from S. aureus and other species, and found to be very highly conserved ([24], and Protein Data Base [PDB] accession numbers 2IA9, 2IA9X, 2IA9Z). However, until our discovery, the biochemistry of SpoVG remained a mystery.

Here we demonstrate that the SpoVG homologues of Borrelia burgdorferi, Staphylococcus aureus, and Listeria monocytogenes all bind to DNA. Further investigations determined that, while SpoVG members are highly similar, they have evolved to bind different consensus sequences. Alanine mutagenesis and domain shuffling revealed residues and microdomains required for generalized DNA binding and for nucleotide sequence specificity.

Results

Identification of B. burgdorferi SpoVG as a Site-specific DNA-binding Protein

As part of our studies of the vlsE system, we postulated that B. burgdorferi expresses a cytoplasmic factor(s) that binds near the recombination site, to help control genetic rearrangement. Addressing that hypothesis, we observed that incubating cell-free B. burgdorferi cytoplasmic extract with vlsE DNA retarded the electrophoretic mobility of DNA, consistent with a DNA-protein complex (Fig. 1). This complex was not evident when cytoplasmic...
extracts were heat denatured or treated with proteinase K, indicating the need for a properly folded, intact protein (data not shown). Additional EMSAs narrowed the protein-binding site even further. The 70 bp labeled probe F27B-R10 bound the cytoplasmic protein, and binding was competed by the unlabeled version of that DNA sequence, fragment F27–R10 (Fig. 1A and B, lane 4). DNAs flanking those 70 bp did not compete for protein binding (Fig. 1B, lanes 3 and 5). These results indicate that the borrelial protein binds a DNA sequence of approximately 70 bp (X on Fig. 1A), and that neither of the repeat regions flanking the recombination site is involved in protein binding.

To identify the unknown factor, we took advantage of a DNA affinity chromatography method developed in our laboratory, which has identified several other novel sequence-specific DNA-binding proteins [25–27]. Using a segment of vlsE that included the high-affinity binding site as bait, a protein of approximately 12 kDa was purified. Buffers containing at least 500 mM NaCl were required to elute the protein off the DNA, indicating that the trans acting factor had a high affinity for vlsE bait DNA (Fig. 2). Matrix assisted laser deionizing- time of flight (MALDI-TOF) MS/MS analysis identified the protein as being encoded by open reading frame BB0785, a hypothetical protein of unknown function, with a corresponding Mascot score of 212. Control reactions that used the same cell-free extracts and different DNA baits did not pull-down this protein (data not shown).

Due to its homology with the SpoVG proteins of many bacterial species, we have retained that name for the borrelial gene and protein (Fig. 3). To confirm that this protein was responsible for the protein-DNA complex formed by cytoplasmic extracts, we purified recombinant B. burgdorferi SpoVG (SpoVG<sub>BB</sub>) and repeated the EMSAs. Indeed, recombinant SpoVG<sub>BB</sub> bound to probe F27B-R10 (Fig. 4A). This 70-mer was dissected and one 18 bp fragment was found to be required and sufficient for SpoVG-binding (Figs. 4B & C). SpoVG<sub>BB</sub> bound to its high-affinity target DNA with an apparent dissociation constant (K<sub>D</sub>) of 308 (±31) nM. Further controls incorporated ctp Operator DNA, a region of DNA known to be bound by other B. burgdorferi DNA-binding proteins [25,28,29]; SpoVG<sub>BB</sub> failed to bind this sequence, confirming its specificity for vlsE DNA (Fig. 4D). The identified SpoVG<sub>BB</sub>-binding sequence does not occur anywhere else in the B. burgdorferi genome, although it is possible that this protein may bind sequences that differ slightly from the site within vlsE. These studies were the first to demonstrate a function for a SpoVG orthologue. The role(s) of SpoVG<sub>BB</sub> in vlsE expression recombination is still under investigation, and is beyond the scope of this communication on the biochemistry of SpoVG-DNA interactions.

**S. aureus SpoVG is also a Site-specific DNA-binding Protein**

Bioinformatics indicate that many spore and non-spore forming bacteria, Gram positive and Gram negative, encode a SpoVG protein (Fig. 3). Given the high degree of sequence conservation, we hypothesized that these orthologues also bind DNA. Compared to wild-type bacteria, *S. aureus* spoVG mutants express significantly less capA-H mRNA and synthesize reduced levels of capsule [9,17]. We hypothesized that *S. aureus* SpoVG (SpoVG<sub>Sa</sub>) might bind DNA near the cap operon promoter. This was confirmed by EMSA, which demonstrated that recombinant SpoVG<sub>Sa</sub> bound to *S. aureus* (Newman) cap5 5′-non-coding DNA in a dose dependent fashion (Fig. 5, lanes 2–4). Heat denaturation or proteinase K treatment eliminated the shifted EMSA band, confirming that this complex contained functional protein (Fig. 5, lanes 11 and 12). In order to determine the relative affinity of the SpoVG<sub>Sa</sub>-DNA interaction, three independent protein preparations and multiple EMSAs with labeled cap probe were performed with saturating concentrations of NaCl. Eluates were separated by SDS PAGE, stained with Sypro Ruby. M. Molecular mass standards. Lanes 1–3. Proteins eluted with 500, 750, and 1000 mM NaCl elution, respectively. doi:10.1371/journal.pone.0066683.g002

![Figure 1. A B. burgdorferi cytoplasmic protein binds DNA within the vlsE open reading frame. (A) Schematic of the vlsE expression locus and DNAs uses for EMSA. In the upper illustration of the vlsE ORF, the gray areas represent the invariable regions, the white area represents the hypervariable region, and the black bars indicate the direct repeat sequences. Below that, the location of each labeled or unlabeled EMSA DNA is represented by thick or thin black horizontal lines, respectively. “DR” indicates the directly-repeated sequences bordering the variable region of vlsE. (B) EMSAs using B. burgdorferi cytoplasmic extracts. Lane 1–6: 1 nM of labeled probe F27B-R10. Lane 2–6: 10 µg cell-free cytoplasmic extract. Lane 3: 100-fold excess unlabeled competitor F35-R14. Lane 4: 100-fold molar excess unlabeled competitor F27–R10. Lane 5: 100-fold excess unlabeled competitor F29–R16. Lane 6: 100-fold molar excess unlabeled competitor F33–R22.](image)
concentrations of SpoVG\textsubscript{Sa}. These experiments indicated an average Ka of 316 (±42) nM.

In a whole transcriptome screen of a \textit{S. aureus} spoVG mutant, significant alterations in several other virulence-related loci were documented, including \textit{fmiB}, \textit{esxA}, and \textit{lukED} [9]. The ability of SpoVG\textsubscript{Sa} to bind near the promoters of those genes was evaluated using each DNA as an unlabeled EMSA competitor against labeled \textit{cap5} DNA. The \textit{fmiB}, \textit{esxA}, and \textit{lukED} 5’ non-coding DNAs each competed with labeled \textit{cap5} probe for binding of SpoVG\textsubscript{Sa} (Fig. 5, lanes 5–7). Control studies using unlabeled competitors derived from the \textit{esxA} or \textit{cap5A} open reading frames had substantially lesser effects on SpoVG\textsubscript{Sa} binding to the labeled \textit{cap5} probe (Fig. 5, lanes 9 and 10). These results indicate that the 5’-non-coding regions of \textit{cap5}, \textit{fmiB}, \textit{esxA}, and \textit{lukED} all contain a unique sequence(s) to which SpoVG\textsubscript{Sa} binds with high affinity and specificity.

Additional EMASs using a smaller probe and unlabeled competitors narrowed down the high-affinity SpoVG\textsubscript{Sa}-binding sequence in \textit{cap5} promoter-proximal DNA. Probe cap41 contains a SpoVG-binding site (Fig. 6, lane 2). Three unlabeled 28 bp DNAs, which span the 62 bp sequence of probe cap41, were included in EMASs at molar excesses over probe cap41. This type of analysis prevents a possible bias towards probe and/or competitor length, while controlling for potential high affinity interactions at the ends of the probe. At a constant concentration of SpoVG\textsubscript{Sa}, addition of competitor A decreased the amount of bound probe and increased the amount of free DNA (Fig. 6, lanes 3–5). In contrast, 5-fold greater concentrations of competitors B or C did not detectably affect SpoVG\textsubscript{Sa} binding to probe cap41 (Fig. 6, lanes 6 and 7). These data indicate that the high-affinity-binding site is contained within the 28 nucleotides of competitor A.

MEME [Multiple Em for Motif Elicitation] analyses of the DNAs bound by SpoVG\textsubscript{Sa} indicated that all contain at least two 5-TAATT\textsuperscript{I}/A\textsuperscript{S}-3’ sequences (Fig. 7A). Competitor A contains two copies of that motif. To evaluate whether this motif is involved with SpoVG binding, a competitor with mutated motifs was incorporated into subsequent EMASs (Fig. 7C). SpoVG\textsubscript{Sa} exhibited greater than five-fold higher affinity for the wild-type competitor over the mutant (Fig 7B). Taken together, these results demonstrate that the \textit{S. aureus} SpoVG protein preferentially binds to DNA containing an TAATT\textsuperscript{I}/A\textsuperscript{S} motif. Whether SpoVG\textsubscript{Sa} will bind to any such sequence or if surrounding DNA sequences/structures contribute to protein binding remains to be determined.

Different SpoVG Homologues Bind to Different DNA Sequences

The \textit{vlsE} probe, to which SpoVG\textsubscript{Sa} binds with high-affinity and specificity, does not possess the SpoVG\textsubscript{Sa} consensus binding motif. These observations suggested that SpoVG homologues might bind to divergent, distinct DNA sequences. With this in mind, we incubated equal concentrations of SpoVG\textsubscript{Sa} or SpoVG\textsubscript{Bb} with labeled \textit{vlsE} and \textit{cap41} probes in independent EMSAs. SpoVG\textsubscript{Bb} bound to the \textit{vlsE} probe, but not \textit{cap41} (Fig. 8A). Likewise, SpoVG\textsubscript{Sa} bound to only the \textit{cap41} probe (Fig. 8B).

To further address our hypothesis that SpoVG homologues act in a similar fashion, but interact with different sequences, we purified the SpoVG homologue from another firmicute, \textit{Listeria monocytogenes} (SpoVG\textsubscript{Lm}). SpoVG\textsubscript{Lm} bound \textit{S. aureus cap41} promoter DNA but not \textit{B. burgdorferi} \textit{vlsE} DNA (Fig. 8C).

Chimeric SpoVG Proteins Identify Residues Involved with Sequence Specificity

Orthologous proteins are under selective pressure to maintain function, but can diverge in amino acid composition to accommodate the needs of the individual species. Protein structural predictions indicated that SpoVG homologues possess a hypervariable alpha helix at the carboxy terminus (Fig. 3). We suspected that it was this variable domain that contributed to the above-described DNA sequence specificity. To address this hypothesis, we created two different chimeric SpoVG proteins. The staphylococcal SpoVG protein was mutated at residues S\textsuperscript{66} through E\textsuperscript{71} and changed to the corresponding borrelial SpoVG sequence. This strategy by exchanging residues Q\textsuperscript{69} through A\textsuperscript{74} of SpoVG\textsubscript{Bb} with those of the \textit{S. aureus} protein, generating the chimeric protein SpoVG\textsubscript{Bb-SA} (Fig 8AB and Fig. 3). For both chimeras, exchanging 6 residues was sufficient to permit binding to the alternative consensus sequence. SpoVG\textsubscript{Bb-SA} bound to the \textit{cap41} probe, but could no longer bind to the \textit{vlsE} probe. The SpoVG\textsubscript{Sa-Bb} protein now bound \textit{vlsE} DNA. That chimaera
a slight ability to interact with the cap41 DNA, albeit at a dramatically reduced affinity for which a $K_D$ could not be calculated (Fig. 8B). Taken together, these results demonstrate that sequence divergence within the alpha helix contributes to DNA sequence specificity.

Conserved Residues Essential for DNA-protein Complexes

Bacterial proteins that perform analogous functions often retain similar biochemical and structural features in order to interact with their respective ligands [30]. We reasoned that, since three different SpoVG proteins interact with DNA, conserved residues common to all SpoVG orthologues might be required for non-specific substrate binding. Recombinant $\text{SpoVG}_{Sa}$ and $\text{SpoVG}_{Bb}$ proteins were produced that included single or double amino acid substitutions at conserved positions (Fig. 3, and Table 1). These mutant proteins were tested for their abilities to interact with their respective high-affinity DNA sequences.

Initial investigations targeted a doublet of positively charged residues (R and K), which were conserved in all SpoVG homologues (Fig. 3). The two charged residues are predicted to project inward from an abbreviated $\beta$-sheet, toward the carboxy-terminal alpha helix. Alanine substitutions at position R53–R54 of $\text{SpoVG}_{Bb}$ or K50–R51 of $\text{SpoVG}_{Sa}$ impaired DNA binding. Addition of mutant proteins at five-fold excess over the dissociation constant of the wild-type protein still did not produce a detectable EMSA shift (Figs. 3 and 9). To assay residues independently, $\text{SpoVG}_{Sa}$ K50A and $\text{SpoVG}_{Sa}$ R51A were created. These variants exhibited the same deficiency in DNA binding as

Figure 4. Identification of a $\text{SpoVG}_{Bb}$ high-affinity binding site. Recombinant $\text{SpoVG}_{Bb}$ binds specifically to an 18 bp sequence of the $\nu B E$ open reading frame. (A) Lanes 1–3: 1 nM of labeled F27B-R10. Lane 2:300 nM $\text{SpoVG}_{Bb}$. Lane 3:600 nM $\text{SpoVG}_{Bb}$. (B) Lanes 1–3: 1 nM of labeled F77B-R10. Lane 2:300 nM $\text{SpoVG}_{Bb}$. Lane 3:600 nM $\text{SpoVG}_{Bb}$. (C) Lanes 1–3: 1 nM of labeled F27B-R4. Lane 2:300 nM $\text{SpoVG}_{Bb}$. Lane 3:600 nM $\text{SpoVG}_{Bb}$. (D) Lanes 1–3: 1 nM of labeled $\text{erp Operator DNA}$. Lane 2:300 nM $\text{SpoVG}_{Bb}$. Lane 3:600 nM $\text{SpoVG}_{Bb}$.
doi:10.1371/journal.pone.0066683.g004
did the double mutant, confirming that both conserved residues are required for DNA binding.

Mutations to other conserved, positively charged residues did not have any significant effects on DNA binding (Fig. 3, Table 1, and data not shown). Additionally, none of the other mutant proteins exhibited altered sequence preference (data not shown).

Site-directed Mutagenesis did not Affect Multimerization

Replacing charged or polar residues with a small, non-polar, uncharged alanine can interfere with protein-protein interactions, or cause protein misfolding [31]. To that end, sizing chromatography and tandem native/denaturing PAGE analysis were used to examine the native state of SpoVG<sub>SA</sub>. The recombinant protein has a molecular mass of 14.6 kDa. By two independent methods, our data indicate that SpoVG<sub>SA</sub> forms a 55–60 kDa complex in solution, consistent with a tetramer (Fig. 10). The complexes disappeared when samples were denatured, demonstrating that these bands were not the results of contamination (Fig. 10C). None of the SpoVG<sub>SA</sub> mutants exhibited diminished multimer formation, suggesting that the mutants which were impaired for DNA binding still retained their ability to fold correctly and form higher ordered species in solution.

Discussion

The current studies yielded several novel findings that impact a broad range of Eubacterial species. First, SpoVG orthologues from three distinct bacteria bound DNA. For several bacterial species, it is known that these small proteins play key roles in critical cellular processes, which we can now hypothesize are due to SpoVG-DNA interactions. Second, these discoveries help explain why SpoVG was found in association with the <i>S. aureus</i> nucleoid, and the involvement of the <i>B. subtilis</i> orthologue with nucleoid organization [4,32]. Third, while SpoVG proteins are highly conserved overall, the <i>S. aureus</i> and <i>B. burgdorferi</i> proteins interact

Figure 5. SpoVG<sub>SA</sub> binds specifically to <i>S. aureus</i> cap5, fmtB, <i>lukED</i>, and <i>saeX</i> DNAs adjacent to the promoter. Illustrated are EMSAs with <i>S. aureus</i> SpoVG<sub>SA</sub> labeled cap5 5’ non-coding DNA, and various unlabeled competitor DNAs. Lanes 1–12.5 nM of labeled capA promoter DNA. Lanes 2–4: Increasing amounts of SpoVG<sub>SA</sub> (0.2 μg, 0.4 μg, and 0.6 μg, respectively). Lanes 5–12: 0.8 μg of SpoVG<sub>SA</sub>. Lanes 5–8: 50-fold molar excess of unlabeled fmtB, lukED, exsA, or cap5 5’ non-coding DNAs, respectively. Lane 9: 50-fold molar excess of exsA ORF DNA. Lane 10: 50-fold molar excess of cap5A ORF DNA. Lane 11: SpoVG<sub>SA</sub> was heated to 99°C for 5 minutes before use in EMSA. Lane 12: SpoVG<sub>SA</sub> was preincubated with 5 mg/ml of Proteinase K before use in EMSA.

doi:10.1371/journal.pone.0066683.g005

Figure 6. SpoVG<sub>SA</sub> interacts specifically with a 28 bp region adjacent to the cap5 promoter. EMSA with a labeled 62 bp probe derived from <i>cap5</i> 5’-non-coding DNA (cap41) and different concentration of 28 bp cold competitors. All lanes contain 5 nM labeled cap41 DNA. Lanes 2–6 also include 0.9 μg of SpoVG<sub>SA</sub>. Lanes 3 and 4: 25 fold molar excess competitor A or mutant competitor A, respectively. Lanes 5 and 6: 50 fold molar excess of competitor A and mutant competitor A, respectively. (C) Sequences of probe cap41 and competitors. The differences between the wild type and mutant competitors are indicated by boldface italics.

doi:10.1371/journal.pone.0066683.g006

Figure 7. The palindromic motif 5’-ATTAA/A-3’ is required for SpoVG<sub>SA</sub> binding. (A) The conserved sequence, 5’ to 3’, identified by multiple motif analysis of promoters bound to and influenced by SpoVG<sub>SA</sub>. (B) EMSA with a labeled probe cap41, derived from cap5 sequence, and two different 28 bp cold competitors. All lanes contain 5 nM labeled cap41 DNA. Lanes 2–6 also include 0.8 μg of SpoVG<sub>SA</sub>. Lanes 3 and 4: 25 fold molar excess competitor A or mutant competitor A, respectively. Lanes 5 and 6: 50 fold molar excess of competitor A and mutant competitor A, respectively. (C) Sequences of probe cap41 and competitors. The differences between the wild type and mutant competitors are indicated by boldface italics.

doi:10.1371/journal.pone.0066683.g007
preferentially with distinct DNA sequences. Given the amino acid divergence among different orthologs' carboxy-terminal alpha helices, we speculate that this feature may also be true for other SpoVG homologues. Finally, we identified two residues, whose biochemical properties are conserved among SpoVGs, that are essential for DNA interactions.

Residues involved with maintaining SpoVG secondary structure model well between species, suggesting that the solved crystal structures are likely to be representative of all orthologous proteins. Merging all of these data, we propose a model for SpoVG binding (Fig. 11). Solvent-accessible, positively charged residues are located adjacent to the alpha helix and can stabilize duplex binding through electrostatic interactions with the phosphate backbone of DNA. These are residues R53 and R54 of SpoVG<sub>Bb</sub>, and K50 and R51 of SpoVG<sub>Sa</sub>. Those residues extend into a pocket, while the alpha helix is arranged perpendicularly to provide base-edge specificity through interactions by residues extending into the pocket (Figs. 10 and 11). Notably, the <i>B. burgdorferi</i> and <i>S. aureus</i> alpha helices are out of phase by approximately one turn of the alpha helix, presenting residues with dissimilar hydrogen-donating and hydrogen-accepting capabilities on the upper helical face ([31], and Fig. 11). Independent evolution of the two studied SpoVG proteins resulted in different nucleic acid binding specificity. Our data suggest that SpoVG homologs of different bacterial species may bind to distinct DNA sequences, and possibly exert different effects on physiology. Similar phenomena have been documented that alter the specificity, diversify the signal, and eliminate unwanted cross-talk between sensor histidine kinases and response regulators in two-component signal transduction systems [33–35].

The mechanisms by which <i>S. aureus</i> controls of virulence-associated genes are poorly understood. The identification of a SpoVG<sub>Sa</sub>-binding site adjacent to the cap promoter suggests that SpoVG<sub>Sa</sub> may play a direct role in controlling capsule production. Indeed, cap transcription is significantly reduced in spoVG mutant <i>S. aureus</i> [9], and <i>S. aureus</i> lacking a SpoVG<sub>Sa</sub>-binding site in the cap promoter exhibit reduced cap transcription [36]. However, expression of the cap operon has been reported to be controlled by at least 12 other regulatory factors [9,17,36–40]. Studies are currently under way to define binding-sites of these many regulators interact with each other and with RNA polymerase to control cap expression.

The role of SpoVG<sub>Bb</sub> in <i>B. burgdorferi</i> vlsE genetic rearrangement remains to be determined. The specialized recombination processes involved are complex and highly regulated, occurring only during mammalian infection but never during tick colonization or in culture [2,41]. Recombination of vlsE is RecA-independent, requires holiday junction resolvases, and may involve G-quadruplex DNA [42–44]. Our preliminary studies suggest that SpoVG may interact with other, as-yet unidentified factors. We are continuing studies to identify other players in the
Table 1. Plasmids used in this study.

<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Backbone</th>
<th>Organism/Target locus</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBLJ340</td>
<td>pET101</td>
<td>L.monocytogenes/spoVG</td>
<td>ΔSpoVG</td>
</tr>
<tr>
<td>pBLJ341</td>
<td>pET101</td>
<td>B.burgdorferi/spoVG</td>
<td>ΔSpoVG R53A, R54A</td>
</tr>
<tr>
<td>pBLJ342</td>
<td>pET101</td>
<td>B.burgdorferi/spoVG</td>
<td>ΔSpoVG R50A, R51A</td>
</tr>
<tr>
<td>pBLJ343</td>
<td>pET101</td>
<td>B.burgdorferi/spoVG</td>
<td>ΔSpoVG R72A</td>
</tr>
<tr>
<td>pBLJ347</td>
<td>pET101</td>
<td>B.burgdorferi/spoVG</td>
<td>ΔSpoVG BbQDFRKA Sa SDMRQE</td>
</tr>
<tr>
<td>pBLJ505</td>
<td>pET101</td>
<td>S.aureus/spoVG</td>
<td>ΔSpoVG K50A, R51A</td>
</tr>
<tr>
<td>pBLJ351</td>
<td>pET101</td>
<td>S.aureus/spoVG</td>
<td>ΔSpoVG K50A</td>
</tr>
<tr>
<td>pBLJ352</td>
<td>pET101</td>
<td>S.aureus/spoVG</td>
<td>ΔSpoVG H62A</td>
</tr>
<tr>
<td>pBLJ353</td>
<td>pET101</td>
<td>S.aureus/spoVG</td>
<td>ΔSpoVG R69A</td>
</tr>
<tr>
<td>pBLJ354</td>
<td>pET101</td>
<td>S.aureus/spoVG</td>
<td>ΔSpoVG R50A</td>
</tr>
<tr>
<td>pBLJ355</td>
<td>pET101</td>
<td>S.aureus/spoVG</td>
<td>ΔSpoVG R51A</td>
</tr>
<tr>
<td>pBLJ357</td>
<td>pET101</td>
<td>S.aureus/spoVG</td>
<td>ΔSpoVG SaSDMRQE BbQDFRKA</td>
</tr>
<tr>
<td>pBLJ506</td>
<td>pCR2.1</td>
<td>S.aureus/cap5A</td>
<td>cap5' UTR</td>
</tr>
<tr>
<td>pBLJ507</td>
<td>pCR2.1</td>
<td>S.aureus/fmt</td>
<td>fmt8 UTR</td>
</tr>
<tr>
<td>pBLJ508</td>
<td>pCR2.1</td>
<td>S.aureus/esxA</td>
<td>esxA UTR</td>
</tr>
<tr>
<td>pBLJ509</td>
<td>pCR2.1</td>
<td>S.aureus/luk</td>
<td>lukED UTR</td>
</tr>
<tr>
<td>pTB7a</td>
<td>pCR2.1</td>
<td>B. burgdorferi/vlsE</td>
<td>vlsE - 80 +240</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0066683.t001

In conclusion, our data suggest that all SpoVG orthologues are controls gene expression in B. burgdorferi orthologs. It is also possible that SpoVG controls gene expression in B. burgdorferi as do the S. aureus and B. subtilis orthologs.

Figure 10. Site directed mutagenesis did not influence SpoVG oligomerization. Results of panels A through D illustrate HPLC sizing column chromatography of wild type and mutant SpoVG proteins. For some preparations, proteins eluted with a retention time of approximately 7 minutes, which corresponds with a molecular mass >440 kDa and are composed of protein aggregates. (A) Wild-type SpoVGsa, (B) SpoVGsa K50A-R51A; (C) SpoVGsa K50A; (D) SpoVGsa R51A. Peaks marked with red asterisks indicate retention volumes corresponding with approximately 55–60 kDa. Panels E and F illustrate proteins separated following native of denaturing PAGE, respectively. M, Molecular mass standards; Lane 1, 1.5 μM mutant SpoVGsa K50A-R51A; Lane 2, 1.5 μM mutant SpoVGsa K50A; Lane 3, 1 μM mutant SpoVGsa R51A; Lane 4, 1 μM wild-type SpoVGsa. doi:10.1371/journal.pone.0066683.g010
DNA-affinity Chromatography

A protein was purified from *B. burgdorferi* cytoplasmic extract based on its affinity for *vlsE* DNA bait, using previously-described procedures [25,27]. Bait DNA was generated by PCR of the *B. burgdorferi* *vlsE* coding region using one 5′-biotin-modified and one unmodified oligonucleotide (Table 2). A single band that eluted in buffer containing 750 mM NaCl was excised and MALDI-TOF

Figure 11. Monomeric structures of SpoVG<sub>Sa</sub> and SpoVG<sub>Bb</sub>, modeled on the solved *S. epidermidis* SpoVG protein structure. Residues required for DNA-binding and those involved with sequence specificity are indicated, with different colors reflecting biochemical properties of amino acids: Gray = positively charged, Red = negatively charged, and Green = polar, uncharged. (A) SpoVG<sub>Sa</sub> (B) SpoVG<sub>Bb</sub>

doi:10.1371/journal.pone.0066683.g011

Table 2. Oligonucleotides used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5′-3′)</th>
<th>Target</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio CapUp F-1</td>
<td>CTA TCT GAT AAT AAT CAT CTA ACT CAC</td>
<td>capA5 5′UTR</td>
<td>5′ Biotin</td>
</tr>
<tr>
<td>CapUp F-2</td>
<td>TAT TTA CCT CCC TTA AAA ATT TTC</td>
<td>capA5 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>CapUp F-1</td>
<td>CTA TCT GAT AAT CAT CTA ACT CAC</td>
<td>capA5 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>fmtUp F-1</td>
<td>GTC ATC CTC CTG GTT GAT TAT TC</td>
<td>fmtB 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>fmtUp R-2</td>
<td>fmtB 5′UTR</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>lukEUp F-1</td>
<td>CTT AAA CAT AAG TTT CAC TTT CTT TC</td>
<td>lukE 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>lukEUp R-2</td>
<td>lukE 5′UTR</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>exsAlUp F-1</td>
<td>CCT TTA TGT ATT TAA TTT AG</td>
<td>exs 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>exsAlUp R-2</td>
<td>CCT CCT GAA TAT TTT AAG TTT ATC</td>
<td>exs ORF</td>
<td>None</td>
</tr>
<tr>
<td>exsAlR-1</td>
<td>CAC CAA TCC AGT TCA TTT CTG</td>
<td>exs ORF</td>
<td>None</td>
</tr>
<tr>
<td>exsAlR-2</td>
<td>CCA AAG ATG GAC ACA ACG ATT AG</td>
<td>exs ORF</td>
<td>None</td>
</tr>
<tr>
<td>CapIn F-1</td>
<td>GCG CTA TTG TTA CAT TTT TCG TC</td>
<td>capA5 ORF</td>
<td>None</td>
</tr>
<tr>
<td>CapIn R-2</td>
<td>GGT GAA TAC TTA CTA TTT AAG TCC</td>
<td>capA5 ORF</td>
<td>None</td>
</tr>
<tr>
<td>BioSpo41</td>
<td>GAG TAT AAT TTT TAA TTT ACA TAT AAA AAG GGA AAT AAT GCG GTT TAA AAG TAA TTA AT</td>
<td>capA5 5′UTR</td>
<td>5′ Biotin</td>
</tr>
<tr>
<td>Spo41-F</td>
<td>GAG TAT AAT TTT TAA TTT ACA TAT AAA AAG GGA AAT AAT GCG GTT TAA AAG TAA TTA AT</td>
<td>capA5 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>Spo42-R</td>
<td>ATT AAT TAC TTT TAA ACC GCA TTA TTT TCG CTT TTT TAA TTA ATA TAC TC</td>
<td>capA5 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>SpoMut33-F</td>
<td>GAG TAG CGG CCT TTT CGC CGG ACA TAT A</td>
<td>capA5 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>SpoMut34-R</td>
<td>TAT ATG TCC GGC GAA AAG CCG GCT ACT C</td>
<td>capA5 5′UTR</td>
<td>None</td>
</tr>
<tr>
<td>Bvls27F</td>
<td>GTT AAT AGT TTT CCT AAG G</td>
<td>vlsE ORF</td>
<td>5′ Biotin</td>
</tr>
<tr>
<td>Bvls77F</td>
<td>GTA CAG GTT CTG GA</td>
<td>vlsE ORF</td>
<td>5′ Biotin</td>
</tr>
<tr>
<td>vlsR10</td>
<td>ACC AAC AGA ACC TGT AC</td>
<td>vlsE ORF</td>
<td>None</td>
</tr>
<tr>
<td>vlsR4</td>
<td>CTG GTC TCC CGG GTC GTA GTA C</td>
<td>vlsE ORF</td>
<td>None</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0066683.t002
MS/MS analysis was performed at the Clinical Proteomics Center, University of Louisville. Peptide masses were compared to the B. burgdorferi strain B31-MI sequence [46,47], using Mascot (Matrix Science, Boston, MA). Significance parameters were fixed at p<0.05, corresponding to Ion score of 81.

**Recombinant Proteins**

Purified B. burgdorferi B31 DNA was used as template to clone the borrelial spoVG gene into pET101, creating pBLJ132. Similarly, the S. aureus and L. monocytogenes spoVG genes were individually cloned into pET101 (Invitrogen, Grand Island, NY), producing pBLJ305 and pBLJ340, respectively. Each cloned insert was completely sequenced to confirm that the spoVG gene was free of mutations and in-frame with the hexa-histidine tag. *Escherichia coli* Rosetta-2 (Novagen, EMB Millipore, Billerica, MA) was independently transformed with pBLJ132, pBLJ340, or pBLJ305. Recombinant proteins were induced by the addition of 1 mM IPTG, and purified using MagnaHis Ni-Particles (Promega, Madison, WI). In order to create conditions conducive to protein-DNA interactions, each SpoVG protein was dialyzed against a buffer containing 100 mM dithiothreitol, 30 mM Tris-HCl, 25 mM KCl, 10% glycerol vol/vol, 0.01% Tween-20, 1 mM phenylmethanesulfonyl fluoride [25,26,29,48]. Protein purities and concentrations were assessed via SDS-PAGE and Bradford assays (Bio-Rad, Hercules, CA) respectively. Protein aliquots were snap frozen in liquid nitrogen and stored at −80°C.

To generate mutant SpoVG proteins, site-directed mutagenesis was performed on wild-type plasmid clones, as previously described [49]. Each plasmid was sequenced to confirm accuracy of mutations. All proteins were expressed, purified, and otherwise handled in the same manner. At least two independent protein preparations were used to evaluate each mutant protein that had a phenotypic difference from the wild-type protein. Tables 1 and 2 describe all probes, competitors, and mutant SpoVG proteins produced in this study follow the text.

**Electromobility Gel Shift Assays (EMSA)**

Sequences of oligonucleotides used in this study are listed in Table 2. Oligonucleotide primers specific for the *B. burgdorferi* csoE coding region or *S. aureus* capA 5’ non-coding region were used to produce labeled probes, with one primer modified to include a 5’ biotin moiety that allowed for chemiluminescent detection. PCR-synthesized probes were purified by gel electrophoresis. Smaller, labeled DNA fragments were annealed by an initial high-temperature melting step, followed by incremental decreases in temperature using a thermocycler [48].

Unlabeled competitor DNAs were also generated via PCR or by annealing oligonucleotides. Larger competitors, consisting of *S. aureus* capA, fimB, exaA, and *lukED* 5’ non-coding DNAs, were PCR amplified, and cloned into pCR2.1 (Invitrogen, Grand Island, NY, USA), generating pBLJ506, 507, 508, and 509, respectively. Each plasmid was sequenced to ensure that the clones were free of mutations. These constructs were then used as templates for PCR generation of specific competitors (Table 2). Amplicons were separated by agarose gel electrophoresis and purified using Wizard DNA Clean-up Systems (Promega, Madison, WI) before use as EMSA competitors. All probe and competitor DNA concentrations were determined spectrophotometrically. When appropriate, competitor concentrations and oligonucleotide annealing efficiencies were also confirmed using relative ethidium bromide-stained band intensity following electrophoresis through native 20% polyacrylamide gels (Invitrogen, Grand Island, NY).

EMSA conditions were essentially the same as those described previously [25,26,29,48]. Protein-DNA binding buffer consisted of 50 mM Tris-HCl (pH 7.5), 1 mM dithiothreitol, 130 mM EDTA, 50 ng/μl poly d-dC, 3.5 μl/ml phosphatase inhibitor (Sigma, St Louis, MO, USA), 9 μl/ml protease inhibitor cocktail 2 (Sigma, St Louis, MO). For reactions involving cell extracts B. burgdorferi B31-MI-16, cells were pelleted, resuspended in the above-described buffer, lysed, and cellular debris cleared by centrifugation. All EMSAs were performed at room temperature (approx. 20°C). Probe concentrations were varied as noted in the text. DNA and protein-bound DNA complexes were separated by electrophoresis through native a 10% polyacrylamide TBE gels (Invitrogen), transferred to a nylon membrane (Thermo Scientific, Waltham, MA), and UV cross linked (Stratagene UV Stratalinker 1800, La Jolla, CA). Nucleic acid probes were detected via chemiluminescence (Thermo Scientific) and visualized by autoradiography. Band densitometry was assessed using ImageJ [http://rsbweb.nih.gov/ij] and dissocation constants (Kd) determined as previously described [50].

**Bioinformatic Analyses**

Promoter motif and structural analyses were performed using MEME (Multiple Em for Motif Elcication) [http://meme.sdsu.edu/meme/cgi-bin/meme.cgi]. SpoVG amino acid sequences were retrieved from GenBank consortium (http://www.ncbi.nlm.nih.gov/sites). Amino acid sequences were Muscle aligned [http://www.ebi.ac.uk/Tools/msa/muscle/], with gap penalties set at 10, and a minimum of 1000 iterations. Images and analysis were generated using Geneious (http://www.geneious.com).

Species, strain, and GenBank accession numbers used for the analysis shown in Figure 4 were as follows: *Staphylococcus aureus* Newman, NP_645270.1; *Staphylococcus epidermidis* ATCC 12289, NP_765840.1; *Abiotrophia deficiens* ATCC 49176, ZP_04425046.1; *Bacillus anthracis* G9241, ZP_00240564.1; *Bacillus cereus* ATCC 14579, NP_829950.1; *Bacillus megaterium* WSH-002, YP_005497349.1; *Bacillus subtilis* 168, NP_387930.1; *Bifidobacterium bacteriocomus* HD100, NP_969591.1; *Borrelia burgdorferi* B31, NP_212919.1; *Borelia hermsii* DAH, YP_001848203.1; *Calicidellopsis sarachaktivis* DSM 8903, YP_001179173.1; *Clostridium botulinum* ATCC 3502, YP_001256027.1; *Clostridium perfringens* SM10, YP_699747.1; *Corallocolus corallodermis* DSM 2259, YP_005368402.1; *Desulfovibrio baarsii* DSM 2075, YP_003806692.1; *Elasmobacterium minutum* Pei191, YP_001875395.1; *Haloblastus contractile* SSD-17B, ZP_08554794.1; *Helibacterium medesiacidicum* Ice1, YP_001679386.1; *Ilyobacter polytropus* DSM 2926, YP_003968016.1; *Kyrpidia turicens* DSM 2912, YP_003587990.1; *Leptotrichia goodfellowii* F0264, ZP_06012807.1; *Listeria monocytogenes* EGDe, NP_463727.1; *Macrocococcus caseolyticus* JSNC5402, YP_002561317.1; *Mycnococcus xanthus* DK 1622, YP_633282.1; *Prevotella dentalis* DSM 3688, EHO38682.1; *Stigmatella aurantia* DW4/3-1, YP_003954342.1; *Spirochaeta thermophila* DSM 6192, YP_003874257.1; and *Trepnema dentale* ATCC 35405, NP_971943.1.

**Multimerization State of SpoVG**

15% SDS PAGE and 10% native PAGE were used to evaluate denatured and native masses of SpoVG preparations. Following electrophoretic separation polyacrylamide gels were stained with Coomassie Brilliant Blue.

Size-exclusion column chromatography was also employed to assess SpoVG multimerization in solution. A Superdex 200 10/300 column (GE Healthcare, Catalog No. 17-5175-01) was prepared per the manufacturer’s instructions with a mobile phase consisting of 300 mM NaCl, 25 mM Na2HPO4 (pH adjusted to 7.0 with 5.0 M HCl), 1 mM Na3N, and 1% glycerol. The mobile
phase was sterile filtered to 0.22 μm. The flow rate was set to 0.10 ml/minute and elution was monitored at A280. The elution of proteins was calibrated using standards of known molecular weight from GE Healthcare LMW and HML Gel Filtration Calibration Kits. (Catalog Nos. 20-4030-41 and 20-4030-42). First, the void volume (V0) of the column was determined by injection of 100 μl of 1 mg/ml blue dextran 2000 (2,000 kDa) in elution buffer with 5% glycerol. Protein standards consisting of thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), ribonuclease A (13.7 kDa), and bovine lung aprotinin (6.5 kDa) were individually prepared in elution buffer with 5% glycerol at 10 mg/ml. These standards were then diluted such that each individual protein had a concentration of 2.0 mg/ml and injected in 100 μl aliquots. The log of the molecular masses of these standards was then graphed against resulting elution volumes (Ve) as Ve/V0 to produce a linear calibration.

References


Acknowledgments

We thank Adriana Rosato, Sarah D’Orazio, and Carol Fickett for providing bacterial strains used in this study. We appreciate the fruitful discussions and comments from Michael Fried, Gavin Ellis, Catherine Brissette, Colleen Foy, and Adriana Rosato.

Author Contributions

Conceived and designed the experiments: BLJ TB BS. Performed the experiments: BLJ AC CLR DC MCM TB. Analyzed the data: BLJ AC MCM TB BS. Contributed reagents/materials/analysis tools: BLJ AC MCM TB BS. Wrote the paper: BLJ BS.