Twin RNA Polymerase–associated Proteins Control Virulence Gene Expression in Francisella tularensis

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi://10.1371/journal.ppat.0030084</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:8156563">http://nrs.harvard.edu/urn-3:HUL.InstRepos:8156563</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Twin RNA Polymerase–Associated Proteins Control Virulence Gene Expression in Francisella tularensis

James C. Charity1, Michelle M. Costante-Hamm1, Emmy L. Balon2, Dana H. Boyd3, Eric J. Rubin2, Simon L. Dove1*

1 Division of Infectious Diseases, Children’s Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, 2 Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts, United States of America, 3 Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts, United States of America

The MglA protein is the only known regulator of virulence gene expression in Francisella tularensis, yet it is unclear how it functions. F. tularensis also contains an MglA-like protein called SspA. Here, we show that MglA and SspA cooperate with one another to control virulence gene expression in F. tularensis. Using a directed proteomic approach, we show that both MglA and SspA associate with RNA polymerase (RNAP) in F. tularensis, and that SspA is required for MglA to associate with RNAP. Furthermore, bacterial two-hybrid and biochemical assays indicate that MglA and SspA interact with one another directly. Finally, through genome-wide expression analyses, we demonstrate that MglA and SspA regulate the same set of genes. Our results suggest that a complex involving both MglA and SspA associates with RNAP to positively control virulence gene expression in F. tularensis. The F. tularensis genome is unusual in that it contains two genes encoding different α subunits of RNAP, and we show here that these two α subunits are incorporated into RNAP. Thus, as well as identifying SspA as a second critical regulator of virulence gene expression in F. tularensis, our findings provide a framework for understanding the mechanistic basis for virulence gene control in a bacterium whose transcription apparatus is unique.

Introduction

Francisella tularensis is a Gram-negative, facultative intracellular pathogen and the aetiologic agent of tularemia, a disease that can be fatal in humans. Although outbreaks of tularemia are thought to be rare, infections caused by F. tularensis have become a public health issue because of the potential for using this organism as a bioweapon. As few as ten organisms can constitute an infectious dose, and the pneumonic form of tularemia, which has a particularly high mortality rate, can be acquired when the organism is aerosolized [1,2]. Relatively little is known regarding the molecular mechanisms of F. tularensis pathogenesis, in part because of the genetic intractability of this organism.

Evidence suggests that the pathogenicity of F. tularensis depends on its ability to survive and replicate within macrophages. However, only a handful of genes that are required for intramacrophage survival have been identified [2–4]. One of these encodes a putative transcription regulator called macrophage growth locus A (MglA) that is responsible for controlling the expression of multiple virulence factors that are themselves required for survival within macrophages [3,5,6]. In particular, in F. tularensis subspecies novicida, MglA positively regulates the expression of multiple virulence genes, some of which are located on a pathogenicity island [5–7]. MglA also regulates the expression of many other genes that may or may not be implicated in virulence [3,6]. However, the mechanism of this MglA-dependent regulation is unknown.

MglA is an ortholog of the stringent starvation protein A (SspA) from Escherichia coli (the two proteins share 21% identity and 34% similarity at the primary amino acid level).

In E. coli, SspA is a RNA polymerase (RNAP)–associated protein that is thought to regulate the expression of a subset of genes under conditions of stress [8,9]. SspA in E. coli is also essential for the lytic growth of bacteriophage P1 [10]. Specifically, Hansen and colleagues have shown that E. coli SspA functions as an activator of P1 late gene expression, acting in concert with a phage-encoded transcription activator called Lpa [11]. SspA homologs have also been found in a number of pathogens other than F. tularensis, and several of these homologs appear to play important roles in virulence gene regulation [12–14].

Recently, the genomes of two different strains of F. tularensis have been sequenced ([15]; GenBank accession number NC_007880). One of these strains is SctHu S4, a highly virulent strain of subspecies tularensis, and the other is the live vaccine strain (LVS), an attenuated derivative of one of the subspecies holarctica strains. The genome sequences reveal that F. tularensis contains an MglA-like protein, which has been annotated as SspA [15]. In addition, the sequences...
Author Summary

The Gram-negative bacterium Francisella tularensis is an intracellular pathogen and the causative agent of tularemia. In F. tularensis, the MglA protein is the only known regulator of virulence genes that are important for intracellular survival, yet it is not known how MglA functions. F. tularensis also contains an MglA-like protein called SspA, whose function is not known. In this study, we show that both MglA and SspA associate with RNA polymerase (RNAP) and positively regulate virulence gene expression in F. tularensis. Our study provides evidence that MglA and SspA interact with one another directly and that the association of MglA with RNAP is dependent on the presence of SspA. We also show that, unlike RNAP from any other bacterium, RNAP from F. tularensis contains two distinct α subunits. Given the fundamental roles the α subunit plays in transcription regulation, this may have far-reaching implications for how gene expression is controlled in F. tularensis. Our study therefore uncovers a new critical regulator of virulence gene expression in F. tularensis (SspA), provides mechanistic insight into how MglA and SspA cooperate to control virulence gene expression, and reveals that the F. tularensis transcription machinery has an unusual composition.

Results

MglA and SspA Associate with RNAP in F. tularensis

To address the question of whether MglA and SspA are associated with RNAP in F. tularensis, we took a directed proteomic approach in which we affinity purified RNAP from F. tularensis and identified candidate RNAP-associated proteins by tandem mass spectrometry (MS/MS). To facilitate affinity purification of F. tularensis RNAP, we adapted the tandem affinity purification (TAP) strategy [17] for use in F. tularensis. Using the integration vector depicted in Figure 1A, we constructed a strain of F. tularensis LVS in which the native chromosomal copy of the rpoC gene (encoding the β′ subunit of RNAP) has been modified to encode a TAP-tagged form of β′ (β′-TAP). The resulting strain (LVS β′-TAP) thus synthesizes, at native levels, β′ with a TAP tag fused to its C-terminus. As a control we constructed, in a similar fashion to the LVS β′-TAP strain, a strain that synthesizes SucD (a subunit of the succinyl-CoA synthetase complex) with a TAP-tag fused to its C-terminus (LVS SucD-TAP). Lysates were made from cells of the LVS SucD-TAP strain and cells of the LVS β′-TAP strain. Proteins were then purified by TAP essentially as described [18], separated by SDS-PAGE, and stained with silver. Two proteins with apparent molecular weights of ~20 kDa were found that specifically co-purified with β′ (Figure 1B). Bands corresponding to these proteins were excised from the gel, and, following in-gel digestion with trypsin, nanoelectrospray MS/MS was used to identify the proteins as MglA and SspA (unpublished data). Among the proteins that appear to specifically co-purify with β′, one had an apparent molecular weight of ~70 kDa, two had apparent molecular weights of ~40 kDa, and one had an apparent molecular weight of ~10 kDa (Figure 1B). Bands corresponding to these proteins were excised from the gel, and, following in-gel digestion with trypsin, nanoelectrospray MS/MS was used to determine the identity of each protein. Accordingly, the ~70- and ~10-kDa proteins were found to be the 70 and 10 kDa (Figure 1B). Bands corresponding to these proteins were excised from the gel, and, following in-gel digestion with trypsin, nanoelectrospray MS/MS was used to determine the identity of each protein. Accordingly, the ~70- and ~10-kDa proteins were found to be the 70 and 10-kDa proteins were found to be the 70 and 10-kDa proteins were found to be z1 and z2, the products of the rpoA1 and rpoA2 genes, respectively (unpublished data). Therefore, F. tularensis RNAP contains two distinct α subunits and associates with both MglA and SspA.

The Association of MglA with RNAP Is Dependent upon SspA

We reasoned that if MglA were associated with RNAP in F. tularensis, then RNAP would be expected to co-purify with TAP-tagged MglA. To test this prediction, we constructed a strain of LVS that synthesized MglA with a TAP-tag fused to its C-terminus (LVS MglA-TAP). Figure 1B shows the results following TAP of MglA from this strain. In support of the hypothesis that MglA associates with RNAP in F. tularensis, subunits of RNAP (including z1, z2, β′, and β′) co-purified with MglA. Furthermore, SspA also co-purified with MglA. If MglA and SspA were to interact with RNAP independently of one another, then SspA might not be expected to co-purify with MglA-TAP. In particular, complexes containing RNAP and MglA would be expected to be distinct from those containing RNAP and SspA if MglA and SspA competed for the same binding site on RNAP. It is of course possible that there is more than one contact site for MglA and/or SspA on RNAP. However, because SspA from Yersinia pestis is a dimer [19], one plausible explanation for our results is that MglA and SspA physically interact with one another and associate with RNAP as a heterodimer. An alternative possibility is that MglA might form two distinct complexes, one with SspA, and another with RNAP; in this case, the SspA that co-purifies with β′-TAP (Figure 1B, lane 2) would represent SspA-RNAP complexes that are distinct from MglA-RNAP complexes. In an attempt to distinguish between the possibilities that MglA and SspA form a single complex or separate complexes with RNAP, we constructed a strain that synthesized MglA-TAP and carried an in-frame deletion of sspA (LVS ΔsspA MglA-TAP). Figure 1C shows the results following TAP of MglA-TAP from cells of this strain. In the absence of SspA, MglA fails to co-purify with RNAP. Our findings are therefore consistent with the hypothesis that MglA and SspA are associated with one another in F. tularensis and that the resulting MglA–SspA complex associates with RNAP.
MgL\(\alpha\) and Ssp\(\alpha\) Interact with One Another Directly

In order to test explicitly whether or not MgL\(\alpha\) and Ssp\(\alpha\) can interact with one another directly, we used a bacterial two-hybrid assay configured to permit the detection of both dimeric and higher order complexes. This two-hybrid assay is based on the finding that any sufficiently strong interaction between two proteins can activate transcription in E. coli, provided that one of the interacting proteins is tethered to the DNA by a DNA-binding protein and the other is tethered to a subunit of E. coli RNAP [20,21]. Specifically, contact between a protein (or protein domain) fused to the \(\alpha\) subunit of E. coli RNAP and another protein fused to a zinc finger DNA-binding protein (referred to as Zif) activates transcription of a lacZ reporter gene (with an upstream Zif binding site) in E. coli cells (Figure 2A) [22]. Because Zif binds its cognate recognition site as a monomer, and because the \(\alpha\) subunit is monomeric in the RNAP holoenzyme complex, this configuration of the assay is ideally suited to detecting interactions between two protein monomers (i.e., dimer formation).

Our strategy for determining whether MgL\(\alpha\) can interact with Ssp\(\alpha\) (thus potentially forming a heterodimer) involved the use of two fusion proteins, one comprising MgL\(\alpha\) fused to Zif and the other comprising F. tularensis Ssp\(\alpha\) fused to the \(\alpha\) subunit of E. coli RNAP (Figure 2A). Accordingly, we fused full-length MgL\(\alpha\) (residues 1–205) to the N-terminus of Zif, and we fused full-length Ssp\(\alpha\) (residues 1–210) to the N-terminus of \(\alpha\). We then tested whether the MgL\(\alpha\)-Zif fusion protein could activate transcription from an appropriate test promoter in cells containing the Ssp\(\alpha\)-\(\omega\) fusion protein. Plasmids expressing the MgL\(\alpha\)-Zif and Ssp\(\alpha\)-\(\omega\) fusion proteins were introduced into E. coli strain KDZif1\(\Delta\omega\), which harbors a test promoter linked to lacZ on an F\(^{\prime}\) episome (Figure 2A; 22). In support of the idea that MgL\(\alpha\) and Ssp\(\alpha\) interact with one another directly, the MgL\(\alpha\)-Zif fusion protein activated transcription of the reporter gene strongly (up to \(\sim\)22-fold) in cells containing the Ssp\(\alpha\)-\(\omega\) fusion protein, whereas Zif (without the fused MgL\(\alpha\) moiety) did not (Figure 2B). The observed activation was dependent upon the ability of the \(\omega\) moiety of the Ssp\(\alpha\)-\(\omega\) fusion to interact with E. coli RNAP (unpublished data). Another control revealed that MgL\(\alpha\)-Zif did not activate transcription from the test promoter in the presence of the unrelated Gal11P-\(\omega\) fusion protein (Figure 2B).

Given that MgL\(\alpha\) and Ssp\(\alpha\) are similar to one another at the primary amino acid level, we sought to determine whether MgL\(\alpha\) or Ssp\(\alpha\) (or both) could also form homomeric complexes. We therefore constructed two additional fusion proteins, one in which full-length MgL\(\alpha\) was fused to the N-terminus of \(\omega\), and another in which full-length F. tularensis Ssp\(\alpha\) was fused to the N-terminus of Zif. Results depicted in Figure 2B show that specific interactions were also detected between the MgL\(\alpha\)-Zif and MgL\(\alpha\)-\(\omega\) fusion proteins, as well as between the Ssp\(\alpha\)-Zif and Ssp\(\alpha\)-\(\omega\) fusion proteins. Our findings suggest that MgL\(\alpha\) and Ssp\(\alpha\) can form both heteromeric and homomeric complexes.

As an additional test of whether MgL\(\alpha\) and Ssp\(\alpha\) interact with one another to form a heteromeric complex, we co-expressed genes encoding S-tagged MgL\(\alpha\) (MgL\(\alpha\)-S) and His-tagged Ssp\(\alpha\) (Ssp\(\alpha\)-His6) in E. coli and asked whether the two proteins co-purified with one another following metal chelate affinity chromatography. Specifically, we made a cell lysate from E. coli, synthesizing both Ssp\(\alpha\)-His6 and MgL\(\alpha\)-S. We
then applied the lysate to a metal chelate affinity resin to capture SspA-His6 and any associated proteins, washed the resin, and eluted proteins that were specifically bound to the resin with imidazole. Following washing, proteins specifically bound to the resin were eluted with imidazole. Cell lysates and eluted proteins were separated on a 4–12% Bis-Tris NuPAGE gel and analyzed by immunoblotting using antibodies that recognize the His-tag on SspA-His6 (anti-His; upper panel), the S-tag on MglA-S (anti-S; middle panel), and the α subunit of E. coli RNAP (anti-α; lower panel). Lane 1, lysate containing MglA-S and SspA-His6. Lane 2, lysate containing MglA-S and SspA. Lane 3, proteins purified from the lysate containing MglA-S and SspA-His6. Lane 4, proteins purified from the lysate containing MglA-S and SspA. doi:10.1371/journal.ppat.0030084.g003

Figure 3. MglA Co-Purifies with His-Tagged SspA

Immunoblot analyses of proteins co-purifying with SspA-His6. Cell lysates of E. coli containing MglA-S and SspA (i.e., without the His-tag) or MglA-S and SspA-His6 were incubated with a metal chelate affinity resin. Following washing, proteins specifically bound to the resin were eluted with imidazole. Cell lysates and eluted proteins were separated on a 4–12% Bis-Tris NuPAGE gel and analyzed by immunoblotting using antibodies that recognize the His-tag on SspA-His6 (anti-His; upper panel), the S-tag on MglA-S (anti-S; middle panel), and the α subunit of E. coli RNAP (anti-α; lower panel). Lane 1, lysate containing MglA-S and SspA-His6. Lane 2, lysate containing MglA-S and SspA. Lane 3, proteins purified from the lysate containing MglA-S and SspA-His6. Lane 4, proteins purified from the lysate containing MglA-S and SspA. doi:10.1371/journal.ppat.0030084.g003

Both MglA and SspA Regulate Virulence Gene Expression in F. tularensis

Previous studies revealed that MglA positively regulates expression of the pdpD, iglA, iglC, iglD, and pdpA genes in F. tularensis subspecies novicida (and of these, iglA, iglC, and pdpA have been implicated in intramacrophage survival and virulence) [4,5,7]. These genes are located on a pathogenicity island [7]. In subspecies novicida there is only one copy of this island [7], whereas in the highly virulent Schu S4 strain (subspecies tularensis), there are two identical copies of the pathogenicity island [15]. Similarly, the attenuated strain LVS (a derivative of one of the subspecies holarctica strains), contains two copies of the pathogenicity island. We have shown that in LVS both MglA and SspA associate with RNAP and that the association of MglA with RNAP is dependent upon SspA. We therefore reasoned that if MglA controls virulence gene expression through its association with RNAP, then the absence of SspA in an sspA mutant strain should result in decreased levels of virulence gene expression. That is to say, we would expect that both MglA and SspA are required for virulence gene expression. To compare the effect on virulence gene expression of removing either MglA or SspA in LVS, we constructed mutant strains of LVS carrying in-frame deletions of either the mglA gene (LVSΔmglA) or the sspA gene (LVSΔsspA). Because both ΔmglA and ΔsspA mutant cells had a growth defect compared to wild-type LVS cells (culture doubling times of ~171, ~182, and ~147 min, respectively, under the conditions of our experiment), we also used, as a control, cells of a mutant containing a mariner transposon in the gene FTL_0951 (LVS::FTL0951); this strain has a more severe growth defect than either the ΔmglA or ΔsspA mutant strain (with a corresponding culture doubling time of ~285 min). Cells of the wild-type LVS strain, cells of the LVS ΔmglA mutant strain, cells of the LVS ΔsspA mutant strain, and cells of the control strain LVS::FTL0951 were grown to mid-log (where, in the wild-type strain, the mglA and sspA genes are highly expressed; see Figure S1), total RNA was isolated, and expression of the iglA, iglC, and pdpA virulence genes in each of the four strains was measured using quantitative real-time reverse transcriptase (RT)–PCR.

The results depicted in Figure 4 show drastically reduced amounts of the iglA, iglC, and pdpA transcripts in both LVS
ΔmglA mutant cells and LVS ΔsspA mutant cells when compared to LVS wild-type cells. Complementation of the LVS ΔmglA and LVS ΔsspA mutant cells with plasmids expressing either the mglA or the sspA gene, respectively, restored the amounts of iglA, iglC, and pdpA transcripts to near wild-type levels (Figure S2). Moreover, the effects of the mglA and sspA deletions on virulence gene expression were not simply due to the effect of these deletions on growth rate; only a modest reduction in the amounts of the iglA, iglC, and pdpA transcripts were evident in cells of the slow growing LVS::FTL0951 mutant when compared to LVS wild-type cells (Figure 4). Thus, both MglA and SspA appear to positively control expression of the iglA, iglC, and pdpA virulence genes in LVS, though we cannot exclude the possibility of indirect effects. Because we showed that the association of MglA with RNAP is dependent on SspA (Figure 1C), we infer that MglA controls virulence gene expression through its association with RNAP. Taken together, our findings suggest that MglA and SspA interact with one another directly to form a complex that associates with RNAP to positively control expression of the iglA, iglC, and pdpA virulence genes located on the F. tularensis pathogenicity island.

MglA and SspA Co-Regulate the Same Set of Genes in F. tularensis

Recently, DNA microarray analyses have revealed that in subspecies novicida, MglA controls the expression of ~100 genes, with the vast majority being positively regulated [6]. In particular, the expression of all of the genes on the novicida pathogenicity island was found to be positively regulated by MglA. In addition, MglA was found to regulate the expression of virulence genes located outside of the pathogenicity island, as well as other genes not thought to play a role in virulence [6]. If MglA and SspA primarily function together to regulate gene expression, then we would predict that MglA and SspA should regulate expression of the same set of genes. To test this prediction on a genome-wide scale, we performed DNA microarray analyses to compare the global gene expression profiles of ΔmglA mutant cells, ΔsspA mutant cells, LVS wild-type cells, and cells of the control strain LVS::FTL0951. The results of these analyses show that in LVS, MglA and SspA regulate expression of essentially the same set of genes (Figure 5; Table 1). Specifically, we found that the expression of 30 genes changed by a factor of 3 or more when mglA was deleted, and the expression of 28 genes changed by a factor of 3 or more when sspA was deleted (Table 1). Furthermore, of the 33 genes whose expression was altered by a factor of 3 or more in either the ΔmglA or the ΔsspA mutant cells compared to wild-type cells, 25 showed at least a 3-fold change in expression in the other mutant background (Figure 5; Table 1), and all showed at least a 2-fold change in expression in the other mutant background. Because expression of five of the MglA- or SspA-regulated genes was also altered in cells of our slow growing control strain (LVS::FTL0951) when compared to wild-type cells, we do not necessarily know whether these particular genes are truly regulated by MglA or SspA; their expression might alter in the ΔmglA or ΔsspA mutants simply because these five genes are growth-rate regulated. The results of microarray analyses thus show not only that MglA and SspA co-regulate expression of essentially the same set of genes in F. tularensis, but also that MglA and SspA exert similar effects on these target genes, as might be expected if MglA and SspA were to function together primarily as a heteromeric complex.

Discussion

We have found that virulence gene expression is controlled by both MglA and SspA in F. tularensis. Specifically, we have shown that both MglA and SspA are RNAP-associated proteins in F. tularensis, and that MglA and SspA can interact with one another directly to form a heteromeric complex. Although MglA and SspA may also form homomeric complexes, we present evidence that the ability of MglA to associate with RNAP is dependent upon SspA. One implication of this finding is that members of the SspA protein family likely associate with RNAP as an oligomer, and
presumably as a dimer [8,19]. Our findings suggest that although MglA homomers might form in the cell, they are unlikely to associate with RNAP. Unfortunately, we were unable to determine whether the association of SspA with RNAP is dependent upon MglA, and so we do not currently

Brotcke and colleagues, who analyzed the effects of an mglA point mutation on gene expression in *F. novicida* [6] (see Table 1). Note that when defining the MglA regulon, we listed those genes whose expression changed by a factor of 3 or more when *mglA* was deleted. If we had defined the regulon by listing those genes whose expression changed by a factor of 2 or more, then the MglA regulon would be composed of more than 100 genes (Table S1).

Although the molecular details of how MglA exerts its effects on virulence gene expression (or on the expression of any other target gene) are not yet known, we have presented evidence that its association with RNAP plays an important, if not essential, role. For example, because expression of the *iglA*, *iglC*, and *pdpA* virulence genes is downregulated in a ΔsspA mutant, and because MglA is present but is not associated with RNAP in a ΔsspA mutant, we disfavor any model whereby MglA upregulates these genes without being associated with RNAP. Accordingly, we suggest two possible models for how MglA, together with SspA, positively controls virulence gene expression in *F. tularensis* based on association of the MglA–SspA complex with RNAP. The first model

### Table 1. Microarray Analysis of Genes Whose Expression Changes by a Factor of 3 or More in either a ΔmglA or a ΔsspA Mutant Background Compared to Wild-Type

<table>
<thead>
<tr>
<th>FTT Locus</th>
<th>FTL Locus</th>
<th>Gene Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTT0026/C0</td>
<td>FTL0951</td>
<td>3-hydroxyisobutyrate dehydrogenase</td>
</tr>
<tr>
<td>FTT0111</td>
<td>FTL0778</td>
<td>Intracellular growth locus, subunit A</td>
</tr>
<tr>
<td>FTT0112</td>
<td>FTL0777</td>
<td>Intracellular growth locus, subunit B</td>
</tr>
<tr>
<td>FTT0113</td>
<td>FTL0776</td>
<td>Intracellular growth locus, subunit C</td>
</tr>
<tr>
<td>FTT0114</td>
<td>FTL0775</td>
<td>Intracellular growth locus, subunit D</td>
</tr>
<tr>
<td>FTT0116</td>
<td>FTL0774</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0117</td>
<td>FTL0773</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0118</td>
<td>FTL0772</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0119</td>
<td>FTL0771</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0120</td>
<td>FTL0770</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0122</td>
<td>FTL0769</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0129</td>
<td>FTL0768</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0131</td>
<td>FTL0767</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0207</td>
<td>FTL0766</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0449</td>
<td>FTL0765</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0499</td>
<td>FTL0764</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0569</td>
<td>FTL0763</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0646</td>
<td>FTL0762</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0671</td>
<td>FTL0761</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0672</td>
<td>FTL0760</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0673</td>
<td>FTL0759</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0674</td>
<td>FTL0758</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0675</td>
<td>FTL0757</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0679</td>
<td>FTL0756</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0950</td>
<td>FTL0755</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT0951</td>
<td>FTL0754</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT1190</td>
<td>FTL0753</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT1213</td>
<td>FTL0752</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT1218</td>
<td>FTL0751</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT1219</td>
<td>FTL0750</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>FTT1273</td>
<td>FTL0749</td>
<td>8-amino-7-oxononanoate synthase</td>
</tr>
<tr>
<td>FTT1546</td>
<td>FTL0748</td>
<td>Major facilitator superfamily transport protein</td>
</tr>
<tr>
<td>FTT1786</td>
<td>FTL0747</td>
<td>Outer membrane protein fragment</td>
</tr>
</tbody>
</table>

Only those genes whose expression changes by a factor of 3 or more, with a p-value < 0.01, in either the ΔmglA or ΔsspA mutant background are listed. Negative values indicate genes that are positively regulated by MglA or SspA, whereas positive values indicate genes that are negatively regulated. LVS ORFs are referred to by the LVS (FTT number) and Schu S4 (FTT number) locus tags for convenience. Gene names and predicted products are as annotated by Entrez Gene at the National Center for Biotechnology Information.

*Indicates those genes whose expression also changes by a factor of 3 or more, with a p-value < 0.01, in the slow growing mutant strain LVS::FTLO951 as compared to wild-type.

doi:10.1371/journal.ppat.0030084.s001

PLoS Pathogens | www.plospathogens.org
0775 | June 2007 | Volume 3 | Issue 6 | e84
Figure 6. Models for how the MglA–SspA complex positively controls virulence gene expression in F. tularensis

(A) Contact between a DNA-bound transcription activator and the RNAP-associated MglA–SspA complex activates transcription from a virulence gene promoter. The arrow indicates the transcription start site.

(B) Contact between promoter DNA and the RNAP-associated MglA–SspA complex activates transcription from a virulence gene promoter. In these models, the virulence gene promoters are depicted as being recognized by RNAP holoenzyme containing \( \sigma^70 \). The RNAP that co-purifies with MglA-TAP appears to contain \( \sigma^70 \) in stoichiometric amounts (Figure 1B), suggesting that the promoters of MglA/SspA-dependent genes may be \( \sigma^70 \)-dependent.

doi:10.1371/journal.ppat.0030084.g006

specifies that a DNA-bound transcription activator(s) contacts the RNAP-associated MglA–SspA complex directly to stimulate transcription initiation at target promoters. Put another way, we hypothesize that the MglA–SspA complex effectively becomes a “subunit” of RNAP that serves as a target for DNA-bound transcription activators that regulate the expression of virulence genes in F. tularensis (see Figure 6A). Precedent for such a model comes from studies of bacteriophage P1 late gene expression. In particular, the P1 late promoter activator (Lpa or gp10) is a phage-encoded, sequence-specific DNA binding protein that is thought to activate P1 late gene expression through a direct interaction with E. coli SspA [11]. In this particular situation, E. coli SspA evidently functions as a co-activator of P1 late gene expression by making simultaneous contact with RNAP and DNA-bound Lpa [11]. In F. tularensis there may be Lpa-like proteins (at least functionally) that bind the MglA–SspA complex and are required for the co-activation of MglA- and SspA-dependent virulence gene expression (see Figure 6A).

Such MglA- and/or SspA-binding proteins, if they do indeed exist, would be predicted to be DNA-binding proteins that bind the promoter regions of MglA- and SspA-dependent virulence genes.

The second model specifies that the RNAP-associated MglA–SspA complex interacts directly with the promoter DNA of MglA- and SspA-dependent virulence genes to activate transcription initiation from the cognate promoters (Figure 6B). However, there is no evidence that any SspA homolog is capable of binding directly to DNA; in particular, E. coli SspA does not appear to bind P1 late promoter DNA [11]. Other models can also be envisioned. These include the MglA–SspA complex modifying the promoter-recognition properties of the associated RNAP holoenzyme without interacting directly with the DNA, or modifying the elongation properties of the associated RNAP. We note that the latter model would be unlikely if the MglA–SspA complex, like SspA in E. coli [8], could interact only with the RNAP holoenzyme and not with the core enzyme and therefore was not a component of mature elongation complexes.

The models discussed above and depicted in Figure 6 are meant only to explain how MglA and SspA positively regulate the expression of target genes. However, MglA and SspA also appear to negatively regulate the expression of at least some target genes (Table 1; [6]). Whether those genes whose expression is negatively controlled by MglA and SspA are controlled directly, or indirectly, remains to be determined.

F. tularensis is unique in that it contains two genes, designated rpoA1 and rpoA2, encoding different \( \alpha \) subunits of RNAP. This situation is without precedent; in all other bacterial genomes sequenced to date there is only one rpoA gene. The two \( \alpha \) subunits are 32% identical and 55% similar to one another at the primary amino acid level and differ from one another in regions that may be important for dimer formation, promoter recognition, and interaction with DNA-bound transcription activators. We have shown that both \( \alpha_1 \) and \( \alpha_2 \) are components of RNAP in F. tularensis. Because \( \alpha \) is a dimer in the RNAP complex, there could be as many as four different species of RNAP core enzyme in F. tularensis; one composed of \( \alpha_1 \) homodimers, one composed of \( \alpha_2 \) homodimers, and two composed of \( \alpha_1 \alpha_2 \) heterodimers (that differ from one another with respect to which \( \beta \) protomer interacts with the \( \beta \) subunit of RNAP and which \( \alpha \) protomer interacts with \( \beta' \)). However, a subset of the predicted dimerization determinants [23] present in the N-terminal domains of \( \alpha_1 \) and \( \alpha_2 \) differ from one another, raising the possibility that \( \alpha_1 \) and \( \alpha_2 \) might exclusively form either homodimers or heterodimers. The experiments we have performed do not allow us to say anything more than that both \( \alpha_1 \) and \( \alpha_2 \) are subunits of RNAP in F. tularensis. Because \( \alpha \) participates in promoter recognition through direct sequence-specific protein–DNA interactions [16,24], and because \( \alpha \) is a common target for transcription activators [25], the presence of RNAP species containing distinct \( \alpha \) subunits might significantly affect the control of gene expression in F. tularensis.

The F. tularensis RNAP complex is also unusual in that it is associated with two different members of the SspA protein family (MglA and SspA). Other bacterial pathogens where SspA homologs play roles in virulence gene control, such as Neisseria gonorrhoeae [12], Yersinia enterocolitica [13], and Vibrio cholerae [14], appear to encode only one SspA homolog. Whether the use of two SspA homologs allows for the integration of multiple environmental signals, or is related to the fact that the F. tularensis RNAP contains two distinct \( \alpha \) subunits, or both, remains to be determined.

Materials and Methods

Plasmids, strains, and growth conditions. F. tularensis subspecies holarctica strain LVS was provided by Karen Elkins (Food and Drug Administration, Rockville, Maryland, United States). LVS was grown at 37°C in modified Mueller Hinton (MH) broth (Difco, http://www.difco.com) supplemented with glucose (0.1%), ferric pyrophosphate (0.025%), and Isovitalex (2%), or on cysteine heart agar (Difco) medium supplemented with 1% hemoglobin solution (VWR, http://www.vwr.com); when appropriate, kanamycin was used for selection at 50 μg/ml. E. coli strains XL1-blue (Stratagene, http://www.stratagene.com) and DH5αF'Q (Invitrogen, http://www.invitrogen.com) were used as recipients for all plasmid constructions. E. coli strain KDZif1ΔZ was used as the reporter strain for the bacterial two-hybrid experiments. KDZif1ΔZ harbors an F' episome containing the lacT promoter derivative placZif1–01 driving expression of a linked lacZ
Plasmid pACTR-AP-Zif directs the synthesis of Zif, the zinc finger DNA-binding domain of the murine Zif268 protein, and has been described previously [22]. This plasmid can be used to create fusions to the N-terminus of Zif; proteins are fused to Zif via a nine-amino acid linker (Ala-Ala-Ala-Pro-Arg-Val-Arg-Thr-Gly) specified in part by a Nosi restriction site. Plasmid pACTR-AP-Zif confers resistance to the antibiotic kanamycin (50 μg/ml) and ribostamycin (0.025 μg/ml). Plasmids pACTR-MglA-Zif and pACTR-SspA-Zif direct the synthesis of either full-length MglA (residues 1–205) or full-length LVS SspA (residues 1–210) fused to Zif. The pACTR-Zif fusion plasmids were made by cloning the appropriate Ndel-NotI-digested PCR products into Ndel-NotI-digested pACTR-AP-Zif. Each of the respective Zif fusion proteins is therefore under the control of an IPTG-inducible lacUV5 promoter.

**TAP.** Cells were grown at 37 °C with aeration in 200 ml of MH broth supplemented with glucose (0.1%), ferric pyrophosphate (0.01%), and Isovitalex (2%). Cells were harvested at an O.D.₆₀₀ of ~0.4 and harvested by centrifugation at 4 °C. TAP was then performed essentially as described earlier [18].

**Bacterial two-hybrid assays.** Cells were grown with aeration at 37 °C in LB supplemented with kanamycin (50 μg/ml), carbenicillin (100 μg/ml), tetracycline (10 μg/ml), and IPTG at the concentration indicated. Cells were permeabilized with SDS-CHCl₃ and assayed for β-galactosidase activity as described previously [29]. Assays were performed at least three times in duplicate on separate occasions.†

**Expression Values.** Time points were determined from one experiment; duplicate measurements differed by less than 10%.

**Co-purification assays.** Plasmid pETDuetMglA-S directs the synthesis of MglA with a C-terminal S-tag (MglA-S) and confers resistance to carbenicillin, while plasmid pRSFDuetSspA-His6 directs the synthesis of SspA with a C-terminal hexa-histidine tag (SspA-His6) and confers resistance to kanamycin. Plasmid pETDuetMglA-S directs the synthesis of wild-type SspA and confers resistance to kanamycin. Plasmid pETDuetMglA-S was introduced into E. coli strain BL21-CodonPlus(DE3)-RP (Stratagene) along with pRSFDuetSspA-His6 or pRSFDuetSspA. Single colonies were used to inoculate Overnight Express Instant TB medium (Novagen, www.emdbiosciences.com/html/NVG/ home.html) supplemented with kanamycin (50 μg/ml), carbenicillin (100 μg/ml), and chloramphenicol (25 μg/ml), and cultures were grown overnight at 30 °C with aeration to allow for induction. Cell pellets from 50 ml of overnight culture (OD₆₀₀ of ~3.0) were resuspended in 1 ml of binding buffer (10 mM imidazole; 10 mM Tris-HCl [pH 8.0]; 300 mM NaCl; 10% glycerol; and 1X EDTA-free protease inhibitor cocktail [Roche, http://www.roche.com]). Each 1-ml suspension was separated into 500 μl aliquots, and 100 μl of lysozyme (10 mg/ml) was added to each aliquot. Cells were then incubated on ice for 1 h and lysed by sonication, and cell debris was removed from each lysate by centrifugation. Lysates were then added to TALON polyhistidine-tag purification resin (Clontech, http://www.clontech.com), and following washing with binding buffer, proteins specifically bound to the resin were eluted in buffer containing 100 mM imidazole; 10 mM Tris-HCl [pH 8.0]; 300 mM NaCl; 10% glycerol; and 1X EDTA-free protease inhibitor cocktail.

**Immunoblots.** Cell lysates and eluted proteins were separated by SDS-PAGE on 4%–12% Bis-Tris NuPAGE gels in MOPS running buffer (Invitrogen) and transferred to nitrocellulose using the iBlot dry blotting system (Invitrogen). Membranes were blocked with 25 ml of SuperBlock Blocking Buffer (Pierce, http://www.piercenet.com) in TBS with 250 ml Surfact-Amps 20 (Pierce). Membranes were then probed with monoclonal antibodies against the S-Tag (Novagen), the His-Tag (Novagen), or the α subunit of E. coli RNAP (NeoClone, http://www. neoclone.com). Lanes were detected by using goat polyclonal anti-mouse IgG conjugated with horseradish peroxidase (Pierce), and visualized using SuperSignal West Pico Chemiluminescent Substrate (Pierce).

**RNA isolation and quantitative RT-PCR.** Cells were grown with aeration at 37 °C in MH broth supplemented with glucose (0.1%), ferric pyrophosphate (0.025%), and Isovitalex (2%). RNA isolation and cDNA synthesis were essentially as described previously for Pseudomonas aeruginosa [30]. Transcript quantities for g3g, g1q, and pspA were determined relative to the amount of rpoS (transcript by RT-PCR with the iTaq SYBR Green kit (Bio-Rad, http://www.bio-rad.com)) and then normalized to the expression of an ABI Prism 7000 Sequence Detection System (Applied Biosystems, http://www.appliedbiosystems.com); the REST 2005 software package was used for statistical analysis of the data [31].

**Array design and construction.** ORFs for LVS were predicted using GLIMMICK (using Ndel-NotI restriction sites) and Ndel-NotI-digested pACTR-AP-Zif. Each of the respective fusion proteins is therefore under the control of an IPTG-inducible lacUV5 promoter.
being those where adjacent genes were fully contained within predicted Schu S4 ORFs.

The DNA microarray consisted of sequence fragments constructed by PCR. Primers for PCR were designed using PRIMER3 ([33]; http:// primer3.sourceforge.net) software using the genome annotation. The 1,932 primer sets produced amplicons of 42 bp to 400 bp in length centered within each gene. They correspond to sequences for 1,678 LVS ORFs with Schu S4 homologs (including the 15 duplicated ORFs in the pathogenicity island), 24 GLMMER-predicted ORFs with no homology to Schu S4, and 230 intergenic regions. Greater than 95% of LVS non-transposable genes were represented on the microarray. The primer sets contained 5’ extensions to each forward and reverse primer, GGAATTCAGAGCAG and CGGACAC-TAGTCCTC, respectively.

All amplifications were performed using a standard protocol. In the first reaction, 20-µl PCR reactions containing 1 ng of LVS genomic DNA as template, 20 pmol of unique primers, and 5 units of Taq polymerase (Takara, http://www. takarabio.usa.com). These reactions were then diluted 1:150. A second round contained 1 µl of the diluted first reaction as template, 20 pmol of universal primers containing 5’ amino modification including a 3’-carbon linker (GAAGCTATCCGATGACAGCAG and GAATTC- CACCAGCAGTAGCAGG), and 5 units of Taq polymerase in a total volume of 100 µl. Aliquots (5 µl) of each first and second round reaction were run on 2% agarose gels to confirm the presence of appropriately sized products. The PCR products were purified using 96-well QiaQuick Multiswell PCR Purification plates (Qiagen, http:// www.qiagen.com), eluted in deionized water, and transferred to 384-well plates. Purified products were air-dried in a biosafety cabinet at room temperature and resuspended in 50 mM sodium phosphate (pH 8.5). Each PCR amplicon was printed in duplicate onto CodeLink Activated Glass slides (GE Healthcare, http://www.amersham.com) using a Qarrayer (Genetix, http://www.genetix.com). The microarrays were then post-processed according to the slide manufacturer’s instructions.

Microarray experiments. Strains were grown to mid-log in supplemented MH medium and RNA was isolated from 15 ml of culture using Qiagen RNeasy Midi columns. Samples were DNase-treated twice: on the columns using a Qiagen RNase-Free DNase Set, and after purification using RQ1 DNase (Promega, http://www. promega.com) and Bio-Rad RNaseout. For each strain, cDNA was synthesized using 1 µg of total RNA treated twice: on the columns using a Qiagen RNase-Free DNase Set, and after purification using RQ1 DNase (Promega, http://www.p promega.com) and Bio-Rad RNaseout, and a random primer, and then labeled with Cy5 dye (Amersham; Cy5 mono-Reactive Dye Pack) in 0.1 M sodium carbonate (pH 9.3), and cDNA from LVS (strain: FTL_1185; YP_513056.1), and SspA (FTL_0616; YP_513375.1). The cDNA was labeled with Cy5 dye (Amersham; Cy5 mono-Reactive Dye Pack) in 0.1 M sodium carbonate (pH 9.3), and cDNA from LVS (strain: FTL_1185; YP_513056.1), and SspA (FTL_0616; YP_513375.1). The Gene Expression (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=GeneExpression) accession number for the nucleotide sequence of the L. monocytogenes genome is NC_007880.1.

Accession Numbers

Proteins discussed in this study, followed by their corresponding accession numbers, are as follows: 21 (FTL_0261; WP_513056.1), 13 (FTL_0616; WP_513375.1), MglA (FTL_1185; WP_513056.1), and SspA (FTL_0616; WP_513375.1). The Entrez Genome (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=GeneExpression) accession number for the nucleotide sequence of the F. tularensis LVS genome is NC_007880.1.

Acknowledgments

We thank Karen Elkins, Simon Dillon, Shite Sebastian, and Igor Golovlev for plasmids and strains, Sandra Castang for help with protein purification, Ross Tomaino for MSMS analysis, Renate Hellmisch for artwork, and Ann Hochschild for discussions and comments on the manuscript. We thank Karen Elkins, Simon Dillon, Shite Sebastian, and Igor Golovlev for plasmids and strains, Sandra Castang for help with protein purification, Ross Tomaino for MSMS analysis, Renate Hellmisch for artwork, and Ann Hochschild for discussions and comments on the manuscript.

Author contributions. JCC and SLD conceived and designed the experiments. JCC and MMCH performed the experiments. JCC, MMCH, and SLD analyzed the data. ELB, DHB, and EJR contributed reagents/materials/analysis tools. JCC, MMCH, ELB, EJR, and SLD wrote the paper. Funding. This work was supported by National Institutes of Health grants AI06495 and AI069219 (to SLD), and AI06296 (to EJR).

Competing interests. The authors have declared that no competing interests exist.

References


www.moleculardevices.com) and data were analyzed with GeneSpring GX.

Supporting Information

Figure S1. Effect of Growth Phase on Abundance of the mglA and sspA Transcripts in F. tularensis

Quantitative RT-PCR analysis of mglA and sspA transcript levels in cells of the LVS wild-type strain at different phases of growth. The mid-logarithmic (mid-log), late-log, and stationary phase cultures had OD600’s of 0.4, 1.0, and 2.6, respectively. Transcripts were normalized to tufL whose expression remains relatively constant throughout the growth curve; compared to cells in mid-log, cells in late-log and stationary phase had 1.03 times (standard error [SE] = 0.21) and 1.64 times (SE = 0.12) the number of tufL transcripts, respectively, as determined by quantitative RT-PCR.

Found at doi:10.1371/journal.ppat.0030084.s001 (136 KB PDF).

Figure S2. Complementation of the Effects of the ΔmglA and ΔsspA Mutations on Virulence Gene Expression in F. tularensis

Quantitative RT-PCR analysis of iglC, iglD, and pdpA transcript levels in wild-type, ΔmglA, and ΔsspA mutant cells containing the indicated plasmids. Transcripts were normalized to tufL whose expression is not influenced by MglA or SspA. Plasmids pF-MglA and pF2-SspA direct the synthesis of MglA and SspA, respectively. Plasmids pF and pF2 served as empty vector controls for plasmids pF-MglA and pF2-SspA, respectively.

Found at doi:10.1371/journal.ppat.0030084.s002 (548 KB PDF).

Table S1. Microarray Analysis of Genes Whose Expression Changes by a Factor of 2 or More in either a ΔmglA or a ΔsspA Mutant Background Compared to Wild-Type

Found at doi:10.1371/journal.ppat.0030084.s001 (247 KB DOC).


