Genetics of germination-arrest factor (GAF) production by *Pseudomonas fluorescens* WH6: identification of a gene cluster essential for GAF biosynthesis

Anne Halgren,1 Maciej Maselko,1 Mark Azevedo,1 Dallice Mills,2 Donald Armstrong2 and Gary Banowetz1

1USDA-ARS National Forage Seed Production Research Center, Corvallis, OR 97331, USA
2Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA

The genetic basis of the biosynthesis of the germination-arrest factor (GAF) produced by *Pseudomonas fluorescens* WH6, and previously identified as 4-formylaminoxyvinylglycine, has been investigated here. In addition to inhibiting the germination of a wide range of grassy weeds, GAF exhibits a selective antimicrobial activity against the bacterial plant pathogen *Erwinia amylovora*. We utilized the in vitro response of *E. amylovora* to GAF as a rapid screen for loss-of-function GAF phenotypes generated by transposon mutagenesis. A Tn5 mutant library consisting of 6364 WH6 transformants was screened in this *Erwinia* assay, resulting in the identification of 18 non-redundant transposon insertion sites that led to loss of GAF production in WH6, as confirmed by TLC analysis. These insertions mapped to five different genes and four intergenic regions. Three of these genes, including two putative regulatory genes (*gntR* and *iopB* homologues), were clustered in a 13 kb chromosomal region containing 13 putative ORFs. A GAF mutation identified previously as affecting an aminotransferase also maps to this region. We suggest that three of the genes in this region (a carbamoyltransferase, an aminotransferase and a formyltransferase) encode the enzymes necessary to synthesize dihydroGAF, the putative immediate precursor of GAF in a proposed GAF biosynthetic pathway. RT-qPCR analyses demonstrated that mutations in the *gntR* and *iopB* regulatory genes, as well as in a *prtR* homologue identified earlier as controlling GAF formation, suppressed transcription of at least two of the putative GAF biosynthetic genes (encoding the aminotransferase and formyltransferase) located in this 13 kb region.

**INTRODUCTION**

We have previously shown that a naturally occurring herbicide is produced by particular isolates of rhizosphere bacteria obtained from soils of the Willamette Valley, Oregon (Banowetz et al., 2008, 2009; Armstrong et al., 2009; McPhail et al., 2010). This compound, which we have termed a germination-arrest factor (GAF), specifically and irreversibly blocks germination of the seeds of a large number of grassy weed species without significantly affecting the growth of established grass seedlings or mature plants. The biological and herbicidal properties of GAF have been described, and we have developed a quantitative bioassay for the herbicide based on its ability to arrest germination of the seeds of annual bluegrass (*Poa annua* L.), one of our target weed species (Banowetz et al., 2008). In addition to its herbicidal properties, GAF was also found to exhibit a selective antimicrobial activity against *Erwinia amylovora*, the bacterial plant pathogen that causes fireblight in orchard crops (Halgren et al., 2011).

TLC of extracts of culture filtrates from the GAF-producing bacterial isolate *Pseudomonas fluorescens* WH6 revealed that GAF activity is associated with a particular ninhydrin-reactive compound (Armstrong et al., 2009). We were subsequently able to purify and identify this compound as 2-amino-4-formylaminoxy-3-butenonic acid (4-formylaminoxyvinylglycine) (Fig. 1), a previously undescribed member of a small group of naturally occurring compounds known as oxyvinylglycines (McPhail et al., 2010). Oxyvinylglycines are known to block reactions catalysed by enzymes dependent upon pyridoxal phosphate as a co-factor (Berkowitz et al.,...
approach, two additional regulatory genes were identified that are controlled by the three putative GAF regulatory genes. Real-time quantitative PCR (RT-qPCR) analysis demonstrated that two genes that encode enzymes involved, directly or indirectly, in GAF biosynthesis have been identified. The recent sequencing of the WH6 genome (Kimbrel et al., 2010) enabled the sites of the mutations in these strains to be mapped. In the mutant strain WH6-2, the Tn5 insertion disrupted a gene that encodes a putative aminotransferase, suggesting that this gene plays a direct role in GAF biosynthesis. The Tn5 insertion in mutant strain WH6-3 disrupted a gene (prtR) that is similar to a putative activator of a sigma factor in Pseudomonas fluorescens LS107d2 (Burger et al., 2000), which is consistent with a role for this gene in the regulation of GAF biosynthesis.

The discovery that GAF has selective antimicrobial activity against E. amylovora (Halgren et al., 2011) indicated that the response of E. amylovora could be used as the basis for a rapid and potentially saturating screen for genes affecting GAF production. The results of a GAF mutational search using an E. amylovora screen are reported here. With this approach, two additional regulatory genes were identified that function, together with the previously identified prtR gene, to regulate the GAF biosynthetic pathway. Moreover, a number of genes that encode enzymes that appear to be involved, directly or indirectly, in GAF biosynthesis have been identified. Real-time quantitative PCR (RT-qPCR) analysis demonstrated that two genes that encode enzymes likely to be directly involved in the GAF biosynthetic pathway are controlled by the three putative GAF regulatory genes.

**METHODS**

**Bacterial strains and growth media.** E. amylovora 153 was obtained from Dr Joyce Loper [United States Department of Agriculture–Agricultural Research Service (USDA-ARS) Horticultural Crops Research Laboratory, Corvallis, OR]. The origin and characterization of P. fluorescens strain WH6 have been described previously (Banowetz et al., 2008; Elliott et al., 1998). All strains used in this study were maintained at −80 °C in Luria–Bertani liquid medium (LB medium) (Sambrook & Russell, 2001) with a final concentration of 15% (v/v) glycerol.

The transposon vector pUTmini-Tn5gfpr in its host Escherichia coli S17-1 Δ pir (Matthysse et al., 1996) was obtained as a gift from Ann Matthysse (Department of Biology, University of North Carolina). The mini-Tn5 derivative contains the inner and outer transposase recognition sequences flanking a promoterless GFP gene and a tetracycline resistance cassette.

**Tn5 mutagenesis.** Both P. fluorescens WH6 and Escherichia coli S17-1 Δ pir containing the transposon delivery vector pUTmini-Tn5gfpr were grown in 4 ml LB medium overnight with shaking (200 r.p.m.) at 28 and 37 °C, respectively. The Escherichia coli medium was supplemented with 10 μg tetracycline ml⁻¹ in order to maintain the vector. A 1.5 ml aliquot of each culture was centrifuged at 3330 g for 2 min, and the cell pellet was then washed and resuspended in 500 μl PBS (Sambrook & Russell, 2001). Both cultures were combined into one tube, spread onto an LB plate containing no antibiotic, and left at 28 °C overnight. The following day, the bacterial lawn was collected from the plate using 3 ml PBS, washed, and streaked onto ‘925 minimal medium’ plates (Halgren et al., 2011) containing 10 μg tetracycline ml⁻¹, in order to select against Escherichia coli and nontransformants.

**Selection of GAF mutants via the E. amylovora screen.** E. amylovora cultures were grown, diluted and plated as previously described (Halgren et al., 2011). Eighteen colonies of transformed P. fluorescens WH6 were spotted with a sterile toothpick onto each plate freshly spread with an E. amylovora lawn. Plates were incubated at 28 °C and examined for antibiotic 48 h after sample application. Colonies lacking a zone of inhibition were selected and retested against E. amylovora, streaked onto plates of LB medium containing 10 μg tetracycline ml⁻¹ and cultured for cryopreservation as described above.

**Culture filtrate production.** Each transformant that had lost the ability to inhibit growth of E. amylovora was inoculated into the modified Pseudomonas minimal salts medium (PMS medium), cultured and harvested as described previously (Banowetz et al., 2008). To prepare culture filtrates, the Pseudomonas culture fluid recovered from 7-day cultures was centrifuged (3000 g, 15 min), and the supernatant was passed through a bacteriological filter (Millipore GP Express SteriTop, 0.22 μm pore size). The resulting sterile culture filtrate was stored at 4 °C.

**TLC analysis and bioassay of GAF activity.** TLC analyses of culture filtrates from mutant bacterial cultures were performed on 90% ethanol extracts of the solids from dried culture filtrates as described previously (Armstrong et al., 2009; Halgren et al., 2011). Silica GHL and microcrystalline cellulose TLC plates were purchased from Analtech.

Bioassays for GAF activity in bacterial culture filtrates were performed with annual bluegrass seeds (ABG, Poa annua L) using the standard Poa germination bioassay protocol and scoring system described by Banowetz et al. (2008). In this scoring system, a score of 4 represents normal germination and plumule development, and a score of 1 represents complete germination arrest immediately after emergence of the radicle and coleoptile. Tests of culture filtrates for antimicrobial activity against E. amylovora were performed as described by Halgren et al. (2011).

**DNA extraction, digestion by restriction enzymes, and self-ligation.** DNA was isolated from mutant strains of P. fluorescens...
WH6 using the ZR Fungal/Bacterial DNA kit (Zymo Research). Purity and concentration were determined using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific).

DNA from putative GAF mutant colonies was digested with EcoRI (NarI), which cuts frequently within the pUT-miniTn5gfp delivery vector and within the P. fluorescens WH6 genome. BstSI was used as a secondary restriction enzyme when EcoRI digests yielded only extremely short inverse PCR (iPCR) products. After the enzyme was inactivated, approximately 12 ng linear digested DNA was circularized by self-ligation using T4 ligase (Fermentas) in a 50 μl reaction mixture at 25 °C for 1 h. The digestion and self-ligation were performed according to the manufacturer’s instructions. The self-ligated DNA was purified using the QIAquick PCR Purification kit (Qiagen). The clean ligation products were then used as template for amplification of transposon-flanking DNA sequences in an iPCR reaction.

iPCR and sequencing. All PCR amplifications were conducted in a Robocycler (Stratagene) thermocycler with Platinum Taq polymerase using the manufacturer’s instructions (Invitrogen). PCR was performed using outward-facing, transposon-specific primers, and iPCR products were sequenced with a nested transposon-specific primer. Primers used to amplify the products of the NarI digest were those developed by Lewenza et al., whereas primers used for amplifying BstII digest products were as noted in Table S1 available with the online version of this paper. Thermocycling consisted of an initial denaturation for 5 min at 94 °C, followed by 40 cycles (30 s, 94 °C, 30 s, 56 °C, 2 min, 72 °C), with a final 10 min extension step at 72 °C. Prior to cycling, samples were heated at 94 °C for 5 min and the extension step was increased to 10 min, 72 °C, as part of the terminal cycle. The samples were subjected to electrophoresis through agarose gels containing 10 μg ethidium bromide ml⁻¹, and amplicons were visualized by exposure to UV light. All samples yielded a single major band were selected for clean-up using either the QIAquick PCR Purification kit or QIAquick Gel Extraction kit (Qiagen). Sequencing was carried out at the Central Services Lab of the Center for Genome Research and Biocomputing (CGRB; Oregon State University, Corvallis, OR). PCR products were identified using BLAST (Altschul et al., 1997), and the region flanking the transposon insertion site was aligned with the annotated genome sequence of P. fluorescens WH6 (Kimbel et al., 2010).

RNA extraction, reverse transcription (RT) and RT-qPCR. Mutant strains WH6-3, WH6-4 and WH6-7 (each containing a Tn5 insertion in a different regulatory gene), and wild-type WH6, were grown as overnight cultures in 5 ml PMS medium at 28 °C for 16 h. The following morning, cultures were diluted 1 : 4 into fresh medium and divided into three replicate cultures with an OD600 of approximately 0.13 each. The bacterial RNA was harvested after 4.5 h, giving a V value of 0.146 (Fig. S2).

These two reference genes had M values of less than 0.3 (Fig. S1) and were expressed reference genes, and recommended normalization of data to a combination of these two genes as a valid endogenous control. geNorm identified the genes encoding sigma factor RpoD as the most stably expressed reference genes, and recommended normalization of data to a combination of these two genes as a valid endogenous control. These two reference genes had M values of less than 0.3 (Fig. S1) and gave a V value of 0.146 (Fig. S2).

Data analysis. All post-run relative expression analyses were performed using REST 2009 (Qiagen) (Pfaffl et al., 2002), which included an efficiency correction for each primer set. For data analysis, the Cq values of the genes were converted to relative quantities and normalized using the geometric mean of the two reference genes.

RESULTS

Mutant yield using the E. amylovora screen

A total of 6364 P. fluorescens WH6 transformants was obtained by Tn5 transposon mutagenesis. From these transformants, 23 putative GAF mutants were identified

no-reverse-transcriptase controls to screen for DNA contamination. Real-time PCR was performed using a Multicolor Real-Time PCR Detection system (model IQ5; Bio-Rad Laboratories) in 20 μl reaction mixtures with iQ SYBR green Supermix (Bio-Rad Laboratories). Cycling conditions were 95 °C for 3 min, and 40 cycles of 95 °C for 10 s and 57 °C for 30 s. Fluorescence readings were recorded after each cycle. A final melting analysis was obtained by slow heating, with 10 s increments of 0.5 °C from 57 to 95 °C, and fluorescence collection at intervals of 0.5 °C. The threshold cycle (Cq) value of each sample was determined during the exponential phase of amplification. Each dissociation curve trace was checked to ensure that no non-specific products were present in the amplification.

Primer design, optimization and validation. Primers for both RT-qPCR genes of interest and candidate reference genes were designed using Primer3Plus software (Untergasser et al., 2007) (Table S1) from Sigma-Aldrich. Four candidate reference genes were selected for evaluation: proc, rpoD, rpsL and tufB. proc (encoding pyruvyl-5-carboxylate reductase) and rpsL (encoding sigma factor RpoD) have been described as stable internal control genes in Pseudomonas aeruginosa (Savil et al., 2003). The rpsL gene, which encodes the 30S ribosomal protein S12, has also been used as a reference gene in P. aeruginosa (Gooderham et al., 2009). tufB, which encodes the elongation factor Tu (EF-Tu), has been used in our laboratory as a reference gene for P. fluorescens (J. A. Kimbel, personal communication).

Prior to validating reference genes and performing expression studies, each primer set (designed for either reference genes or genes of interest) was tested for efficiency against a standard curve, consisting of five fivefold serial dilutions, covering a range from 15.4 ng to 4.9 pg using WH6 cDNA. For each primer pair, amplification efficiency was calculated as $E = -1 + 10^{-1/\text{slope}}$. At least one sample from each primer pair was analysed by agarose gel electrophoresis to verify that the product was a single band of the correct size. Additionally, one sample from each primer pair was sequenced to verify its identity.

To validate the candidate reference genes for use under the conditions of the current study, each candidate primer set was tested against cDNA representing WH6 and each of the mutants used in the analysis. All candidate reference genes were quantified on the same batch of cDNA as used in the experimental analysis. For these validation experiments, triplicate sets of each cDNA were amplified using the same conditions described above. The Cq was averaged for each triplicate and transformed to quantities relative to the sample with the highest expression. These values were imported into geNorm v3.5 software (Vandesompele et al., 2002) for analysis of gene expression stability and subsequent confirmation of optimal reference genes. geNorm identified the genes tufB and rpsL as the most stably expressed reference genes, and recommended normalization of data to a combination of these two genes as a valid endogenous control. These two reference genes had M values of less than 0.3 (Fig. S1) and gave a V value of 0.146 (Fig. S2).

DNA purity and concentration were determined using a NanoDrop ND-1000 spectrophotometer. The RNA extraction, reverse transcription (RT) and RT-qPCR. For each primer set (designed for either reference genes or genes of interest) was tested for efficiency against a standard curve, consisting of five fivefold serial dilutions, covering a range from 15.4 ng to 4.9 pg using WH6 cDNA. For each primer pair, amplification efficiency was calculated as $E = -1 + 10^{-1/\text{slope}}$. At least one sample from each primer pair was analysed by agarose gel electrophoresis to verify that the product was a single band of the correct size. Additionally, one sample from each primer pair was sequenced to verify its identity.

To validate the candidate reference genes for use under the conditions of the current study, each candidate primer set was tested against cDNA representing WH6 and each of the mutants used in the analysis. All candidate reference genes were quantified on the same batch of cDNA as used in the experimental analysis. For these validation experiments, triplicate sets of each cDNA were amplified using the same conditions described above. The Cq was averaged for each triplicate and transformed to quantities relative to the sample with the highest expression. These values were imported into geNorm v3.5 software (Vandesompele et al., 2002) for analysis of gene expression stability and subsequent confirmation of optimal reference genes. geNorm identified the genes tufB and rpsL as the most stably expressed reference genes, and recommended normalization of data to a combination of these two genes as a valid endogenous control. These two reference genes had M values of less than 0.3 (Fig. S1) and gave a V value of 0.146 (Fig. S2).

Data analysis. All post-run relative expression analyses were performed using REST 2009 (Qiagen) (Pfaffl et al., 2002), which included an efficiency correction for each primer set. For data analysis, the Cq values of the genes were converted to relative quantities and normalized using the geometric mean of the two reference genes.

RESULTS

Mutant yield using the E. amylovora screen

A total of 6364 P. fluorescens WH6 transformants was obtained by Tn5 transposon mutagenesis. From these transformants, 23 putative GAF mutants were identified...
using the E. amylovora screen, and 18 of these mutations occurred at non-redundant insertion sites (Table 1). Colonies of two of these mutants, WH6-10 and WH6-12, exhibited slight antimicrobial activity in the E. amylovora screen, but culture filtrates from all mutants were consistently antibiosis-deficient against E. amylovora (Table 2). Loss of GAF production in the mutant culture filtrates was confirmed in all 23 mutants by tests in the Poa germination bioassay for GAF activity (Table 2) and by TLC analysis (Figs S3 and S4). iPCR analysis revealed that these 23 mutations involved five unique ORFs and four intergenic regions. The locations of these mutations and the two separate ORF mutations obtained in our previous screening efforts were determined from the draft of the complete WH6 genome obtained by Kimbrel et al. (2010) and are illustrated in Fig. 2. The annotations of these mutations, as derived from both this source and from independent iPCR analyses, are summarized in Table 1.

### Identification of regulatory genes controlling GAF biosynthesis

Our previous mutant search using the Poa germination bioassay as a primary screen resulted in the identification of one regulatory gene (prtR, PFWH6_3687) controlling GAF biosynthesis. In the current screen, two genes with homology to known regulatory elements were among the five unique ORFs identified as essential for GAF biosynthesis. One of these genes, PFWH6_5251 in mutant WH6-7, was disrupted by Tn5 insertions at five independent loci (Table 1, Fig. 3). This gene appears to be a homologue of a gene in Pseudomonas syringae pv. maculicola strain ES4326 that is annotated as producing a hypothetical protein (79% identity). However, the next most closely related gene (65% identity) is a gene designated iopB, which regulates the production of phenazine 1-carboxamide in P. chlororaphis strain PCL1391 (Van Rij, 2006). In strain PCL1391, this gene occurs downstream of another regulatory gene designated iopA. The putative iopB gene in WH6 also occurs downstream of an iopA homologue (PFWH6_5249). In both PCL1391 and WH6, the iopA and iopB genes are separated by a small putative ORF (PFWH6_5250 in WH6).

The second putative regulatory element identified in the Erwinia screen occurs immediately upstream of the iopA homologue in WH6, but in the opposite orientation. This gene, PFWH6_5248 in mutant WH6-4, was disrupted by Tn5 insertion at four independent loci (Table 1, Fig. 3). This gene encodes a putative GntR transcriptional regulator. GntR transcriptional regulators constitute a large family of regulatory proteins that contain both DNA-binding

### Table 1. Mutations causing loss of GAF production in P. fluorescens WH6

<table>
<thead>
<tr>
<th>Mutation designation</th>
<th>ORF site of transposon insertion*</th>
<th>Putative ORF product</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WH6-2::Tn5†</td>
<td>PFWH6_5256</td>
<td>Aminotransferase</td>
<td>ZP_07777818</td>
</tr>
<tr>
<td>WH6-3::Tn5†</td>
<td>PFWH6_3687</td>
<td>PrtR regulator</td>
<td>ZP_07776266</td>
</tr>
<tr>
<td>WH6-4A::Tn5</td>
<td>PFWH6_5248</td>
<td>GntR regulator</td>
<td>ZP_0777810</td>
</tr>
<tr>
<td>WH6-4N::Tn5</td>
<td>PFWH6_5719</td>
<td>HemY protein</td>
<td>ZP_07778272</td>
</tr>
<tr>
<td>WH6-4AA::Tn5</td>
<td>PFWH6_5251</td>
<td>IopB regulator</td>
<td>ZP_0777813</td>
</tr>
<tr>
<td>WH6-6C::Tn5</td>
<td>5’ of PFWH6_4690</td>
<td>HIT protein</td>
<td>ZP_07777256</td>
</tr>
<tr>
<td>WH6-7D::Tn5</td>
<td>5’ of PFWH6_0984</td>
<td>Hypothetical protein</td>
<td>ZP_0777306</td>
</tr>
<tr>
<td>WH6-7R::Tn5</td>
<td>PFWH6_1695</td>
<td>dGTPase</td>
<td>ZP_07774306</td>
</tr>
<tr>
<td>WH6-7U::Tn5</td>
<td>PFWH6_5254</td>
<td>Carbamoyltransferase</td>
<td>ZP_07777816</td>
</tr>
<tr>
<td>WH6-7DD::Tn5</td>
<td>5’ of PFWH6_5556</td>
<td>ThiG protein</td>
<td>ZP_07778113</td>
</tr>
<tr>
<td>WH6-7CC::Tn5</td>
<td>5’ of PFWH6_5255</td>
<td>Unknown</td>
<td>ZP_07777817</td>
</tr>
</tbody>
</table>

*ORF designations are based on the WH6 draft genome published by Kimbrel et al. (2010).
†Mutations identified in earlier screen using the Poa bioassay (Armstrong et al., 2009).
‡,§,¶,|| Redundant insertion sites.
helix–turn–helix domains and aminotransferase domains (Haydon & Guest, 1991; Hoskisson & Rigali, 2009). Regulators of this family possess a conserved N-terminal domain that is involved in the DNA binding. Based on homology, the specific GntR regulator disrupted in our screen belongs to the MocR-like subfamily, which contains an aminotransferase domain that requires pyridoxal 5'-phosphate as a co-factor (Sung et al., 1991; Rigali et al., 2002). The closest homologue is a gene from Marinobacter sp. ELB1, with which it shares 90% identity. The gntR homologue in WH6 and the iopA and iopB homologues are on opposite strands of the chromosome and separated by an intergenic region with two predicted promoters that also occur on opposite strands of the chromosome (Fig. 3).

### Identification of a gene cluster affecting GAF biosynthesis

The regulatory genes iopB and gntR reside within a 13 kb region of the *P. fluorescens* WH6 genome that includes two other genes (PFWH6_5254 and PFWH6_5256) and an intergenic region whose disruption by Tn5 insertion resulted in loss of GAF production (Table 1, Fig. 3). The WH6-2::Tn5 mutation identified in our earlier screen using the Poa bioassay system also occurs within this 13 kb region. The latter mutation disrupted a putative aminotransferase gene (PFWH6_5256) (Armstrong et al., 2009; Kimbrel et al., 2010). Immediately downstream of this gene is a gene (PFWH6_5527) encoding a putative formyltransferase. The proximity of these two genes suggests that they both may play a role in the synthesis of the formylamino side-chain of GAF (see the structure illustrated in Fig. 1).

In our current screen, two additional types of mutations leading to loss of GAF production were observed in this 13 kb region. Multiple mutations were observed in PFWH6_5254 (mutations labelled WH6-11::Tn5). This gene encodes a putative carbamoyltransferase, which has a closely related homologue (85% identity) in *Pseudomonas chlororaphis* (Van Rij, 2006). Tn5 insertions resulting in loss of GAF production occurred at three independent sites

<table>
<thead>
<tr>
<th>Source of culture filtrate</th>
<th>Poa germination assay (score ± SEM)*</th>
<th>E. amylovora assay area of zone of inhibition (cm² ± SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WH6 (wild-type)</td>
<td>1.0 ± 0.0</td>
<td>15.1 ± 0.2</td>
</tr>
<tr>
<td>Water control</td>
<td>4.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>Culture medium</td>
<td>3.8 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-2</td>
<td>3.8 ± 0.1</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-3</td>
<td>3.7 ± 0.1</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-4A</td>
<td>3.0 ± 0.1</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-4N</td>
<td>3.6 ± 0.3</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-4AA</td>
<td>3.6 ± 0.3</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-4BB</td>
<td>4.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-6C</td>
<td>3.1 ± 0.1</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-7D</td>
<td>3.2 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-7R</td>
<td>3.5 ± 0.1</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-7U</td>
<td>4.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-7DD</td>
<td>3.9 ± 0.1</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-7CC</td>
<td>3.9 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-8F</td>
<td>3.3 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-9H</td>
<td>3.3 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-10K</td>
<td>3.2 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-11L</td>
<td>3.2 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-11M</td>
<td>3.4 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-11O</td>
<td>3.6 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-11P</td>
<td>3.2 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-11T</td>
<td>3.8 ± 0.2</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-11W</td>
<td>4.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-11X</td>
<td>4.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-12Q</td>
<td>3.1 ± 0.1</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-15V</td>
<td>4.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>WH6-15Y</td>
<td>4.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
</tbody>
</table>

*The scoring system is described in detail by Banowetz et al. (2008). A score of 4 indicates normal development, equivalent to the water control, at the end of the 7-day assay period. A score of 1 indicates germination arrest immediately after emergence of the radicle and coleoptile.

Table 2. Biological activity of culture filtrates from mutant strains of *P. fluorescens* WH6
in this gene, and insertions at each site were observed more than once. An additional Tn5 insertion (WH6-15::Tn5) that disrupted GAF biosynthesis was within an intergenic region immediately downstream of the putative carbamoyltransferase gene and immediately upstream of a small ORF (PFWH6_5255) of unknown function and significance that preceded the putative aminotransferase gene (PFWH6_5256). It seems likely that mutation in this intergenic region is affecting GAF biosynthesis by disrupting transcription of downstream genes such as the aminotransferase gene.

**Characterization of other mutations affecting GAF biosynthesis**

The *E. amylovora* screen identified five GAF mutations that are located outside the 13 kb gene cluster described above. (Table 1, Fig. 2). The WH6-3::Tn5 mutation identified in our earlier screen as occurring in a homologue of *prtR* (PFWH6_3687) (Armstrong et al., 2009; Kimbrel et al., 2010) also lies outside this gene cluster. The genomic context of all of these mutations is illustrated in Fig. 4.

Tn5 insertions affecting GAF production occurred in two ORFs external to the 13 kb gene cluster. WH6-6::Tn5 (located in PFWH6_5719) disrupted a gene encoding a HemY-like protein presumed to function in haem biosynthesis. Immediately downstream of this gene are two ORFs encoding homologues of DsbH, a disulfide-bond reductase, and AlgQ, a global transcriptional regulator. WH6-10::Tn5 (located in PFWH6_1695) disrupted a gene encoding a putative deoxyguanosine-triphosphatase triphosphohydrolase-like protein. Downstream of this gene are three additional ORFs, encoding a putative response regulator, a LuxR-family two-component transcriptional regulator and a histidine kinase.

Three GAF mutations external to the 13 kb gene cluster occurred in intergenic regions and presumably affect
expression of downstream genes. These genes encode both putative regulatory and enzymic products. The WH6-8::Tn5 mutation mapped 658 bp upstream of a gene (oprD) encoding a protein of unknown function. However, this insertion may have affected a gene on the antisense strand that occurs just 85 bp upstream of the site of insertion. This gene (PFWH6_4690) encodes a putative HisTriad (HIT) protein that likely functions as a nucleotide hydrolase and transferase acting on the alpha-phosphate of ribonucleotides. Immediately 3' of the HIT protein-encoding gene is an ORF encoding a product, SlyX, of unknown function. The WH6-9::Tn5 insertion occurred 230 bp upstream of the start site of another gene of unknown function (PFWH6_0984). Downstream of this gene, and in the same orientation, are two more genes of unknown function and a gene encoding a putative dehydrogenase. WH6-12::Tn5 mapped to an intergenic site 253 bp upstream of a gene with homology to thiG (PFWH6_5556), which has a putative role in thiazole biosynthesis. The single co-directional ORF located 3' of thiG is trmB, which encodes a putative tRNA methyltransferase.

**RT-qPCR analysis of putative regulatory gene functions affecting GAF biosynthesis**

Tn5 insertions in three putative regulatory genes (prtR, gntR and iopB) affected GAF production. RT-qPCR was performed to determine whether these three genes actually control transcription of two genes presumed to encode enzymes involved in GAF biosynthesis. The two presumptive biosynthetic genes examined, PFWH6_5256 and PFWH6_5257, encode putative aminotransferase and formyltransferase activities, respectively, and appear likely to be involved in the formation of the formylamino group that characterizes the GAF side-chain located distal to the oxygen bridge in the GAF molecule (Fig. 1). Both of these genes reside within the 13 kb gene cluster discussed above. When expression of these genes in the prtR, iopB and gntR mutants (WH6-3, WH6-4 and WH6-7, respectively) was compared with their transcription in wild-type WH6, a substantial decrease in transcription of both genes was observed in the mutants (Fig. 5), with the iopB mutation having the greatest effect in downregulating transcription.

**DISCUSSION**

The ability to utilize *E. amylovora* as a mutant screen has facilitated our efforts to identify the genes in *P. fluorescens* WH6 that are involved in GAF biosynthesis. Our earlier mutagenic efforts, using the *Poa* germination bioassay as a direct screen for loss of GAF activity, resulted in the recovery of two GAF mutants (WH6-2 and WH6-3) after screening 1214 Tn5-induced mutations (Armstrong et al., 2009). In the current study, we used *E. amylovora* to screen 6364 Tn5-induced mutations and identified 23 mutations affecting GAF biosynthesis. These mutations, combined with the two mutations generated in our earlier study, occurred in seven unique genes and four intergenic regions. Interestingly, none of the mutations observed in the most recent screen occurred in the two GAF-related genes identified in our earlier screen. Thus, neither screen was
GAF biosynthesis is consistent with the function of other regulatory factors involved with the transport of amino acids (or amino acid-related compounds) out of the bacterial cells. At least two of the genes with putative biosynthetic functions in this cluster, genes encoding a carbamoyltransferase and the aminotransferase, appear to contain genes that could encode the enzymes involved in the formation of the vinyl double bond. Tentative support for the pathway is provided by the results of our RT-qPCR studies of the effects of mutations in the regulatory genes prtR, iopB and gntR, which have been shown to regulate the temperature-sensitive production of an extracellular protease. PrtI appears to have structural features similar to the sigma factors involved in regulating extracytoplasmic functions (ECF sigma factors), and prtR has been suggested to encode a novel transmembrane activator of PrtI. The role of the WH6 iopB homologue in the regulation of GAF production is unclear, but the gene in P. chlororaphis PCL1391 has been reported to regulate phenazine production in that organism through effects on quorum sensing (Van Rij, 2006). The WH6 gntR gene identified in our mutant screen belongs to a very large family of transcriptional regulators with over 8500 members described to date (Hoskisson & Rigali, 2009). These regulatory factors control a wide variety of metabolic processes and environmental responses, including antibiotic production (Haydon & Guest, 1991; Hoskisson & Rigali, 2009). Therefore, the role of the WH6 gntR gene in regulating GAF biosynthesis is consistent with the function of other members of this family of transcriptional regulators.

The WH6 iopB and gntR homologues identified in the present study occur in a 13 kb gene cluster that includes at least three other genes that appear likely to be involved in GAF biosynthesis (Fig. 3). This region includes an additional putative regulatory gene (an iopA homologue), five genes that appear to encode enzymes involved in biosynthetic processes, two very small ORFs whose functions and significance are unclear, and three genes whose products appear to be involved with the transport of amino acids (or amino acid-related compounds) out of the bacterial cells. At least two of the genes with putative biosynthetic functions in this cluster, genes encoding a carbamoyltransferase and the aminotransferase described above, are essential for GAF production as demonstrated by the loss of GAF production that occurred upon disruption of these genes by Tn5 insertions. A putative formyltransferase gene (PFWH6_5257) suspected to be involved in GAF biosynthesis is located immediately downstream of the aminotransferase gene.

On balance, the genetic elements identified in this study are consistent with the basic outline of a hypothetical GAF biosynthetic pathway proposed in our earlier publication (McPhail et al., 2010). Based on precedents provided by pathways established for other naturally occurring oxyvinylglycines, we are reasonably confident that GAF biosynthesis begins with the amino acid homoserine (2-amino-4-hydroxybutanoic acid). Homoserine has been demonstrated as the starting point for the biosynthesis of the vinylglycine rhizobitoxine in Burkholderia andropogonis (Mitchell & Coddington, 1991) and Bradyrhizobium elkanii (Yasuta et al., 2001), and for AVG in Streptomyces sp. NRRL 5331 (Fernández et al., 2004). Addition of an amino group to the hydroxyl group of homoserine would produce canaline (2-amino-4-aminoxybutanoic acid), a naturally occurring compound known to be produced in certain legumes (Miersch, 1967; Rosenthal, 1997). The transfer of a formyl group to the 4-amino group of canaline would then result in the formation of a compound that can be termed dihydroGAF (2-amino-4-formylaminoxybutanoic acid). Finally, insertion of the vinyl double bond would then result in the formation of GAF itself (2-amino-4-formylaminoxy-3-butenoic acid) (Fig. 1). The GAF-related 13 kb gene cluster identified here appears to contain genes that could encode the enzymes necessary to make the dihydroGAF precursor of GAF, but this gene cluster does not contain a candidate for a gene encoding the desaturase required for insertion of the vinyl double bond. Tentative support for the pathway is provided by the results of our RT-qPCR studies of the effects of mutations in the regulatory genes prtR, iopB and gntR. Mutations in any one of these three genes, in addition to eliminating GAF production, also suppressed transcription of the aminotransferase and formyltransferase genes that we suspect encode the enzymes involved in the formation of the formylamino group found in the GAF side-chain.

To our knowledge, the 13 kb GAF-related gene cluster is absent from other sequenced strains of pseudomonads, with one exception. P. syringae pv. maculicola strain ES4326 (GenBank accession no. AEAK0000000) contains a gene...
cluster that is almost identical to the WH6 gene cluster. Every gene of the WH6 cluster, with the exception of the gene immediately downstream of the gntR homologue, is present in high homology in ES4326, although not unanimously as the first BLAST hit. The missing gene (corresponding to PFWH6_5427) encodes a putative trans-aconitase 2-methyltransferase, which is unlikely to be involved in GAF biosynthesis. In spite of the similarity of this particular gene cluster in WH6 and ES436, the latter strain does not produce GAF (data not shown). Therefore, we suspect that ES436 lacks one or more genes essential for the formation of the GAF vinyl double bond.

The functional diversity of the GAF-related genes identified in the present study indicates that the details of the GAF biosynthetic pathway may be somewhat more complicated than the simple model postulated by McPhail et al. (2010). Moreover, the biosynthesis of another oxyvinylglycine, 4-methoxyvinylglycine (2-amino-4-methoxy-3-butenonic acid) produced by P. aeruginosa, has been shown to involve a five-gene cluster that includes non-ribosomal peptide synthetases (Lee et al., 2010), suggesting that significant diversity may exist in the biosynthetic pathways by which vinylglycines are produced. In the case of GAF, the enzyme encoded by the carbamoyltransferase gene (corresponding to PFWH6_5254), located in the GAF-related 13 kb gene cluster, may function in the synthesis of amino and/or formyl donor compounds required by the pathway, but the roles of other genes shown to affect GAF biosynthesis, particularly those located outside of the GAF-related gene cluster, are uncertain. In some instances, the effects of Tn5 insertions on GAF production could be due to polar effects on genes downstream of the point of insertion. Based on the context in which these mutations occurred, some may have affected additional regulatory functions (e.g. WH6-10::Tn5) or processes related to the desaturase function yet to be identified (e.g. WH6-9::Tn5). To clarify these issues, site-specific mutagenic studies, combined with transcriptional analyses and appropriate gene transformations, are under way.

ACKNOWLEDGEMENTS

Support from the United States Department of Agriculture Cooperative State Research, Education and Extension Service (USDA CSREES) Grass Seed Cropping Systems for Sustainable Agriculture Special Grant Program and from the Oregon State University Agricultural Research Foundation is gratefully acknowledged by D.M and D.A. Technical assistance for parts of this study was provided by Donald D. Chen. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

REFERENCES


Edited by: G. Preston