The *sypA*, *sypB*, and *sypC* Synthetase Genes Encode Twenty-Two Modules Involved in the Nonribosomal Peptide Synthesis of Syringopeptin by *Pseudomonas syringae* pv. *syringae* B301D

Brenda K. Scholz-Schroeder,¹ Jonathan D. Soule,¹ and Dennis C. Gross²

¹Department of Plant Pathology, Washington State University, Pullman, Washington 99164-6430, U.S.A.; ²Department of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas 77843-2132, U.S.A.

Submitted 27 August 2002. Accepted 5 December 2002.

Syringopeptin is a necrosis-inducing phytotoxin, composed of 22 amino acids attached to a 3-hydroxy fatty acid tail. Syringopeptin, produced by *Pseudomonas syringae* pv. *syringae*, functions as a virulence determinant in the plant-pathogen interaction. A 73,800-bp DNA region was sequenced, and analysis identified three large open reading frames, *sypA*, *sypB*, and *sypC*, that are 16.1, 16.3, and 40.6 kb in size. Sequence analysis of the putative SypA, SypB, and SypC sequences determined that they are homologous to peptide synthetases, containing five, five, and twelve amino acid activation modules, respectively. Each module exhibited characteristic domains for condensation, aminoacyl adenylation, and thiolation. Within the aminoacyl adenylation domain is a region responsible for substrate specificity. Phylogenetic analysis of the substrate-binding pockets resulted in clustering of the 22 syringopeptin modules into nine groups. This clustering reflects the substrate amino acids predicted to be recognized by each of the respective modules based on placement of the syringopeptin NRPS (nonribosomal peptide synthetase) system in the linear (type A) group. Finally, SypC contains two C-terminal thioesterase domains predicted to catalyze the release of syringopeptin from the synthetase and peptide cyclization to form the lactone ring. The syringopeptin synthetases, which carry 22 NRPS modules, represent the largest linear NRPS system described for a prokaryote.

Additional keywords: lipopeptide phytotoxin, thiotemplate mechanism.

The syringopeptin (*syp*) and syringomycin (*syr*) gene clusters are localized to an approximately 145-kb *DraI* fragment on the chromosome of *Pseudomonas syringae* pv. *syringae* B301D (Scholz-Schroeder et al. 2001b). The *syr* and *syp* clusters were phenotypically and genotypically mapped relative to the position of the known *syr* and *syp* genes. It was demonstrated that the *syp* gene cluster lies adjacent to the *syr* gene

cluster and that almost the entire *syp* and *syr* gene clusters are located on this approximately 145-kb *DraI* fragment. The *syp* gene cluster was estimated to encompass approximately 80 kb in addition to syringomycin and syringopeptin biosynthesis genes (*syrE*, *syrB1*, *syrC*, and *sypA*), the *syr-syp* cluster contained genes for phytotoxin regulation (*syrP*, *salA*, *syrF*, and *syrG*) and secretion (*syrD*) (Lu et al. 2002; Scholz-Schroeder et al. 2001b). Together the *syr-syp* cluster was estimated to account for approximately 2% of the *P. syringae* pv. *syringae* genome (Scholz-Schroeder et al. 2001b). The *syp-syr* gene cluster appears to constitute a genomic island (Hacker and Kaper 2000) in *P. syringae* pv. *syringae*. The *syp-syr* gene cluster is a unique characteristic of *P. syringae* pv. *syringae* and a few closely related pathovars within *P. syringae* genomospecies 1 and contribute to its genetic distinctiveness (Gardan et al. 1999; Quigley and Gross 1994).

Syringopeptins are a class of necrosis-inducing lipopeptide phytotoxins produced by *P. syringae* pv. *syringae* (Ballio et al. 1991; Iacobellis et al. 1992). Syringopeptin SP22A is the major form produced by *P. syringae* pv. *syringae* B301D (Ballio et al. 1991) and is characterized as possessing a 3-hydroxydecanoyl fatty acid tail attached to a cyclic peptide 22 amino acids in length (Fig. 1). Syringopeptin is cytotoxic and capable of lysing erythrocytes at low concentrations. For example, less than 10 ng per ml is sufficient to lyse tobacco protoplasts (Hutchison and Gross 1997; Iacobellis et al. 1992; Lavermicocca et al. 1997). Based on its ability to form pores in artificial membranes, it is hypothesized that syringopeptin functions by targeting the plant plasma membrane in which at least three monomers of syringopeptin aggregate to form a pore that allows free passage of cations, to disrupt the electrical potential across the plant cell membrane, and to ultimately lead to plant cell death (Agner et al. 2000; Dalla Serra et al. 1999; Hutchison and Gross 1997). Accordingly, syringopeptin is an important virulence determinant of *P. syringae* pv. *syringae* (Scholz-Schroeder et al. 2001a). Strains carrying *sypA* mutations produced significantly smaller lesions in immature sweet cherry fruits and were reduced approximately 60% in virulence compared with the wild-type strain (Scholz-Schroeder et al. 2001a).

Syringopeptin is synthesized via a thiotemplate mechanism of biosynthesis (Grgurina and Benincasa 1994; Scholz-Schroeder et al. 2001a). This is a nonribosomal mechanism that utilizes multifunctional enzymes, called peptide synthetases, to catalyze the activation and addition of amino acids into the peptide chain (Marahiel et al. 1997). These peptide synthetases are composed of repeated amino acid activation modules, contain-

Corresponding author: D. C. Gross; Telephone: (979) 845-8288; E-mail: dgross@ppserver.tamu.edu

Nucleotide sequence data reported are available in the GenBank/EMBL database under the accession number AF286216

*The e-Xtra logo stands for “electronic extra” and indicates the figures and captions appear in color on-line with corresponding captions.

ing domains for condensation, aminoacyl adenylation, and thiolation. Modules are responsible for activation and incorporation of the amino acids into the growing peptide. A large number of prokaryotic and eukaryotic organisms synthesize peptide metabolites via this nonribosomal mechanism of biosynthesis (Marahiel et al. 1997; Mootz et al. 2002). In particular, many peptide antibiotics are synthesized by a nonribosomal peptide synthetase (NRPS) system, including gramicidin S, surfactin, and tyrocidine production by *Bacillus* spp. (Menkhaus et al. 1993; Mittenhuber et al. 1989; Mootz and Marahiel 1997; Turgay et al. 1992). In addition, there are several *Pseudomonas* spp. that utilize the thiotemplate mechanism for the biosynthesis of secondary metabolites, such as pyoverdine and syringomycin (Guenzi et al. 1998; Lehoux et al. 2000; Merriman et al. 1995; Zhang et al. 1995).

A colinear organization of the amino acid activating modules is observed for the linear (type A) NRPS system, and this is reflected in the sequence of the peptide product (Mootz et al. 2002). Genetically, syringomycin synthetases digress slightly from this, because the *syrB1* synthetase gene, responsible for incorporation of the last amino acid in the syringomycin peptide chain, is located upstream of the *syrE* synthetase gene (Guenzi et al. 1998). Stachelhaus and associates (1999) used computational analysis of the substrate-binding pockets of a large collection of NRPS modules, including all nine syringomycin modules, to construct a phylogenetic tree that showed clustering according to the specificity of individual modules. Mootz and associates (2002) recently classified the syringomycin NRPS system in the nonlinear (type C) group. Currently, only the first amino acid activation module of SypA (SypA-M1), involved in syringopeptin synthesis, has been sequenced and characterized (Scholz-Schroeder et al. 2001a). Based on the model for nonribosomal synthesis of peptides (Stachelhaus and Marahiel 1995; Marahiel et al. 1997), it is predicted that 22 modules will be required for the biosynthesis of SP22A and SP22B, the two forms of syringopeptin produced by *P. syringae* pv. *syringae* B301D (Ballio et al. 1995). It was demonstrated that the syringopeptin biosynthesis genes mapped to an 80-kb region of a 145-kb *DraI* fragment of the *P. syringae* pv. *syringae* B301D genome (Scholz-Schroeder 2001b). Because about 3 kb of DNA is required to code for each amino acid activation module, this region could easily encode the required amino acid activation modules necessary for syringopeptin synthesis.

In this report, we present the entire sequence of the *syp* synthetase genes obtained from cosmids identified in a previous study (Scholz-Schroeder et al. 2001b). Sequence analysis identified three large open reading frames (ORFs), *sypA*, *sypB*, and *sypC*, encompassing 73.8 kb of DNA, responsible for encoding three separate peptide synthetases. Phylogenetic analysis of the

primary sequence of the substrate binding pockets of all 22 syringopeptin and nine syringomycin modules was used to predict substrate specificity.

RESULTS

Nucleotide sequences of cosmids that map to the syringopeptin gene cluster.

Fragment assembly demonstrated that cosmids pYM1, p116, p149, p174, and p753, which encompass the syringopeptin gene cluster identified in a previous study (Table 1 and Fig. 2) (Scholz-Schroeder et al. 2001b), contained overlapping DNA sequence. Two gaps in this sequence were identified, and primers 31T3R2 and 13T7R1 and primers 71F1 and 58T3R1 amplified DNA fragments that were subcloned and designated pJS101 and pJS114, respectively. The sequence from these DNA fragments filled the gaps in the nucleotide sequence, generating an approximately 73,800-bp *syp* contig.

Sequence analysis of the *sypA*, *sypB*, and *sypC* genes.

Analysis of the 73,800-bp contig identified three prominent ORFs, designated *sypA*, *sypB*, and *sypC*. The first ORF, *sypA*, is 16,140 nucleotides in length and encompasses a previously described 3,920-bp region downstream of *syrD* (Scholz-Schroeder et al. 2001a). A Shine-Dalgarno sequence (GAGAAG) was previously identified (Scholz-Schroeder et al. 2001a), and the *sypA* gene is translated beginning at nucleotide 138 (ATG) and ending at nucleotide 16,277. The proposed translational start site of the second ORF, *sypB*, is preceded by a consensus ribosomal binding site (GAGAAG) and begins (GTG) 39 bases from the *sypA* stop codon at nucleotide 16,316. The *sypB* gene is 16,374 nucleotides in length ending at nucleotide 32,689. The third ORF, *sypC*, is 40,611 nucleotides in length. The proposed translational start site is preceded by a Shine-Dalgarno sequence (AGCAGG) and begins (ATG) 426 nucleotides downstream from the *sypB* stop codon at nucleotide 33,115 and ends at nucleotide 73,725. No potential rho-independent transcriptional terminators were identified for *sypA*, *sypB*, or *sypC*, demonstrating that the three ORFs constitute an operon, and the deduced protein products of the *sypA*, *sypB*, and *sypC* genes are 5,380, 5,458, and 13,537 amino acids in length, respectively. The *sypA*, *sypB*, and *sypC* genes are 62.0, 62.1, and 62.1% G+C, respectively. This is similar to other *P. syringae* genes in the GenBank/EMBL database, including the syringomycin biosynthesis genes *syrB1* (accession number U25130) and *syrE* (accession number AF047828) and the *P. aeruginosa* genome (Stover et al. 2000).

Previously, it was reported that several mini-Tn5 insertions disrupted the production of syringopeptin (Scholz-Schroeder et al. 2001b). DNA flanking these mini-Tn5 insertions was sequenced (Scholz-Schroeder et al. 2001b), and it was determined that the mini-Tn5 insertions carried in strains B301D.4 and B301D.5, respectively, were located within *sypC* at nucleotides 35,678 and 45,591.

Analysis of the syringopeptin synthetases.

Analysis of the deduced amino acid sequences of *sypA*, *sypB*, and *sypC* indicated that SypA, SypB, and SypC exhibit significant homology to a family of proteins involved in the thiotemplate mechanism of biosynthesis. In particular, SypA, SypB, and SypC exhibit 45.4, 45.6, and 45.9% identity, respectively, to SyrE and 42.7, 44.3, and 43.3% identity, respectively, to SyrB1, both of which are syringomycin synthetases (Guenzi et al. 1998; Zhang et al. 1995). The SypA, SypB, and SypC proteins also exhibited significant homology to many other peptide synthetases, including CYSYN responsible for cyclosporin A synthesis (Weber et al. 1994), GrsA involved in

Fig. 1. Structure of **A**, syringomycin form SRE and **B**, syringopeptin form SP22A, the major toxin forms produced by *Pseudomonas syringae* pv. *syringae* B301D. L-isomers of the respective amino acids are denoted as blue ovals, and D-isomers are denoted as white ovals. Abbreviations: Dab = 2,4-diaminobutyric acid; Dhb = 2,3-dehydroaminobutyric acid; and αThr = allothreonine. All other amino acids are identified by standard three-letter biochemical notation.

gramicidin S synthesis (Hori et al. 1989; Krätzschar et al. 1989), HC-toxin synthetase (Scott-Craig et al. 1992), and surfactin synthetases (Cosmina et al. 1993). These homologies were significant, as the calculated Z scores for each of these proteins with SypA, SypB, or SypC were greater than 114.7 (>6 indicates significance).

The deduced SypA, SypB, and SypC proteins are very similar in structural organization. A detailed analysis of each protein determined that they contain repeating amino acid activation modules (Marahiel et al. 1997) composed of conserved motifs for the condensation, aminoacyl adenylation, and thiolation domains (Stein and Vater 1996) (Fig. 3). The condensation domains within SypA, SypB, and SypC encompass approximately 450 amino acids containing seven conserved core sequences, including the conserved motif HHxxxDG (C3) in each of the 22 condensation domains (Fig. 3) (deCrécy-Lagard et al. 1995). The condensation domains are followed by aminoacyl adenylation domains that contain the ten conserved core sequences previously described by Marahiel and associates (1997) (Fig. 3). Some of these conserved sequences include TSGSTGLPKG (A3) (Fig. 3) and the highly conserved TGD motif (A7) (Fig. 3). The thiolation domains were observed to follow the aminoacyl adenylation domains in SypA, SypB, and SypC (Fig. 3), and they contain the conserved thioester-binding motif LGGHS (T) (Fernández-Moreno et al. 1992; Gocht and Marahiel 1994; Marahiel et al. 1997; Mootz et al. 2001; Stein and Vater 1996).

As predicted, a carboxy-terminal thioesterase (TE) domain of approximately 260 amino acid residues was observed at the C-terminal end of the SypC protein. Alignment of the C-terminal portion of SypC with the corresponding region of PvdD (accession number U07359), CmaT (accession number U14657), GrsB (accession number X61658), EntF (accession number M60177), TycC (accession number AF004835), and SyrE (accession number AF047828) identified the conserved motif

GxSxG (amino acids 13,053 to 13,057), the conserved aspartic acid residue at amino acid 13,082, and the conserved histidine residue at amino acid 13,229 in SypC (data not shown). These three conserved motifs compose the 'nucleophile elbow' of the TE domain (Bruner et al. 2002; Mootz and Marahiel 1997; Schneider and Marahiel 1998). Quite unexpectedly, a second TE domain of approximately 250 amino acids was found to start 11 amino acids after the first TE domain in the SypC protein. This second TE domain contains the GxSxG motif (amino acids 13,347 to 13,351) and the aspartic acid residue (amino acid 13,376), but it is lacking the conserved histidine residue (Bruner et al. 2002).

Despite the fact that the syringopeptin structure contains several modified amino acids including the D-isomers of proline, valine, alanine, and serine, no racemase or epimerase domains were observed (deCrécy-Lagard et al. 1995). SypA and SypB are comparable in size, each containing five amino acid activation modules. SypC is much larger than SypA or SypB, containing 12 amino acid activation modules followed by two TE domains. These data indicate that the syringopeptin synthetases exhibit the modular organization characteristic of peptide synthetases (Marahiel et al. 1997; Stachelhaus and Marahiel 1995).

Analysis of the substrate specificity regions of SypA, SypB, and SypC.

The putative substrate-binding pocket (Challis et al. 2000; Conti et al. 1997; Stachelhaus et al. 1999) was identified for each of the 22 amino acid activation modules of SypA, SypB, and SypC. The phylogenetic analysis of the 22 putative substrate-binding pockets from the amino acid-activating modules of SypA, SypB, and SypC, with the substrate-binding pockets from the amino acid-activating module of SyrB1 and the eight amino acid-activating modules of SyrE (Fig. 4) determined that these regions clustered with those that incorporate identical or similar

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics ^a	Source
<i>E. coli</i>		
DH10B	F ⁻ <i>mcrA</i> Δ lacX74 (ϕ 80 Δ lacZ Δ M15) Δ (<i>mrr-hsdRMS-mcrB</i>) <i>deoR recA1 endA1 araD139</i> Δ (<i>ara, leu</i>)7697 <i>galU galK</i> λ ⁻ <i>rpsL nup</i>	Invitrogen Life Technologies, Carlsbad, CA, U.S.A.
XL1-Blue MR	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1</i> [F ⁺ <i>proAB lacI</i> ^q Z Δ M15 Tn10 (Tet ^r)]	Stratagene, La Jolla, CA, U.S.A.
Cosmids		
p116	pLAFR3 carrying a 20-kb <i>P. syringae</i> pv. <i>syringae</i> B301D genomic DNA fragment; Tc ^r	Scholz-Schroeder et al. 2001b
p149	pLAFR3 carrying a 30-kb <i>P. syringae</i> pv. <i>syringae</i> B301D genomic DNA fragment; Tc ^r	Scholz-Schroeder et al. 2001b
p174	pLAFR3 carrying a 23-kb <i>P. syringae</i> pv. <i>syringae</i> B301D genomic DNA fragment; Tc ^r	Scholz-Schroeder et al. 2001b
p753	pLAFR3 carrying an 8-kb <i>P. syringae</i> pv. <i>syringae</i> B301D genomic DNA fragment; Tc ^r	Scholz-Schroeder et al. 2001b
Plasmids		
pBluescript II SK	High-copy-number cloning vector; Ap ^r	Stratagene
pLAFR3	IncP cos ⁺ rlx ⁺ ; Tc ^r	Staskawicz et al. 1987
pUC18	High-copy-number cloning vector; Ap ^r	Yanisch-Perron et al. 1985
pBS008	pUC18 carrying a 5.1-kb <i>EcoRI-KpnI</i> fragment of pYM1; Ap ^r	Scholz-Schroeder et al. 2001a
pJS048	pBSK carrying a 7.5-kb <i>KpnI</i> fragment of p149; Ap ^r	This study
pJS037	pBSK carrying a 1.5-kb <i>BamHI/ HindIII</i> fragment of p149; Ap ^r	This study
pJS018	pBSK carrying a 3.1-kb <i>BamHI</i> fragment of p149; Ap ^r	This study
pJS028	pBSK carrying a 4.6-kb <i>KpnI/ BamHI</i> fragment of p149; Ap ^r	This study
pJS026	pBSK carrying a 5.0-kb <i>HindIII</i> fragment of p149; Ap ^r	This study
pJS031	pBSK carrying a 1.5-kb <i>BamHI/ HindIII</i> fragment of p149; Ap ^r	This study
pJS114	pBSK carrying a 1.2-kb PCR-amplified clone from <i>P. syringae</i> pv. <i>syringae</i> B301D; Ap ^r	This study
pJS013	pBSK carrying a 7.8-kb <i>EcoRI</i> fragment of p753; Ap ^r	Scholz-Schroeder et al. 2001b
pJS069	pBSK carrying a 9.7-kb <i>HindIII</i> fragment of p174; Ap ^r	This study
pJS049	pBSK carrying an 8.1-kb <i>BamHI</i> fragment of p174; Ap ^r	Scholz-Schroeder et al. 2001b
pJS071	pBSK carrying an 11.8-kb <i>HindIII</i> fragment of p174; Ap ^r	This study
pJS101	pBSK carrying a 3.4-kb PCR-amplified clone from <i>P. syringae</i> pv. <i>syringae</i> B301D; Ap ^r	This study
pJS058	pBSK carrying a 2.5-kb <i>HindIII</i> fragment of p116; Ap ^r	This study
pJS057	pBSK carrying a 3.9-kb <i>HindIII</i> fragment of p116; Ap ^r	This study
pJS065	pBSK carrying a 12.9-kb <i>HindIII</i> fragment of p116; Ap ^r	This study
pJS067	pBSK carrying a 7.1-kb <i>BamHI</i> fragment of p116; Ap ^r	This study
pYM1	pGS72 carrying a 16-kb <i>HindIII</i> fragment from strain B301D; Tc ^r	Mo and Gross 1991

^a Ap^r and Tc^r resistance to ampicillin and tetracycline, respectively.

TycC (accession number AF004835), indicating that they are both TE domains. It is unclear why SypC would contain two TE domains and exactly how the two domains function to release the syringopeptin molecule from SypC. The first TE domain is complete, while the second domain does not contain the conserved histidine residue in the conserved region of the domain, which indicates that the second domain may not be functional, as this is an important component of the catalytic triad for TE domains (Bruner et al. 2002). While the conserved histidine was not located in the conserved region of the second TE domain, there are other histidine residues in close proximity that could complete the ‘nucleophile elbow’ (Bruner et al. 2002), possibly enabling the second TE domain to be functional.

Despite the presence of multiple D-isomers of various amino acids in syringopeptin (Fig. 1), no epimerization domains were

observed in SypA, SypB, or SypC. This was previously observed for the SyrE syringomycin synthetase (Guenzi et al. 1998). The lack of an epimerase domain may suggest that a separate racemase modifies the amino acids to their D-isomeric form, which is then directly bound by the aminoacyl adenylation domain and incorporated into the syringopeptin structure, as suggested for the cyclosporin or HC-toxin synthesis (Scott-Craig et al. 1992; Weber et al. 1994). The lack of epimerization domains in SypA, SypB or SypC may also indicate that the amino acid is modified after it is bound to the amino acid-activating module, as suggested by Guenzi and associates (1998) for syringomycin synthesis.

The synthesis of antibiotic peptides via the nonribosomal thio-template mechanism of biosynthesis is the focus of intense research. Understanding the genetic information that encodes these proteins enables the manipulation of their modular domain

Fig. 3. Diagrammatic representation of an amino acid activation module illustrating the condensation, aminoacyl adenylation, and thiolation domains characteristic of peptide synthetases as defined by Marahiel and associates (1997). The condensation domain contains the characteristic HHxxxDG motif (C3). The aminoacyl adenylation domain contains five conserved core sequences. A3 corresponds to the Walker type A motif involved in ATP binding (Saraste et al. 1990), the motif of A7 is associated with ATPase activity, and A6 and A8 are predicted to be involved in catalyzing aminoacyl adenylate formation. T, identified by a characteristic LGGHSL motif, is located in the thiolation domain. The conserved sequences for each region are as reported by Stachelhaus and Marahiel (1995), as well as the actual amino acid sequence for the domains involved in syringopeptin biosynthesis.

structure to form hybrid peptide synthetases that will produce new peptides (Doekel and Marahiel 2000; Mootz et al. 2000). The selection of the amino acids to be incorporated in the new peptides by a peptide synthetase is dependent on the region of substrate specificity (Conti et al. 1997; Marahiel et al. 1997). X-ray crystallography was used to identify the region of substrate binding in the aminoacyl adenylation domain of GrsA, responsible for the activation and incorporation of phenylalanine in gramicidin S synthesis (Conti et al. 1997). The region of substrate binding from GrsA was used as a model to identify the substrate-binding regions for all 22 modules from syringopeptin synthetases and nine modules from syringomycin synthetases. Phylogenetic analysis of these substrate-binding regions identified clusters that reflect the predicted substrate bound by each amino acid activation module (Fig. 4). The phylogenetic analysis demonstrates that SyrE-M1, SyrE-M2, and SypC-M16 are similar to one another and form the Ser cluster (Fig. 4). This is significant because the SyrE-M1 and SyrE-M2 amino acid activation modules were previously shown to incorporate serine via ATP/PP_i exchange reactions (Guenzi et al. 1998). The clustering of SypC-M16 indicates that this amino acid activation module is responsible for the activation and incorporation of serine, the sixteenth amino acid, into the syringopeptin structure. The Thr/Dhb cluster (SypA-M1, SyrE-M7, SyrB1-M9, SypB-M9, SypC-M14, SypC-M15, and SypC-M18) is of particular interest. SypA-M1 was analyzed previously in a similar manner, and it was observed that it clustered with SyrE-M7 and SyrB1-M9 (Scholz-Schroeder et al. 2001a). In addition, it was demonstrated that SyrB1 activates and incorporates L-Thr in ATP/PP_i exchange assays (Guenzi et al. 1998), and it is known that the formation of Dhb occurs via the dehydration of L-[¹⁴C]-Thr in *P. syringae* (Grgurina and Marriotti 1999). Thus, it appears that there is a common substrate specificity for the Thr and Dhb amino acid activation modules. Correspondingly, the modular organization of syringopeptin synthetases, SypB-M9, SypC-M14, SypC-M15, and SypC-M18 are predicted to incorporate Dhb into the syringopeptin structure, and this is reflected in their clustering with SypA-M1, SyrE-M7, and SyrB1-M9 (Scholz-Schroeder et al. 2001a).

It also was demonstrated that the SyrE-M3, SyrE-M4, SypC-M20, and SypC-M21 modules, predicted to incorporate 2,4-diaminobutyric acid into their respective phytotoxin structures, are closely related (Fig. 4). Similar observations can be made for the substrate-binding regions of the Val cluster composed of SypA-M3, SypA-M4, SypB-M7, SypB-M8, and SypC-M11. The clustering observed within the Ala cluster, including SypA-M5, SypB-M6, SypB-M10, SypC-M12, SypC-M13, SypC-M17, and SypC-M19, is not as strong as that observed for other clusters within this dendrogram (64% on the basis of 1,000 samplings). However, the observed clustering of these substrate-binding regions indicates that they are more closely related to each other than to others. Finally, SypC-M22 and SyrE-M6 form the final cluster, which was designated as the Tyr/Phe cluster (Fig. 4). Based on the modular organization of nonribosomal peptide synthesis, SypC-M22 is predicted to incorporate Tyr, and SyrE-M6 is predicted to incorporate Phe into the respective phytotoxin structures. The clustering observed reflects significant homology between the two substrate-binding regions (86% identity, 90% similarity) and is most likely the result of the structural similarity between the two aromatic amino acids, Tyr and Phe. The phylogenetic analysis encompassing 31 substrate-binding regions from the amino acid-activating modules from an individual organism (*P. syringae* pv. *syringae* B301D) effectively removes evolutionary bias from the analysis, providing clear and concise results.

Phylogenetic analysis of NRPS modules can be confounded by the evolutionary origin of the sequences analyzed, especially if modules from eukaryotes and prokaryotes are included

in comparisons, as noted in the studies by Stachelhaus and associates (1999). With the exception of *Bacillus* spp., very few studies of nonribosomal peptide synthetase systems have multiple examples of amino acid activation module sequence available from the same genus, let alone the same species (Konz et al. 1999). In order to circumvent this problem, extensive analyses of the substrate-binding region from amino acid activation modules from numerous peptide synthetases of various origin were analyzed, and signature sequences were identified that could be used to predict which amino acid would be activated and subsequently incorporated into the growing peptide (Conti et al. 1997; Stachelhaus et al. 1999). Stachelhaus and associates (1999) observed the same clustering of SyrB1 and SyrE-M7 and of SyrE-M3 and SyrE-M4 as observed in our analysis. However, an effect of evolutionary origin was observed in their analysis (Stachelhaus et al. 1999). In an effort to reduce the effect of organismal bias of these analyses, the signature sequences were further refined, and a database was developed using eight residues composing the signature sequences of those amino acid activation modules whose substrate specificity was determined by enzymatic assay or by logical inference. This database could then be used to predict the substrate specificity of newly identified amino acid activation modules (Challis et al. 2000). Using this program, we analyzed the substrate-binding regions of the syringopeptin synthetase amino acid activation modules. Good agreement was observed between the predictive program and the phylogenetic clustering of the amino acid activation modules predicted to activate and incorporate Val, Ser, Dab, and Thr/Dhb. Little agreement was observed between the predictive program and the phylogenetic clustering of the amino acid activation modules predicted to activate and incorporate Ala. These modules were consistently predicted to activate and incorporate Gly by the program (data not shown). Challis and associates (2000) indicated that the substrate-binding regions of Ala and Gly exhibit low homology among themselves, and this lower homology was observed for the Ala cluster in our phylogenetic analysis (Fig. 4; 64% on the basis of 1,000 samplings). Recently, May and associates (2002) further refined the specificity-conferring code to include 11 conserved residues, enabling the prediction of both amino and aryl acid substrate-binding regions. This refinement may aid in the identification of unknown substrate specificity regions in which a large number of modules from the same organism are not available. Thus, this emphasizes the strength of our phylogenetic analysis using 31 modules from the same genus and species.

Syringopeptin and syringomycin exhibit notable similarities in relation to their structure and function, but they differ in their relative contribution to the virulence of *P. syringae* pv. *syringae* B301D (Ballio et al. 1991; Iacobellis et al. 1992; Scholz-Schroeder et al. 2001a). A syringomycin synthetase mutant (*syrB1*) is reduced 23%, while a syringopeptin synthetase mutant (*sypA*) is reduced 59% in virulence, respectively, relative to the wild-type *P. syringae* pv. *syringae* B301D on immature sweet cherry fruits. Interestingly, the contribution of the toxins to virulence is not synergistic, as a syringopeptin-syringomycin (*sypA-syrB1*) double mutant is reduced 76% in virulence relative to the wild-type strain B301D (Scholz-Schroeder et al. 2001a). Two more syringopeptin mutants, B301D.4 and B301D.5, carrying mini-Tn5 insertions were previously reported but were not completely characterized (Scholz-Schroeder et al. 2001b). In this study, it was determined that the mini-Tn5 insertions carried within *P. syringae* pv. *syringae* B301D.4 and B301D.5 are located within *sypC* at nucleotides 35,678 and 45,591, respectively, relative to the GenBank sequence. This demonstrates that mutations within the first syringopeptin synthetase gene *sypA* and the last

Fig. 4. Phylogenetic analysis of the 22 substrate specificity regions from the syringopeptin synthetases, SypA, SypB, and SypC and the regions of substrate specificity from the syringomycin synthetases SyrE and SyrB1. The clustering observed reflects the putative substrates incorporated by the various modules. The numbers at the nodes indicate the levels of bootstrap support based on neighbor-joining analysis as a percentage of 10,000 resampled data sets. Abbreviations: SypA-M1 to SypA-M5 = the respective amino acid activating modules of the SypA protein; SypB-M6 to SypB-M10 = the respective amino acid activating modules of the SypB protein; SypC-M11 to SypC-M22 = the respective amino acid activating modules of the SypC protein; SyrE-M1 to SyrE-M8 = the respective amino acid activating modules of the SyrE protein; and SyrB1-M9 = the amino acid activation module of SyrB1. The physical distance indicated represents the evolutionary divergence in the number of expected amino acid substitutions per 100. Amino acid abbreviations: Dab = 2,4-diaminobutyric acid; Dhb = 2,3-dehydroaminobutyric acid; and α Thr = *allo*threonine. All other amino acids are identified by standard three-letter biochemical notation.

syringopeptin synthetase gene *sypC* have an equivalent detrimental effect on syringopeptin production.

The first step towards the characterization of the syringopeptin gene cluster indicates that the *sypA*, *sypB*, and *sypC* genes are located within the cluster, and they encode peptide synthetases responsible for the nonribosomal synthesis of syringopeptin. As a genomic cluster, it is predicted that the *syp* region carries genes involved in toxin secretion and regulation, as well as ancillary functions important to the NRPS system. For example, proteins responsible for epimerization of amino acids incorporated into syringopeptin and acyltransferase activity, responsible for attachment of the fatty acid tail to the peptide, remain to be identified. Analysis of the DNA region downstream of *sypC* may identify genes important to phyto-toxin production, which is significant to plant pathogenicity of *P. syringae* pv. *syringae* B301D.

MATERIALS AND METHODS

Bacterial strains and media.

The bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* DH10B (Gibco-BRL Technologies, Inc., Grand Island, NY, U.S.A.) and XL1-Blue MR (Stratagene U.S.A., La Jolla, CA, U.S.A.) were cultured at 37°C in Luria Bertani (LB) broth or agar. Antibiotics (Sigma Chemical Co., St. Louis) were added to the media at the following final concentrations: 100 µg of ampicillin per ml, 100 µg of kanamycin per ml, and 25 µg tetracycline per ml.

DNA manipulations, in vitro mutagenesis, and sequencing.

Routine procedures (Sambrook et al. 1989) were used to isolate and subclone cosmid DNA (Table 1) (Scholz-Schroeder et al. 2001b) into pBlueScript II SK (Stratagene). Plasmid DNA was isolated using a CTAB boiling plasmid prep (Del Sal et al. 1989). Subclones were mutagenized using the EZ::TN <Kan-2> insertion kit according to manufacturer's instructions (Epicentre Technologies Corp., Madison, WI, U.S.A.). Transformants of each subclone, carrying the EZ::TN transposon, were selected for growth on LB plus kanamycin. On average, at least five transformants per kb of target DNA were selected for sequencing reactions. Forward and reverse primers supplied with the EZ::TN <Kan-2> Insertion Kit were used for initial sequence reactions off each plasmid carrying an EZ::TN transposon. Sequence reactions were completed using fluorescence-based dideoxy terminators and Ampli-Taq polymerase and were run on an Applied Biosystems sequencer model 377 (Perkin Elmer Applied Biosystems, Inc., Norwalk, CT, U.S.A.). Primer walking was used to complete the sequence when noncontiguous sequences were present. Primers were designed using the PRIME program within the Wisconsin Sequence Analysis programs of Genetics Computer Group Package (release beta 10.0; GCG, Madison, WI, U.S.A.) (Devereux et al. 1984) and were synthesized by Operon Technologies Inc. (Alameda, CA, U.S.A.). Using this procedure, at least three sequence reactions were completed for the contig, with at least one reaction for each strand of a DNA region to confirm the nucleotide sequence. Polymerase chain reaction (PCR) was used to amplify DNA corresponding to two regions of the syringopeptin gene cluster not present on cosmid clones. DNA regions were amplified from 300 ng *P. syringae* pv. *syringae* B301D chromosomal DNA using 1 U VentPolymerase (New England Biolabs, Beverly, MA, U.S.A.), and 0.3 mM dNTPs, 1.5 mM MgCl₂, and 0.5 µM of each primer. The following parameters were used to amplify both DNA fragments: 94°C for 1 min for one cycle, followed by 24 cycles of 94°C for 40 s, 50°C for 45 s, and 72°C for 3.5 min; the 24 cycles were followed by 72°C for 2 min.

Primers 31T3R2 (5' TCTGCTCAATGGTTATGG 3') and 13T7R1 (5' GTGTTGAACGGGTTATCG 3') were used to amplify a DNA region between nucleotides 25,261 and 28,773, located in *sypB*. Primers 71F1 (5' AGCACTGCGAACACATCAC 3') and 58T3R1 (5' GTTCTGAGCCTTCACCACCTTG 3') were used together to amplify a DNA region between nucleotides 55,991 and 56,932, located in *sypC*. PCR and sequence reactions were performed using a M. J. Research PTC-200 DNA Engine (M J Research, Inc., San Francisco). PCR products were gel purified using the Qiaex II gel extraction kit (Qiagen Inc., Valencia, CA, U.S.A.) and were blunt-end ligated into the *EcoRV* site of pBlueScript II SK. Three independent subclones were sequenced for each PCR product to determine the validity of the obtained sequence.

Computer analysis of the *sypA*, *sypB*, and *sypC* nucleotide sequences.

Sequence data were analyzed using GCG and Omega version 2.0 software package (Oxford Molecular Ltd., Oxford, U.K.). The following programs were used: FRAGMENT ASSEMBLY, for assembly of the nucleotide sequence; FINDPATTERNS, for identification of Shine-Dalgarno sequences for the translational start site; TERMINATOR, for identification of rho-independent transcription terminators; and BESTFIT, for comparison of two sequences. Database searches for genes and proteins homologous to obtained nucleotide sequence and putative proteins were performed using the BLAST server at the National Center for Biotechnology Information (Bethesda, MD, U.S.A.). The significance of protein sequence similarity was evaluated by sequence randomization and calculation of Z scores using GAP. The GAP program was run using the randomization option of 100 to produce the necessary data for Z score calculations. The Z score was calculated using the actual quality score minus the mean quality value of the randomized scores divided by the standard deviation of randomized score distribution. Protein sequence homology was considered significant when the Z score value was greater than six.

The amino acid-activating modules from SypA, SypB, SypC, SyrE, and SyrB1 were compared to GrsA, using BESTFIT in order to identify the regions within each protein encompassing substrate-binding pockets (Conti et al. 1997, Stachelhaus et al. 1999, Scholz-Schroeder et al. 2001a). The substrate-binding region of each module was located within the amino acid adenylation domain and encompassed the region between motifs A3 and A6 (Marahiel et al. 1997). An alignment of the substrate-binding pockets of the 22 amino acid-activating modules from the SypA, SypB, and SypC proteins, along with the substrate-binding pockets from the amino acid-activating module of SyrB1 (accession number U25130) and the eight amino acid-activating modules of the SyrE protein (accession number AF047828) obtained from GenBank, was completed using Clustal-X version 1.8. TreeView version 1.5 (Page 1996) was used to construct a dendrogram. TE domains were predicted with PROSITE, and alignment of TE domains was completed using Clustal-X Version 1.8.

Nucleotide sequence and accession numbers.

Nucleotide and amino acid sequence data for *sypA*, *sypB*, and *sypC* are available from the GenBank/EMBL database under accession number AF286216.

ACKNOWLEDGMENTS

We are grateful to S.-E. Lu for helpful discussions and critical evaluation of this manuscript. This work was supported by grant 2001-35319-10400 from the National Research Initiative Competitive Grants Program of the U.S. Department of Agriculture, Science, and Education Administration.

LITERATURE CITED

- Agner, G., Kaulin, Y. A., Gurnev, P. A., Szabo, Z., Schagina, L. V., Takemoto, J. Y., and Blasko, K. 2000. Membrane-permeabilizing activities of cyclic lipopeptides, syringopeptin 22A and syringomycin E from *Pseudomonas syringae* pv. *syringae* in human red blood cells and in bilayer lipid membranes. *Bioelectrochemistry* 52:161-167.
- Ballio, A., Barra, D., Bossa, F., Collina, A., Grgurina, I., Marino, G., Moneti, G., Paci, M., Pucci, P., Segre, A., and Simmaco, M. 1991. Syringopeptins, new phytotoxic lipopeptides of *Pseudomonas syringae* pv. *syringae*. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 291:109-112.
- Ballio, A., Bossa, F., DiGiorgio, D., Dinola, A., Manetti, C., Paci, M., Scaloni, A., and Segre, A. L. 1995. Solution conformation of the *Pseudomonas syringae* pv. *syringae* phytotoxic lipopeptide syringopeptin 25-A—Two-dimensional NMR, distance geometry and molecular dynamics. *Eur. J. Biochem.* 234:747-758.
- Bruner, S. D., Weber, T., Kohli, R. M., Schwarzer, D., Marahiel, M. A., Walsh, C. T., and Stubbs, M. T. 2002. Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. *Structure* 10:301-310.
- Challis, G. L., Ravel, J., and Townsend, C. A. 2000. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem. Biol.* 7:211-224.
- Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. 1997. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* 16:4174-4183.
- Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., and van Sinderen, D. 1993. Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. *Mol. Microbiol.* 8:821-831.
- Dalla Serra, M., Fagiuoli, G., Nordera, P., Bernhart, I., Della Volpe C., DiGiorgio, D., Ballio, A., and Menestrina, G. 1999. The interaction of lipopeptide toxins from *Pseudomonas syringae* pv. *syringae* with biological and model membranes: A comparison of syringotoxin, syringomycin, and two syringopeptins. *Mol. Plant-Microbe Interact.* 12:391-400.
- deCrécy-Lagard, V., Marlière, P., and Saurin, W. 1995. Multienzymatic nonribosomal peptide biosynthesis: Identification of the functional domains catalysing peptide elongation and epimerisation. *C. R. Acad. Sci. Paris* 318:927-936.
- de Ferra, F., Rodriguez, F., Tortora, O., Tosi, C., and Grandi, G. 1997. Engineering of peptide synthetases. Key role of the thioesterase-like domain for efficient production of recombinant peptides. *J. Biol. Chem.* 272:25304-25309.
- Del Sal, G., Manfioletti, G., and Schneider, C. 1989. The CTAB-DNA precipitation method: A common mini-scale preparation of template DNA from phagmids, phages or plasmids suitable for sequencing. *Biotechniques* 7:514-520.
- Devereux, J., Haeberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-395.
- Doekel, S. and Marahiel, M. A. 2000. Dipeptide formation on engineered hybrid peptide synthetases. *Chem. Biol.* 7:373-384.
- Fernández-Moreno, M. A., Martínez, E., Boto, L., Hopwood, D. A., and Malpartida, F. 1992. Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J. Biol. Chem.* 267:19278-19290.
- Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimont, F., and Grimont, P.-A. D. 1999. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremiae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *Int. J. Syst. Bacteriol.* 49:469-478.
- Gocht, M., and Marahiel, M. A. 1994. Analysis of core sequences in the D-phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis. *J. Bacteriol.* 176:2654-2662.
- Grgurina, I., and Benincasa, M. 1994. Evidence of the nonribosomal biosynthetic mechanism in the formation of syringomycin and syringopeptin, bioactive lipopeptides of the phytopathogenic bacterium *Pseudomonas syringae* pv. *syringae*. *Ital. Biochem. Soc. Trans.* 5:143.
- Grgurina, I. and Mariotti, F. 1999. Biosynthetic origin of syringomycin and syringopeptin 22, toxic secondary metabolites of the phytopathogenic bacterium *Pseudomonas syringae* pv. *syringae*. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 462:151-154.
- Guenzi, E., Galli, G., Grgurina, I., Gross, D. C., and Grandi, G. 1998. Characterization of the syringomycin synthetase gene cluster—A link between prokaryotic and eukaryotic peptide synthetases. *J. Biol. Chem.* 273:32857-32863.
- Hacker, J., and Kaper, J. B. 2000. Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* 54:641-679.
- Hori, K., Yamamoto, Y., Minetoki, T., Kurotsu, T., Kanda, M., Miura, S., Okamura, K., Furuyama, J., and Saito, Y. 1989. Molecular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene. *J. Biochem. (Tokyo)* 106:639-645.
- Hutchison, M. L., and Gross, D. C. 1997. Lipopeptide phytotoxins produced by *Pseudomonas syringae* pv. *syringae*: Comparison of the biosurfactant and ion channel-forming activities of syringopeptin and syringomycin. *Mol. Plant-Microbe Interact.* 10:347-354.
- Iacobellis, N. S., Lavermicocca, P., Grgurina, I., Simmaco, M., and Ballio, A. 1992. Phytotoxic properties of *Pseudomonas syringae* pv. *syringae* toxins. *Physiol. Mol. Plant Pathol.* 40:107-116.
- Konz, D., Doekel, S., and Marahiel, M. A. 1999. Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. *J. Bacteriol.* 181:133-140.
- Krätzschmar, J., Krause, M., and Marahiel, M. A. 1989. Gramicidin S biosynthesis operon containing the structural genes *grsA* and *grsB* has an open reading frame encoding a protein homologous to fatty acid thioesterases. *J. Bacteriol.* 171:5422-5429.
- Lavermicocca, P., Iacobellis, N. S., Simmaco, M., and Graniti, A. 1997. Biological properties and spectrum of activity of *Pseudomonas syringae* pv. *syringae* toxins. *Physiol. Mol. Plant Pathol.* 50:129-140.
- Lehoux, D. E., Sanschagrin, F., and Levesque, R. C. 2000. Genomics of the 35-kb *pvd* locus and analysis of novel *pvdIJK* genes implicated in pyoverdine biosynthesis in *Pseudomonas aeruginosa*. *FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Lett.* 190:141-146.
- Lu, S.-E., Scholz-Schroeder, B. K., and Gross, D. C. 2002. Characterization of the *salA*, *syfF*, and *sygR* regulatory genes located at the right border of the syringomycin gene cluster of *Pseudomonas syringae* pv. *syringae*. *Mol. Plant-Microbe Interact.* 15:43-53.
- Marahiel, M. A. 1997. Protein templates for the biosynthesis of peptide antibiotics. *Chem. Biol.* 4:561-567.
- Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. 1997. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* 97:2651-2673.
- May, J. J., Kessler, N., Marahiel, M. A., and Stubbs, M. T. 2002. Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U.S.A.* 99:12120-12125.
- Menkhaus, M., Ullrich, C., Kluge, B., Vater, J., Vollenbroich, D., and Kamp, R. M. 1993. Structural and functional organization of the surfactin synthetase multienzyme system. *J. Biol. Chem.* 268:7678-7684.
- Merriman, T. R., Merriman, M. E., and Lamont, I. L. 1995. Nucleotide sequence of *pvdD*, a pyoverdine biosynthetic gene from *Pseudomonas aeruginosa*: PvdD has similarity to peptide synthetases. *J. Bacteriol.* 177:252-258.
- Mittenhuber, G., Weckermann, R., and Marahiel, M. A. 1989. Gene cluster containing the genes for tyrocidine synthetases 1 and 2 from *Bacillus brevis*: Evidence for an operon. *J. Bacteriol.* 171:4881-4887.
- Mo, Y.-Y., and Gross, D. C. 1991. Expression *in vitro* and during plant pathogenesis of the *syfB* gene required for syringomycin production by *Pseudomonas syringae* pv. *syringae*. *Mol. Plant-Microbe Interact.* 4:28-36.
- Mootz, H. D., and Marahiel, M. A. 1997. The tyrocidine biosynthesis operon of *Bacillus brevis*: Complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* 179:6843-6850.
- Mootz, H. D., Schwarzer, D., and Marahiel, M. A. 2000. Construction of hybrid peptide synthetases by module and domain fusions. *Proc. Natl. Acad. Sci. U.S.A.* 97:5848-5853.
- Mootz, H. D., Finking, R., and Marahiel, M. A. 2001. 4'-phosphopantetheine transfer in primary and secondary metabolism of *Bacillus subtilis*. *J. Biol. Chem.* 276:37289-37298.
- Mootz, H. D., Schwarzer, D., and Marahiel, M. A. 2002. Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *ChemBio. Chem.* 3:491-504.
- Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12:357-358.
- Quigley, N. B., and Gross, D. C. 1994. Syringomycin production among strains of *Pseudomonas syringae* pv. *syringae*: Conservation of the *syfB* and *syfD* genes and activation of phytotoxin production by plant signal molecules. *Mol. Plant-Microbe Interact.* 7:78-90.
- Sambrook J., Fritsch, E. F., and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY, U. S. A.
- Saraste, M., Sibbald, P. R., and Wittinghofer, A. 1990. The P-loop: A common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* 15:430-434.
- Schlumbohm, W., Stein, T., Ullrich, C., Vater, J., Krause, M., Marahiel, M. A., Kruff, V., and Wittmann-Liebold, B. 1991. An active serine is

- involved in covalent substrate amino acid binding at each reaction center of gramicidin S synthetase. *J. Biol. Chem.* 266:23135-23141.
- Schneider, A., and Marahiel, M. A. 1998. Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*. *Arch. Microbiol.* 169:404-410.
- Scholz-Schroeder, B. K., Hutchison, M. L., Grgurina, I., and Gross, D. C. 2001a. The contribution of syringopeptin and syringomycin to virulence of *Pseudomonas syringae* pv. *syringae* strain B301D on the basis of *sypA* and *syrB1* biosynthesis mutant analysis. *Mol. Plant-Microbe Interact.* 14:336-348.
- Scholz-Schroeder, B. K., Soule, J. D., Lu, S.-E., Grgurina, I., and Gross, D. C. 2001b. A physical map of the syringomycin and syringopeptin gene clusters localized to an approximately 145-kb DNA region of *Pseudomonas syringae* pv. *syringae* strain B301D. *Mol. Plant-Microbe Interact.* 14:1426-1435.
- Schwarzer, D., Mootz, H. D., and Marahiel, M. A. 2001. Exploring the impact of different thioesterase domains for the design of hybrid peptide synthetases. *Chem. Biol.* 8:997-1010.
- Scott-Craig, J. S., Panaccione, D. G., Pocard, J.-A., and Walton, J. D. 1992. The cyclic peptide synthetase catalyzing HC-toxin production in the filamentous fungus *Cochliobolus carbonum* is encoded by a 15.7-kilobase open reading frame. *J. Biol. Chem.* 267:26044-26049.
- Stachelhaus, T., and Marahiel, M. A. 1995. Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *J. Biol. Chem.* 270:6163-6169.
- Stachelhaus, T., Mootz, H. D., Bergendahl, V., and Marahiel, M. A. 1998. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J. Biol. Chem.* 273:22773-22781.
- Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* 6:493-505.
- Staskawicz, B., Dahlbeck, D., Keen, N. T., and Napoli, C. 1987. Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. *J. Bacteriol.* 169:5789-5794.
- Stein, T., and Vater, J. 1996. Amino acid activation and polymerization at modular multienzymes in nonribosomal peptide biosynthesis. *Amino Acids* 10:201-227.
- Stover, C. K., Pham, X. Q., Erwin, A. L., Mizoguchi, S. D., Warrenner, P., Hickey, M. J., Brinkman, F. S., Hufnagle, W. O., Kowalik, D. J., Lagrou, M., Garber, R. L., Goltry, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L. L., Coulter, S. N., Folger, K. R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G. K., Wu, Z., and Paulsen, I. T. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature* 406:959-964.
- Trauger, J. W., Kohli, R. M., Mootz, H. D., Marahiel, M. A., and Walsh, C. T. 2000. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* 407:215-218.
- Turgay, K., Krause, M., and Marahiel, M. A. 1992. Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes. *Mol. Microbiol.* 6:529-546.
- Weber, G., Schörgendorfer, K., Schneider-Scherzer, E., and Leitner, E. 1994. The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genetics* 26:120-125.
- Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.
- Zhang, J.-H., Quigley, N. B., and Gross, D. C. 1995. Analysis of the *syrB* and *syrC* genes of *Pseudomonas syringae* pv. *syringae* indicates that syringomycin is synthesized by a thiotemplate mechanism. *J. Bacteriol.* 177:4009-4020.

AUTHOR-RECOMMENDED INTERNET RESOURCES

National Center for Biotechnology Information: www.ncbi.nlm.nih.gov/
 Prosite Database: www.expasy.org/prosite/
 Prediction of amino acid recognition by nonribosomal peptide synthetase adenylation domains database: raynam.chm.jhu.edu/~nrps/
 TIGR Microbial Database: www.tigr.org/tldb/mdb/mdbcomplete.html;
 NRPS prediction blast server, Johns Hopkins University website: raynam.chm.jhu.edu/~nrps/