

DOI: 10.1002/cbic.200500425

Deciphering the Biosynthetic Codes for the Potent Anti-SARS-CoV Cyclodepsipeptide Valinomycin in *Streptomyces tsusimaensis* ATCC 15141

Yi-Qiang Cheng^{*[a]}

Valinomycin was recently reported to be the most potent agent against severe acute respiratory-syndrome coronavirus (SARS-CoV) in infected Vero E6 cells. Aimed at generating analogues by metabolic engineering, the valinomycin biosynthetic gene cluster has been cloned from *Streptomyces tsusimaensis* ATCC 15141. Targeted disruption of a nonribosomal peptide synthetase (NRPS) gene abolishes valinomycin production, which confirms its predicted nonribosomal-peptide origin. Sequence analysis of the

NRPS system reveals four distinctive modules, two of which contain unusual domain organizations that are presumably involved in the generation of biosynthetic precursors D- α -hydroxyisovaleric acid and L-lactic acid. The respective adenylation domains in these two modules contain novel substrate-specificity-conferring codes that might specify for a class of hydroxyl acids for the biosynthesis of the depsipeptide natural products.

Introduction

Severe acute respiratory syndrome (SARS) was the first pandemic disease of the twenty-first century, which caused significant morbidity, mortality, and economic loss worldwide.^[1] The causal agent, a new coronavirus (CoV), is believed to be of animal origin.^[2–6] SARS-CoV has a large RNA genome that mutates rapidly.^[7,8] The possibility of a SARS resurgence warrants the need for anti-SARS-CoV therapeutics.^[9,10] Great efforts have been devoted to the development of effective vaccines and antiviral drugs against SARS-CoV.^[11–13] In one study, a cell-based assay was employed to screen a large chemical library and 50 compounds were identified to have anti-SARS-CoV activities.^[14] Among those compounds, valinomycin (VLM, C₅₄H₉₀N₆O₁₈, MW = 1111.5; Scheme 1), a natural product produced by several *Streptomyces* isolates,^[15–19] demonstrated the highest potency against SARS-CoV in infected Vero E6 cells with an EC₅₀ = 0.85 μ M.^[14]

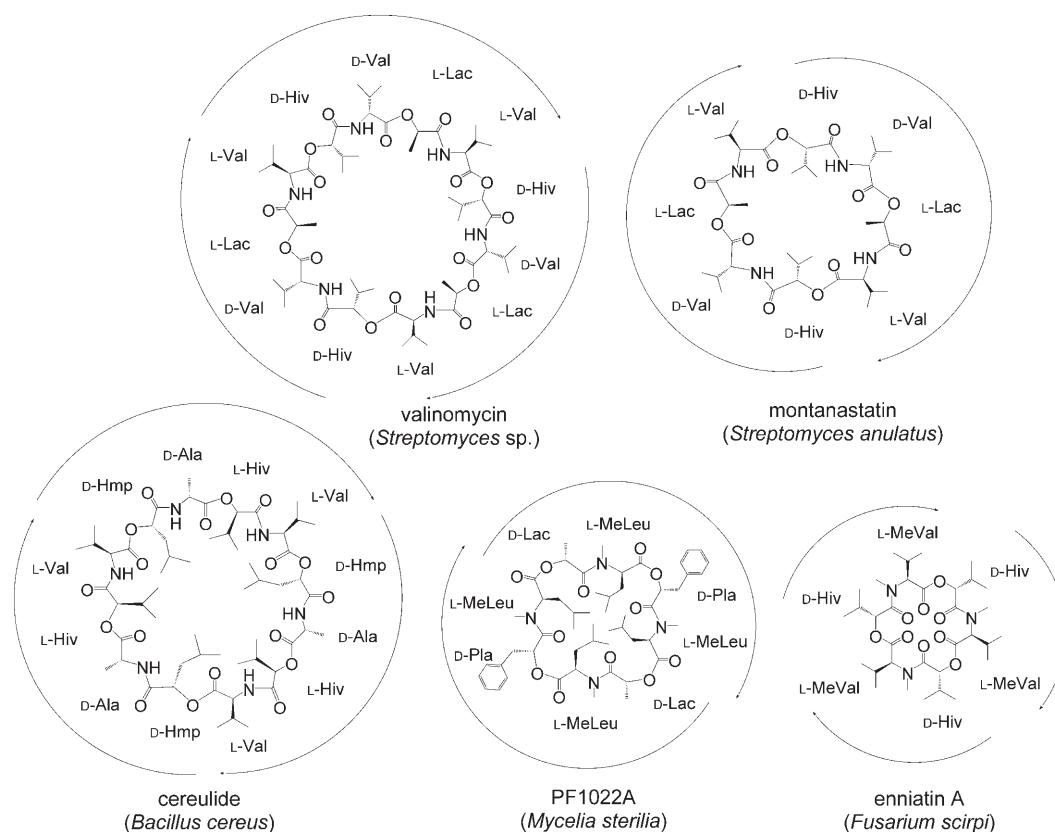
VLM's mode of anti-SARS-CoV action is not yet known, but its antiviral activity is just one of its many inhibitory effects. Previous studies have revealed VLM's antifungal,^[20] insecticidal-nematocidal,^[21] and antibacterial activities (moderately against *Mycobacterium tuberculosis* in vitro).^[22] It also displays apoptosis-inducing activity in mammalian cell types, inhibition of human NK cell function,^[23] and antitumor activities.^[18,24] Unfortunately, the cytotoxicity of VLM has so far prevented its clinical use but structural analogues of VLM could provide an alternative.

VLM has a cyclododecdepsipeptide structure that contains a three-repeating sequence of a tetradepsipeptide basic unit, D- α -hydroxyisovaleryl-D-valyl-L-lactyl-L-valyl, to form a symmetric 36-membered ring molecule.^[25–27] Feeding experiments have

identified D- α -hydroxyisovaleric acid (D-Hiv), L-lactic acid (L-Lac) and L-valine (L-Val) as the biosynthetic precursors in VLM biosynthesis. VLM forms an ionophore that specifically modulates potassium ion transport across biological membranes.^[28,29] Many peptide antibiotics act as ion channels.^[30]

The cyclic peptide structural features of VLM suggest that it is likely biosynthesized by a nonribosomal peptide synthetase (NRPS) system.^[31,32] A previous genetic study claimed to have identified two genetic loci that appeared to be involved in VLM biosynthesis in *S. levoris* A-9,^[19] but to date no further results (e.g., gene sequence) have been published or patented. In this study, the gene cluster for VLM biosynthesis (designated as *vlm*) in *S. tsusimaensis* ATCC 15141 has been cloned, sequenced, and partially characterized. Sequence analysis revealed good alignment between modular NRPS organization and the biosynthetic precursor incorporation for VLM biosynthesis. In addition, the molecular basis for nonribosomal peptide intermediate oligomerization mediated by a C-terminal iterative type I thioesterase (TE) is discussed. This work sets the stage for generating VLM analogues through metabolic engineering of the VLM biosynthetic pathway and is aimed at acquiring potent, highly specific, and low-cytotoxic anti-SARS-CoV agents and/or other anti-infective therapeutics.

[a] Prof. Dr. Y.-Q. Cheng
Department of Biological Sciences, Biotechnology Program
University of Wisconsin–Milwaukee
Milwaukee, WI 53211 (USA)
Fax: (+1) 414-229-3926
E-mail: ycheng@uwm.edu



Scheme 1. Chemical structures and the producing strains of valinomycin and related symmetric cyclodepsipeptides. Curved arrows indicate the repetition of basic biosynthetic units. D-Hiv = D- α -hydroxyisovaleric acid; D-Hmp = D- α -hydroxy-4-methylpentanoic acid; L-Lac = L-lactic acid; L-MeLeu = L-N-methyl-leucine; D-Pla = D-phenyl-lactic acid.

Results and Discussion

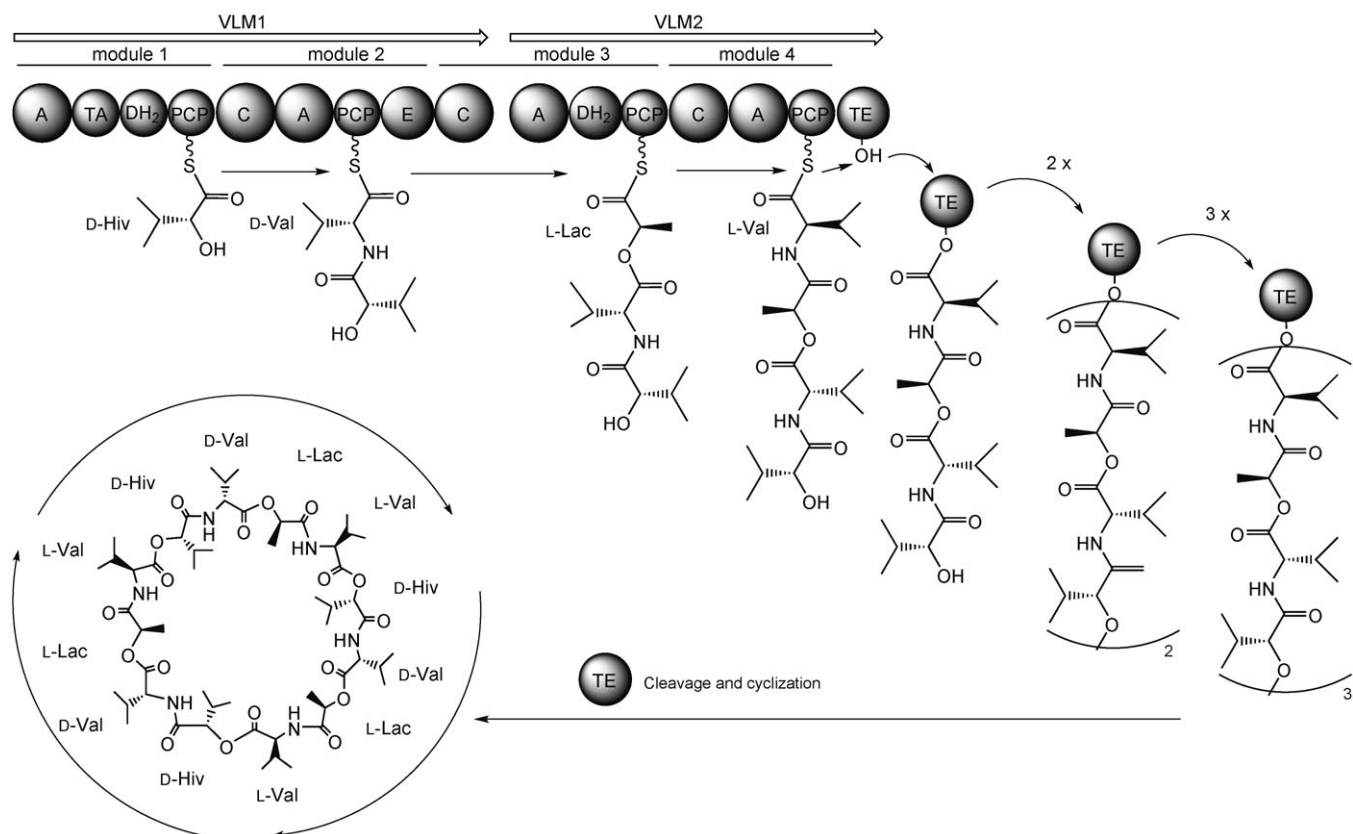
Identification of a gene fragment essential for valinomycin biosynthesis in *Streptomyces tsusimaensis* ATCC 15141

Based on the accumulated knowledge about the biosynthesis of nonribosomal peptide natural products,^[31,32] particularly of enniatins^[33,34] and related cyclodepsipeptides^[35–38] (Scheme 1), biosynthesis of VLM by a type B NRPS system^[39] was predicted as follows. Such an NRPS system would contain four distinctive modules. Each module would be responsible for the selection, activation, epimerization (when necessary), and incorporation of a precursor (also called building block: D-Hiv, L-Val, L-Lac, or L-Val) in a linear fashion, to form a tetradepsipeptide basic unit covalently tethered as an ester to the active-site serine of the C-terminal iterative TE domain. Two additional rounds of basic unit buildup would generate a linear 12-building block intermediate stalled on the TE. Subsequently, TE would cleave off the intermediate and catalyze the head-to-tail cyclization to afford the mature VLM product (Scheme 2). Alternatively, a terminal peptidyl carrier protein–condensation (PCP–C) didomain might be substituted for the TE for basic unit buildup and cyclization—a mechanism seen in the biosynthesis of enniatins.^[39] Considering that a typical NRPS module is encoded by an average 4.5 kb DNA, the putative VLM NRPS assembly line would be expected to require approximately 18 kb of DNA. Genes involved in regulation and self-resistance would also be anti-

ciated in the putative *vlm* gene cluster.^[40] With the prediction of a large *vlm* gene cluster, the “genome-scanning” approach^[41] was adopted to clone the cluster. The key element of this approach is to sequence a certain number of random genomic clones and to identify candidate gene(s) involved in natural product biosynthesis.

A genome sampling library was constructed and subjected to sequencing. Among 175 random genomic sequence tags obtained from 192 reads, one sequence tag, Plate1E10-T7 (810 bp), showed significant homology to genes encoding typical NRPSs, such as the TeiC in the teicoplanin biosynthetic gene cluster (accession no. CAG15011) and pristnamycin I synthetase 3 and 4 (accession no. CAA67249). The putative NRPS gene that contains the Plate1E10-T7 tag was therefore designated as “*St-nrps*”.

To test whether this candidate *St-nrps* gene is essential for VLM biosynthesis, mutants were created by targeted gene disruption of the *St-nrps* gene in *S. tsusimaensis*. Subsequently VLM production in wild type *S. tsusimaensis* and in four independent mutants was analyzed by fermentation, methanol extraction, solid-phase fractionation, and finally by liquid chromatography–mass spectrometry (LC–MS). VLM gives a very weak UV signal at a wavelength of 210 nm because it does not contain a chromophore, but its positive ion [VLM+H]⁺ has a strong MS signal of 1112.6 (*m/z*). Wild type *S. tsusimaensis* produced about 0.2 mg mL⁻¹ of VLM under the fermentation con-



Scheme 2. Proposed model for the biosynthesis of VLM in *S. tsusimaensis* ATCC 15141. The subprotein organization of VLM1 and VLM2 is dissected into modules and domains sequentially. Domain name abbreviations are defined in the text. TA and DH₂ domains are hypothetical. An alternative model where the terminal TE domain is replaced by a PCP–C didomain for intermediate buildup and cyclization is explained in the text but excluded from the scheme for simplicity.

ditions tested when calibrated with standard VLM. It was found that disruption of *St-nrps* did not affect the vegetative growth or sporulation of *S. tsusimaensis* mutants but completely abolished VLM production (data not shown). This experiment unambiguously confirmed that the *St-nrps* gene is essential for VLM biosynthesis in *S. tsusimaensis*.

Cloning, sequencing, and in silico analysis of the VLM biosynthetic gene cluster (*vlm*)

Plate1E10 insert DNA was labeled and used as a probe to screen an *S. tsusimaensis* cosmid library and three overlapping cosmids that cover the *vlm* gene cluster region (~80 kb) were identified (Figure 1). End sequencing of cosmid DNAs suggested that cosmid 15 might contain the entire *vlm* gene cluster. Sequencing and assembling of a shotgun library constructed with nebulized cosmid 15 DNA generated a

contig with a calculated 5.8-fold coverage. This contig contained a 39345 bp DNA sequence that included the entire *vlm* gene cluster. Sequence analysis of the contig revealed eighteen open reading frames (ORFs). The overall GC content of this contig is 70.68%; this indicates a typical *Streptomyces* origin.^[42] The function of deduced gene products of individual ORFs was assigned based on similarity to known proteins in the database

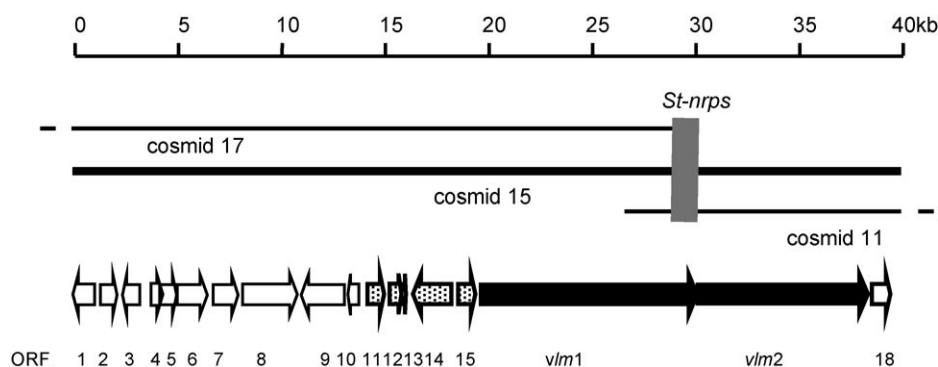


Figure 1. Genetic map and gene arrangement of the *vlm* biosynthetic gene cluster and flanking regions in *S. tsusimaensis* ATCC 15141. Genes in black (*vlm1* and *vlm2*) are critical VLM biosynthetic genes, ORFs in gray (ORF11–15) are putative genes within the *vlm* gene cluster, and ORFs in white are genes outside of the *vlm* gene cluster. *St-nrps* was the original sequence tag identified to be essential for VLM biosynthesis and used as a molecular probe to obtain a series of overlapping cosmids; cosmid 15 was shotgun sequenced.

Table 1. Deduced functions for the ORFs in the *vlm* biosynthetic gene cluster and flanking regions.

ORF (bp ^[a] bp)	Size (amino acids)	Protein homologue by accession no.	Proposed function
Upstream region:			
ORF1 (1 < 949 ^[b])	316+	CAA19245 from <i>Streptomyces coelicolor</i>	unknown
ORF2 (1366 > 2244)	292	NP_767597 from <i>Bradyrhizobium japonicum</i>	unknown
ORF3 (2789 < 3520)	243	CAB93439 from <i>S. coelicolor</i>	TetR-family regulatory protein
ORF4 (3781 > 4410)	209	none	unknown
ORF5 (4427 > 5272)	281	BAC75195 from <i>Streptomyces avermitilis</i>	N5, N10-MTHMP reductase ^[c]
ORF6 (5475 > 6959)	494	AAF61246 from <i>Streptomyces griseus</i>	penicillin-binding protein
ORF7 (7184 > 8311)	375	CAD55532 from <i>S. coelicolor</i>	unknown
ORF8 (8509 > 11103)	864	Q82V2 of <i>S. avermitilis</i>	FO synthase ^[d]
ORF9 (11 670 < 13 808)	712	BAC70248 from <i>S. avermitilis</i>	unknown
ORF10 (13 860 < 14 288)	142	BAC70249 from <i>S. avermitilis</i>	unknown
The <i>vlm</i> gene cluster:			
ORF11 (14 516 > 15 310)	264	NP_631398 from <i>S. coelicolor</i>	necrosis inducing factor
ORF12 (15 380 > 15 877)	165	CAC42141 from <i>S. coelicolor</i>	unknown membrane-associated protein
ORF13 (15 877 > 16 086)	69	CAB58311 from <i>S. coelicolor</i>	transcriptional regulator
ORF14 (16 138 < 18 186)	682	BAC71762 from <i>S. avermitilis</i>	unknown
ORF15 (18 814 > 19 548)	244	AAP09421 from <i>Bacillus cereus</i>	discrete type II thioesterase
<i>vlm1</i> (19 526 > 29 812)	3428	AAV98629 from <i>B. cereus</i>	NRPS (A-TA-DH ₂ -PCP-C-A-PCP-E-C) ^[e]
<i>vlm2</i> (29 835 > 37 802)	2655	AAV98629 from <i>B. cereus</i>	NRPS (A-DH ₂ -PCP-C-A-PCP-TE) ^[e]
Downstream region:			
ORF18 (37 938 > 39 266)	442	BAD83980 from <i>Corynebacterium glutamicum</i>	transposase

[a] < or > indicates the orientation of transcription; [b] incomplete; [c] N5,N10-methylenetetrahydromethanopterin reductase; [d] 7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase; [e] abbreviations for domain names are defined in the text; TA and DH₂ domains are hypothetical.

(Table 1). Among the ORFs identified, *vlm1* and *vlm2* encode NRPSs that constitute an assembly line for VLM biosynthesis (see later for details); *vlm1* contains the original sequence tag Plate1E10-T7 (equivalent to the 28 911–29 719 bp DNA on the contig).

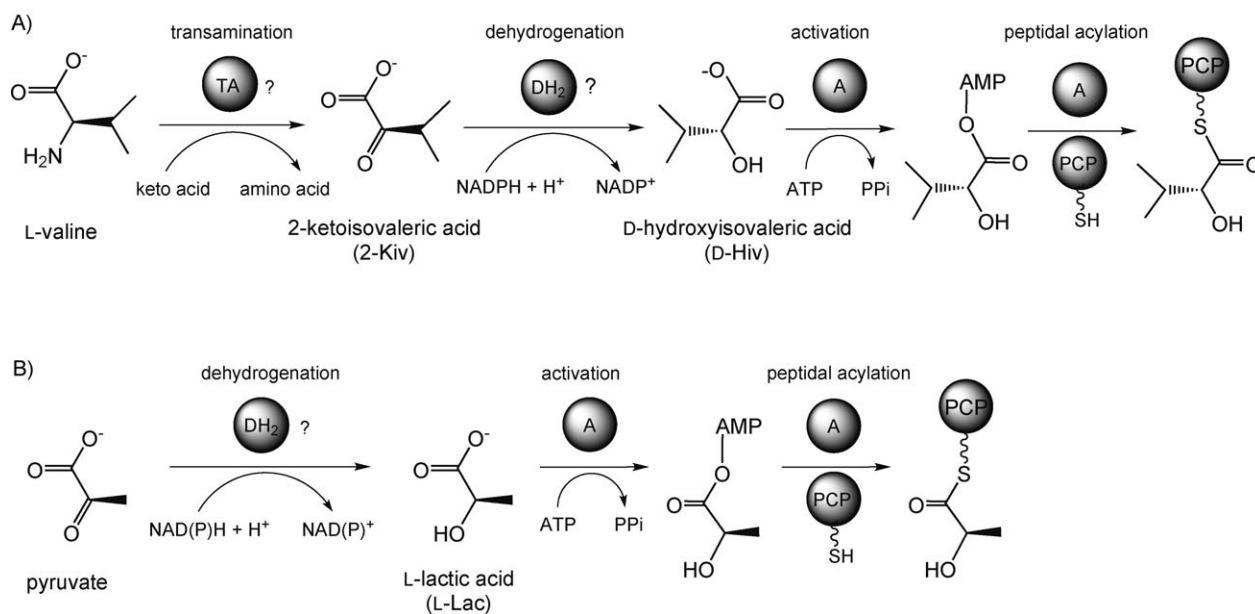
The *vlm* gene cluster was predicted to encompass seven ORFs, bordered by ORF11 and *vlm2*, respectively (Figure 1), based on the following observations: 1) downstream of *vlm2*, ORF18 encodes a putative transposase. Many natural product biosynthetic gene clusters have boundaries flanked by transposable elements, which suggests an origin of horizontal gene transfer; 2) upstream of *vlm1*, ORF11/12/13 appear to be co-transcribed. ORF13 encodes a transcriptional regulator containing a helix-turn-helix motif, which indicates its DNA-binding property. Regulatory genes are usually present in biosynthetic gene clusters; 3) ORF15 encodes a discrete type II TE, a component often found in biosynthetic pathways for nonribosomal peptides and polyketides. Type II TEs are believed to have a proof-reading function during the cleavage of misprimed thioesters from carrier proteins;^[43] 4) ORF9/10 appear to be cotranscribed in an orientation opposite to that of ORF11/12/13; 5) further upstream, ORF8 is a housekeeping gene.

A correlation between the modular organization of the VLM assembly line and structure

NRPSs encoded by *vlm1/2* contain sixteen distinctive domains organized into four modules followed by a C-terminal iterative TE (Scheme 2). This modular organization is consistent with the biosynthesis of the tetradepsipeptide basic unit $\text{D-}\alpha\text{-hydroxyisovaleryl-D-valyl-L-lactyl-L-valyl}$. Module 1 is the initiation module that contains four functional domains, adenylation (A; designated as VLM1A1), hypothetical transaminase (TA), hypo-

thetical dehydrogenase (DH₂), and PCP. The presence of these two hypothetical domains in module 1 correlates well with a proposed mechanism for the generation of the biosynthetic precursor D-Hiv (Scheme 3). According to a study of enniatin biosynthesis in *Fusarium sambucinum*, D-Hiv was postulated to be derived from L-Val by sequential transamination (catalyzed by an uncharacterized TA) and dehydrogenation (catalyzed by a purified discrete D-Hiv DH_2).^[44] Generation of precursor D-Hiv in VLM biosynthesis might follow the same route, catalyzed by discrete TA and DH₂ enzymes yet to be discovered, or by the integral hypothetical TA and DH₂ domains in this module. Module 2 contains four standard NRPS domains: C, A (designated as VLM1A2), PCP, and epimerase (E). Module 3 contains four domains: C, A (designated as VLM2A3), hypothetical DH₂, and PCP. This hypothetical DH₂ domain is speculated to act as a L-lactate dehydrogenase to convert pyruvate into L-Lac which is then acylated onto PCP (Scheme 3).^[44] Generation of lactate from pyruvate is a common biochemical reaction in many microorganisms that undergo fermentation. However, dedicated L-Lac DH_2 has not been characterized from any known natural product biosynthetic pathway. Module 3 is split into two proteins, a phenomenon increasingly found in natural product biosynthetic pathways.^[45] Module 4 contains a minimal set of three NRPS domains: C, A (designated as VLM2A4), and PCP.

Collectively, amino acid sequence alignment of four A domains on VLM1/2 revealed that they all contain the ten conserved signature motifs,^[46] albeit considerable amino-acid substitutions have occurred in VLM1A1 and VLM2A3 (data not shown). Phylogenetic display of the alignment indicates that VLM1A1 and VLM2A3 form a distinctive branch whereas VLM1A2 and VLM2A4 cluster with GrsA_A, which serves as a model A domain with a defined substrate specificity for Phe and a known structure;^[47] DhbE, a stand-alone A domain which



Scheme 3. Scheme of generation and incorporation of two precursors in VLM biosynthesis. A) D-Hiv generation, activation, and peptidyl acylation; B) L-Lac generation, activation, and peptidyl acylation.

activates the carboxyl acid DBH (2,3-dihydroxybenzoic acid) from the bacillibactin biosynthetic pathway,^[48] is distantly related to all of them. At substrate specificity level, VLM1A2 and VLM2A4 contain conserved NRPS codes^[49,50] (DALWLGGTFK and DAFWVGGTFK, respectively) that are predicted to specify for L-Val. In contrast, extracted NRPS codes from VLM1A1 or VLM2A3 (AALWIADSGK or VVIWIAEQHK, respectively) do not yield any predictable substrate. Remarkably, both VLM1A1 and VLM2A3 have the highly conserved Asp at position 235 (first position in the NRPS codes; numbered by GrsA_A sequence) replaced by an Ala or Val. Similarly, DhbE also has the conserved Asp residue replaced by an Asn. This finding suggests that VLM1A1 and VLM2A3, together with DhbE, might represent a divergent group of A domains that specify for a class of hydroxyl acid substrates other than proteinogenic amino acids or carboxyl acids. In the proposed model for VLM biosynthesis (Scheme 2) module 1 and 3 are responsible for the incorporation of D-Hiv and L-Lac, respectively. The A domains in these two modules might have adapted themselves to select and activate hydroxyl acids. The fact that D-Hiv and L-Lac were identified by feeding experiments as the biosynthetic precursors for VLM biosynthesis^[25–27] supports the hypothesis that both D-Hiv and L-Lac might have been synthesized from the VLM assembly line before being activated by their respective A domains. Additional biochemical experiments are needed to prove the VLM1A1 and VLM2A3 substrate specificities for D-Hiv and L-Lac, respectively.

Immediately following module 4 is a C-terminal iterative TE domain (designated as VLM2_TE). Terminal TE domains are known to control the termination, release, and cyclization of growing chains in the biosynthetic process of many nonribosomal peptides and polyketides.^[51,52] An unsolved mystery regarding the biosynthesis of symmetric cyclo(depsi)peptides is the control of repetition of the basic units. As shown in

Scheme 1, VLM contains a three-repeating sequence of a tetrapeptide basic unit (designated as a 3×4 mode), whereas montanastatin (C₃₆H₆₀N₄O₁₂, M_w = 740.9)^[18] contains the exact same tetrapeptide basic unit that repeats only twice (a 2×4 mode). This observation raises the question of whether VLM and montanastatin are produced by the same biosynthetic pathway with relaxed control of basic unit repetition, or by two independent pathways, each of which has a stringent control of basic-unit repetition. A low level of montanastatin was detected by LC-MS in the wild-type strain of *S. tsusimaensis* but not in the *St-nrps*⁻ mutants. This result suggests that montanastatin is likely the shunt product of VLM, produced by the same VLM biosynthetic pathway with a relaxed TE control of basic unit repetition. In comparison to VLM and montanastatin, enniatin A has a 3×2 mode, whereas cereulide and PF1022A have a 3×4 and 2×4 mode, respectively. Not included in Scheme 1 are additional known symmetric nonribosomal peptide natural products—bacillibactin, enterobactin, and gramicidin S^[39]—which have a 3×2, 3×2, or 2×5 mode, respectively. Iterative TE domains in the pathways for the biosynthesis of these symmetric oligomeric products are believed to hold the basic unit until an appropriate number of units are condensed and then proceed with cleavage and cyclization.^[51] However, the exact mechanism of controlling the number of repetitions remains unknown. An amino acid sequence alignment (data not shown) of VLM2_TE with DEBS3_TE (accession no. Q03133), SrfA-C_TE (accession no. Q08787), Bs-EntF_TE (accession no. P45745), Ec-EntF_TE (accession no. P11454), CesB_TE (accession no. AY691650), and GrsB_TE (accession no. P14688) showed that VLM2_TE contains a standard catalytic triad, Ser80/His207/Asp107 (residues numbered corresponding to those of SrfA-C_TE) and a conserved substrate-binding motif, Gly-Xxx-Ser-Xxx-Gly, where Xxx stands for any amino acid.^[53] No residues stand out to suggest a correlation between con-

served amino acid residues, domain, or motif, and the unique oligomerization function of VLM2_TE and other TEs (Bs-EntF_TE, Ec-EntF_TE, CesB_TE, and GrsB_TE) that are from the pathways for the biosynthesis of oligomerized natural products VLM, bacillibactin, enterobactin, cereulide, and gramicidin S, respectively. The biochemical mechanism for the control of the number of basic unit repetitions in the biosynthesis of cyclo-(depsi)peptides remains of interest for future studies.

In conclusion, the VLM biosynthetic gene cluster (*vlm*) was cloned from *S. tsusimaensis* ATCC 15141. This *vlm* gene cluster contains two biosynthetic genes encoding type B iterative NRPSs. These NRPSs contribute an assembly line that consists of sixteen distinctive domains organized into four modules; each module is expected to incorporate a biosynthetic precursor to form a tetradepsipeptide basic unit. A terminal TE domain might mediate the oligomerization and cyclization of three-repeating sequence of the basic unit to afford the mature VLM product (Scheme 2). The cloning and decoding of VLM biosynthetic machinery offers an opportunity for genetic manipulation of the pathway. Genetic engineering of the VLM biosynthesis pathway serves to complement chemical synthesis to generate novel VLM analogues for screening potent, highly specific, and low cytotoxic anti-SARS-CoV agents and/or other anti-infective therapeutics.

Experimental Section

Bacterial strains and culture conditions: *E. coli* strains DH5 α , XL1-Blue MR (Stratagene), and S17-1^[54] were used in this work, according to standard procedures.^[55] *S. tsusimaensis* ATCC 15141, a VLM producer,^[19] was purchased from the American Type Culture Collection. *S. tsusimaensis* was grown at 30 °C in tryptic soy broth (TSB) medium supplemented with sucrose (34%) and MgCl₂ (25 mM) for mycelia harvesting; on ISP-2 agar for sporulation; on modified ISP-4 agar supplemented with tryptone (0.1%) and yeast extract (0.05%) for conjugation between *E. coli* S17-1 and *S. tsusimaensis*; and in a composite fermentation medium containing Diaion HP-20 (5%)^[56] for VLM production. TSB, ISP-2, and ISP-4 media were from Difco Laboratories. *S. tsusimaensis* was found to be sensitive to apramycin (50 $\mu\text{g mL}^{-1}$), nalidixic acid (100 $\mu\text{g mL}^{-1}$), and thiostrepton (10 $\mu\text{g mL}^{-1}$), which suggested that common *Streptomyces* genetic practices^[42] could be feasible in this strain.

DNA manipulations and library construction: Plasmid DNA extraction and DNA fragment recovery from agarose gels were carried out with commercial kits (Qiagen). DNA digestion, agarose gel electrophoresis, ligation, and transformation were done by standard protocols.^[55] Total *S. tsusimaensis* DNA was extracted from mycelia by the Kirby mix procedure.^[42] Three genomic DNA libraries were constructed in this work. A genome sampling library^[41] was constructed with the 2–4 kb fraction of *Mbol* partially digested total DNA into the *Bam*HI site of pGEM-3zf (Promega). A cosmid library was constructed with the 35–45 kb fraction of *Mbol* partially digested total DNA into the *Bam*HI site of SuperCos 1 vector (Stratagene). A shotgun sequencing library was constructed with the 1–2 kb fraction of nebulized cosmid 15 DNA into the pCR[®]4Blunt-TOPO vector according to manufacturer's instructions (Invitrogen). For library screening and Southern blot analysis, digoxigenin labeling of the Plate1E10 probe DNA, hybridization, and detection were performed according to the manufacturer's protocols (Roche).

DNA sequencing and sequence analysis: For the identification of candidate gene fragments encoding NRPSs, two 96-well plates (192 samples) of the genome sampling library were submitted for end sequencing with the T7 primer to generate random genomic sequence tags (DNA Sequencing Laboratory, University of Wisconsin–Madison, WI, USA). Individual sequencing reads were manually examined and subjected to homology search against GenBank by using the Blastx algorithm.^[57] The candidate *St-nrps* gene which contained clone Plate1E10, was subjected to additional sequencing by primer walking and mapping. For shotgun sequencing of cosmid 15, eight 96-well plates of the shotgun library (768 samples) were submitted for end-sequencing with T3 and T7 primers at Rexagen Co. (Seattle, WA, USA). Contig assembling and sequence polishing were performed by the service company. Sequence quality of the contig containing the *vlm* gene cluster was examined by PCR amplification with genomic DNA as template and resequencing of any regions of interest on the contig. ORF prediction, DNA translation, and protein sequence alignment were performed by using the Lasergene Package (DNASTar, Madison, WI).

Construction of targeted gene disruption mutants of *S. tsusimaensis* and examination of VLM production by using LC–MS: A 2.3 kb *Eco*RI–*Pst*I fragment from Plate1E10 was cloned into a suicide conjugative vector pOJ260 (replicative in *E. coli* but not in *Streptomyces*)^[58] to make a targeted gene disruption construct pYC02-36d. The construct was first transformed into *E. coli* S17-1 cells and then mobilized into *S. tsusimaensis* cells by conjugation.^[56] Four independent apramycin-resistant conjugants were selected and were verified by Southern blot analysis for the correct integration of plasmid DNA into the host chromosome. Wild-type *S. tsusimaensis* and the four mutants YC02-38a/b/c/d were then grown in the fermentation medium for 6 days at 30 °C; their mycelia and the HP-20 resin were collected by filtration and the whole mass was lyophilized and extracted twice with two volumes of methanol. Methanol extracts were pooled and a portion was further fractionated by an Oasis[®] HLB Plus SPE cartridge (Waters, Milford, MA, USA). A fraction eluted by methanol (90%) was subjected to LC–MS analysis. LC–MS spectra were obtained with an Agilent 1000 LC/MSD Trap SL instrument (Agilent, Palo Alto, CA, USA). The eluents used contained formic acid (0.1%) in water (eluent A) or in acetonitrile (eluent B). A 15 min gradient from 10% eluent B to 90% eluent B was achieved in a 2.1 \times 110 mm Zorbax SB-AQ column (Agilent) with a flow rate of 0.5 mL min⁻¹. UV_{210nm} LC signal and positive ion MS signal were recorded. The VLM peak appeared at 12.7 min. The VLM standard used in calibration was purchased from Sigma–Aldrich.

Nucleotide sequence accession number: The nucleotide sequence reported here has been deposited in the GenBank database under accession no. DQ174261.

Acknowledgements

The work was supported by funds from the University of Wisconsin–Milwaukee. The author thanks Mary Lynne Collins (University of Wisconsin–Milwaukee) for providing *E. coli* S17-1 strain, and Gary Girdaukas (University of Wisconsin–Madison, School of Pharmacy Analytical Instrumentation Center) for performing LC–MS analysis. The author further thanks Mark McBride and Andrea Matter for critical reading of this manuscript and anonymous reviewers for their many suggestions.

Keywords: antiviral agents · biosynthesis · natural products · nonribosomal peptide synthetase · valinomycin

- [1] B. B. Finlay, R. H. See, R. C. Brunham, *Nat. Rev. Microbiol.* **2004**, *2*, 602–607.
- [2] W. Li, Z. Shi, M. Yu, W. Ren, C. Smith, J. H. Epstein, H. Wang, G. Cramer, Z. Hu, H. Zhang, J. Zhang, J. McEachern, H. Field, P. Daszak, B. T. Eaton, S. Zhang, L. F. Wang, *Science* **2005**, *310*, 676–679.
- [3] C. Drosten, W. Preiser, S. Gunther, H. Schmitz, H. W. Doerr, *Trends Mol. Med.* **2003**, *9*, 325–327.
- [4] T. G. Ksiazek, D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, L. J. Anderson, *N. Engl. J. Med.* **2003**, *348*, 1953–1966.
- [5] R. A. Fouchier, T. Kuiken, M. Schutten, G. van Amerongen, G. J. van Doornum, B. G. van den Hoogen, M. Peiris, W. Lim, K. Stohr, A. D. Osterhaus, *Nature* **2003**, *423*, 240.
- [6] T. Kuiken, R. A. Fouchier, M. Schutten, G. F. Rimmelzwaan, G. van Amerongen, D. van Riel, J. D. Laman, T. de Jong, G. van Doornum, W. Lim, A. E. Ling, P. K. Chan, J. S. Tam, M. C. Zambon, R. Gopal, C. Drosten, S. van der Werf, N. Escriou, J. C. Manuguerra, K. Stohr, J. S. Peiris, A. D. Osterhaus, *Lancet* **2003**, *362*, 263–270.
- [7] K. Stadler, V. Masignani, M. Eickmann, S. Becker, S. Abrignani, H. D. Klenk, R. Rappuoli, *Nat. Rev. Microbiol.* **2003**, *1*, 209–218.
- [8] S. M. E. C. Chinese, *Science* **2004**, *303*, 1666–1669.
- [9] L. A. Meyers, B. Pourbohloul, M. E. Newman, D. M. Skowronski, R. C. Brunham, *J. Theor. Biol.* **2005**, *232*, 71–81.
- [10] D. M. Skowronski, C. Astell, R. C. Brunham, D. E. Low, M. Petric, R. L. Roper, P. J. Talbot, T. Tam, L. Babiuk, *Annu. Rev. Med.* **2005**, *56*, 357–381.
- [11] S. Jiang, Y. He, S. Liu, *Emerging Infect. Dis.* **2005**, *11*, 1016–1020.
- [12] S. Shigeta, T. Yamase, *Antiviral Chem. Chemother.* **2005**, *16*, 23–31.
- [13] G. H. Du, W. Wang, K. D. Cheng, L. L. Cao, X. Y. Yang, P. Zhu, *Yaouxue Xuebao* **2004**, *39*, 236–240.
- [14] C. Y. Wu, J. T. Jan, S. H. Ma, C. J. Kuo, H. F. Juan, Y. S. Cheng, H. H. Hsu, H. C. Huang, D. Wu, A. Brik, F. S. Liang, R. S. Liu, J. M. Fang, S. T. Chen, P. H. Liang, C. H. Wong, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10012–10017.
- [15] H. Brockmann, G. Schmidt-Kastner, *Chem. Ber.* **1955**, *88*, 57–61.
- [16] H. Brockmann, H. Geeren, *Justus Liebigs Ann. Chem.* **1957**, *603*, 216–232.
- [17] M. A. Andersson, R. Mikkola, R. M. Kroppenstedt, F. A. Rainey, J. Peltola, J. Helin, K. Sivonen, M. S. Salkinoja-Salonen, *Appl. Environ. Microbiol.* **1998**, *64*, 4767–4773.
- [18] G. R. Pettit, R. Tan, N. Melody, J. M. Kieley, R. K. Pettit, D. L. Herald, B. E. Tucker, L. P. Mallavia, D. L. Doubek, J. M. Schmidt, *Bioorg. Med. Chem.* **1999**, *7*, 895–899.
- [19] J. B. Perkins, S. K. Guterman, C. L. Howitt, V. E. Williams II, J. Pero, *J. Bacteriol.* **1990**, *172*, 3108–3116.
- [20] H. Nishimura, U.S. Patent no. 3365363, **1968**.
- [21] E. L. Patterson, D. P. Wright, U.S. Patent no. 3520973, **1970**.
- [22] W. A. Taber, L. C. Vining, *Can. J. Microbiol.* **1957**, *3*, 953–965.
- [23] A. Paananen, K. Jarvinen, T. Sareneva, M. S. Salkinoja-Salonen, T. Timonen, E. Holtta, *Toxicology* **2005**, *212*, 37–45.
- [24] S. Kuroda, *Gan To Kagaku Ryoho* **1984**, *11*, 2625–2632.
- [25] J. C. MacDonald, G. P. Slater, *Can. J. Biochem.* **1968**, *46*, 573–578.
- [26] H. Ristow, J. Salnikow, H. Kleinkauf, *FEBS Lett.* **1974**, *42*, 127–130.
- [27] T. Anke, F. Lipmann, *FEBS Lett.* **1977**, *82*, 337–340.
- [28] M. S. Frant, J. W. Ross Jr., *Science* **1970**, *167*, 987–988.
- [29] B. C. Pressman, *Annu. Rev. Biochem.* **1976**, *45*, 501–530.
- [30] B. A. Wallace, *Bioessays* **2000**, *22*, 227–234.
- [31] R. Finking, M. A. Marahiel, *Annu. Rev. Microbiol.* **2004**, *58*, 453–488.
- [32] C. T. Walsh, *Science* **2004**, *303*, 1805–1810.
- [33] A. Haese, M. Schubert, M. Herrmann, R. Zocher, *Mol. Microbiol.* **1993**, *7*, 905–914.
- [34] R. Pieper, A. Haese, W. Schroder, R. Zocher, *Eur. J. Biochem.* **1995**, *230*, 119–126.
- [35] W. Weckwerth, K. Miyamoto, K. Iinuma, M. Krause, M. Glinski, T. Storm, G. Bonse, H. Kleinkauf, R. Zocher, *J. Biol. Chem.* **2000**, *275*, 17909–17915.
- [36] P. F. Horwood, G. W. Burgess, H. J. Oakey, *FEMS Microbiol. Lett.* **2004**, *236*, 319–324.
- [37] M. Toh, M. C. Moffitt, L. Henrichsen, M. Raftery, K. Barrow, J. M. Cox, C. P. Marquis, B. A. Neilan, *J. Appl. Microbiol.* **2004**, *97*, 992–1000.
- [38] M. Ehling-Schulz, N. Vukov, A. Schulz, R. Shaheen, M. Andersson, E. Martlbauer, S. Scherer, *Appl. Environ. Microbiol.* **2005**, *71*, 105–113.
- [39] H. D. Mootz, D. Schwarzer, M. A. Marahiel, *ChemBioChem* **2002**, *3*, 490–504.
- [40] M. F. Martin, P. Liras, *Annu. Rev. Microbiol.* **1989**, *43*, 173–206.
- [41] E. Zazopoulos, K. Huang, A. Staffa, W. Liu, B. O. Bachmann, K. Nonaka, J. Ahlert, J. S. Thorson, B. Shen, C. M. Farnet, *Nat. Biotechnol.* **2003**, *21*, 187–190.
- [42] T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, John Innes Foundation, Norwich, England, **2000**.
- [43] M. L. Heathcote, J. Staunton, P. F. Leadlay, *Chem. Biol.* **2001**, *8*, 207–220.
- [44] C. Lee, H. Gorisch, H. Kleinkauf, R. Zocher, *J. Biol. Chem.* **1992**, *267*, 11741–11744.
- [45] S. C. Wenzel, R. Muller, *Curr. Opin. Chem. Biol.* **2005**, *9*, 447–458.
- [46] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem. Rev.* **1997**, *97*, 2651–2674.
- [47] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO. J.* **1997**, *16*, 4174–4183.
- [48] J. J. May, T. M. Wendrich, M. A. Marahiel, *J. Biol. Chem.* **2001**, *276*, 7209–7217.
- [49] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493–505.
- [50] G. L. Challis, J. Ravel, C. A. Townsend, *Chem. Biol.* **2000**, *7*, 211–224.
- [51] T. A. Keating, D. E. Ehmman, R. M. Kohli, C. G. Marshall, J. W. Trauger, C. T. Walsh, *ChemBioChem* **2001**, *2*, 99–107.
- [52] T. A. Keating, C. T. Walsh, *Curr. Opin. Chem. Biol.* **1999**, *3*, 598–606.
- [53] S. D. Bruner, T. Weber, R. M. Kohli, D. Schwarzer, M. A. Marahiel, C. T. Walsh, M. T. Stubbs, *Structure* **2002**, *10*, 301–310.
- [54] R. Simon, U. Priefer, A. Puehler, *Biotechnology (N.Y.)* **1983**, *1*, 784–791.
- [55] J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York, **2000**.
- [56] Y. Q. Cheng, G. L. Tang, B. Shen, *J. Bacteriol.* **2002**, *184*, 7013–7024.
- [57] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403–410.
- [58] M. Bierman, R. Logan, K. O'Brien, E. T. Seno, R. N. Rao, B. E. Schoner, *Gene* **1992**, *116*, 43–49.

Received: October 14, 2005