

# Functional Analysis of the Validamycin Biosynthetic Gene Cluster and Engineered Production of Validoxylamine A

Linquan Bai,<sup>1</sup> Lei Li,<sup>1,4</sup> Hui Xu,<sup>1,4</sup>  
Kazuyuki Minagawa,<sup>2,4</sup> Yi Yu,<sup>1</sup> Yirong Zhang,<sup>1</sup>  
Xiufen Zhou,<sup>1</sup> Heinz G. Floss,<sup>3</sup> Taifo Mahmud,<sup>2,\*</sup>  
and Zixin Deng<sup>1,\*</sup>

<sup>1</sup>Laboratory of Microbial Metabolism and School  
of Life Science & Biotechnology  
Shanghai Jiaotong University  
Shanghai 200030  
China

<sup>2</sup>Department of Pharmaceutical Sciences  
Oregon State University  
Corvallis, Oregon 97331

<sup>3</sup>Department of Chemistry  
University of Washington  
Seattle, Washington 98195

## Summary

A 45 kb DNA sequencing analysis from *Streptomyces hygroscopicus* 5008 involved in validamycin A (VAL-A) biosynthesis revealed 16 structural genes, 2 regulatory genes, 5 genes related transport, transposition/integration or tellurium resistance; another 4 genes had no obvious identity. The VAL-A biosynthetic pathway was proposed, with assignment of the required genetic functions confined to the sequenced region. A cluster of eight reassembled genes was found to support VAL-A synthesis in a heterologous host, *S. lividans* 1326. In vivo inactivation of the putative glycosyltransferase gene (*valG*) abolished the final attachment of glucose for VAL production and resulted in accumulation of the VAL-A precursor, validoxylamine, while the normal production of VAL-A could be restored by complementation with *valG*. The role of *valG* in the glycosylation of validoxylamine to VAL-A was demonstrated in vitro by enzymatic assay.

## Introduction

Validamycin A (VAL-A; Figure 1A), a weakly basic C<sub>7</sub>-aminocyclitol-containing antibiotic, was first isolated from *Streptomyces hygroscopicus* var. *limoneus*, and later from *S. hygroscopicus* var. *jinggangensis* 5008 (hereafter *S. hygroscopicus* 5008 or strain 5008). It is widely used, especially in Asia, to control sheath blight disease of rice plants and dumping-off of cucumber seedlings caused by the fungus *Pellicularia sasakii* (*Rhizoctonia solani*) [1]. In the presence of VAL-A, normal extension of the main fungal hyphae is switched to an abnormal branching at the tips, and further development of those growing fungi is repressed [2]. VAL-A was also found to have insecticidal activity, most likely resulting from its strong inhibition of trehalase, the trehalose-degrading enzyme. Trehalose is a primary storage carbohydrate in fungi (8%–10% of the dry cell

weight of *P. sasakii*) and is recognized as a characteristic blood sugar of insects. The enzyme trehalase appears to play an important metabolic role in these organisms in generating glucose for energy supply or for other physiological purposes. Its inhibition results in disruption of the glucose supply system of the fungus, leading to growth abnormality and death. Although in vitro experiments revealed that the pseudodisaccharide, validoxylamine A, is more active than VAL-A against trehalase [3], results of in vivo experiments suggest the opposite. It is proposed that the presence of the glucose moiety on VAL-A is essential for its efficient entry into fungal mycelia, in which it is hydrolyzed by an  $\alpha$ -glucosidase to yield the active pharmacophore, validoxylamine A [4]. VAL-A is also used as the source of valienamine, a pharmaceutically important precursor for the production of the antidiabetic drug, Voglibose (Basen), whose mechanism of action is to inhibit  $\alpha$ -glucosidase in the intestine, comparable to another antidiabetic agent, acarbose.

The biosynthesis of the C<sub>7</sub>-N-aminocyclitol moiety has been investigated to some extent in *Actinoplanes* sp. SE50/110 (acarbose-producer) [5–7], in *S. hygroscopicus* var. *limoneus* (VAL-A producer) [8–10], and, more recently, in *Microtetraspora spiralis* (pyralomicin producer) [11]. Feeding experiments with a number of isotopically labeled potential intermediates to *S. hygroscopicus* var. *limoneus* [8, 10, 12] demonstrated that 2-*epi*-5-*epi*-valiolone, 5-*epi*-valiolone, valienone, and validone were incorporated into VAL-A, leading to a proposed biosynthetic pathway to VAL-A, as shown in Figure 1B. Similar feeding experiments were also performed with *Actinoplanes* sp. SE50/110 [6, 7] and *Microtetraspora spiralis*, but with surprisingly different incorporation patterns. Only 2-*epi*-5-*epi*-valiolone and 5-*epi*-valiolone were found to be efficiently incorporated into pyralomicin [11], whereas 2-*epi*-5-*epi*-valiolone was the only one of these compounds incorporated into acarbose.

The gene cluster for acarbose biosynthesis has been cloned and a number of the proteins involved have been characterized, resulting in a proposed biosynthetic pathway to acarbose [13]. The pathway is initiated by a cyclization of sedoheptulose 7-phosphate, a C<sub>7</sub> sugar phosphate involved in the pentose phosphate pathway, catalyzed by the dehydroquinase synthase-like protein, AcbC, to give 2-*epi*-5-*epi*-valiolone (Figure 1B) [5]. The latter compound is then phosphorylated to its 7-phosphate derivative catalyzed by the kinase AcbM. Furthermore, the product of *acbO*, which was found in the same operon as *acbC* and *acbM*, was identified as a 2-*epi*-5-*epi*-valiolone 7-phosphate 2-epimerase [14]. Efficient incorporation of 2-*epi*-5-*epi*-valiolone into both VAL and acarbose strongly suggested that both biosynthetic pathways share the same cyclization step. However, further downstream the pathways to the two compounds may be different, as feeding experiments with the producing strains gave two different patterns. This leaves many open questions surrounding the biosynthesis of acarbose and VAL-A, which cannot be addressed by conventional feeding experiments alone.

\*Correspondence: taifo.mahmud@oregonstate.edu (T.M.); zxdeng@sjtu.edu.cn (Z.D.)

<sup>4</sup>These authors contributed equally to this work.

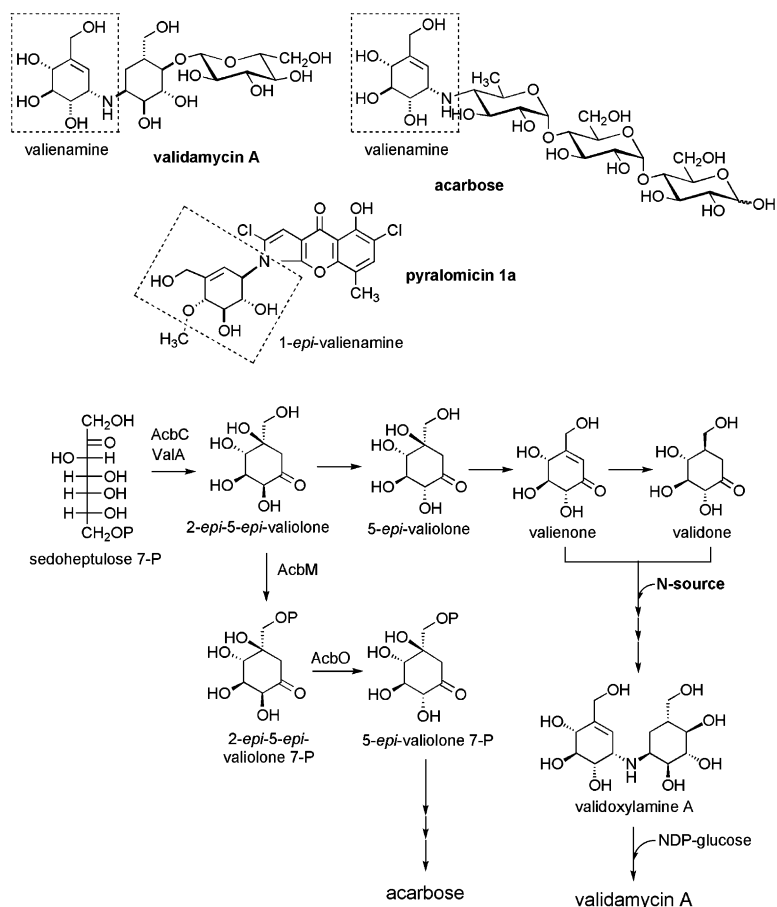


Figure 1. Valienamine-Containing C<sub>7</sub>N Amino-cyclitols and Proposed Biosynthetic Pathways to VAL-A and Acarbose

(Top) Chemical structures of VAL-A, acarbose, and pyralomicin 1a. Dashed-boxed regions show the cyclitol moiety shared by all three compounds.

(Bottom) Previously proposed biosynthetic pathways to VAL-A [8] and acarbose [13, 14].

Using the *acbC* gene from the acarbose cluster as a heterologous probe, we have successfully cloned and identified the VAL-A biosynthetic genes from a genomic library of *S. hygroscopicus* 5008. Sequencing of a 6 kb BamHI fragment (Figure 2A), in vivo gene inactivation, and in vitro biochemical characterization resulted in the identification of 2-epi-5-epi-valiolone synthase (ValA), a homolog of AcbC, paving the way for further sequencing of the entire cluster [15].

Here we describe the sequencing and functional analysis of 45 kb of DNA, which led to the identification of 8 out of 27 genes as essential for validamycin biosynthesis through reconstituted heterologous gene expression in *S. lividans* 1326. We also demonstrate the glycosyltransferase function of the identified *valG* gene, first by its in vivo inactivation and functional complementation in strain 5008, second by in vitro enzymatic characterization of the ValG protein expressed in *Escherichia*

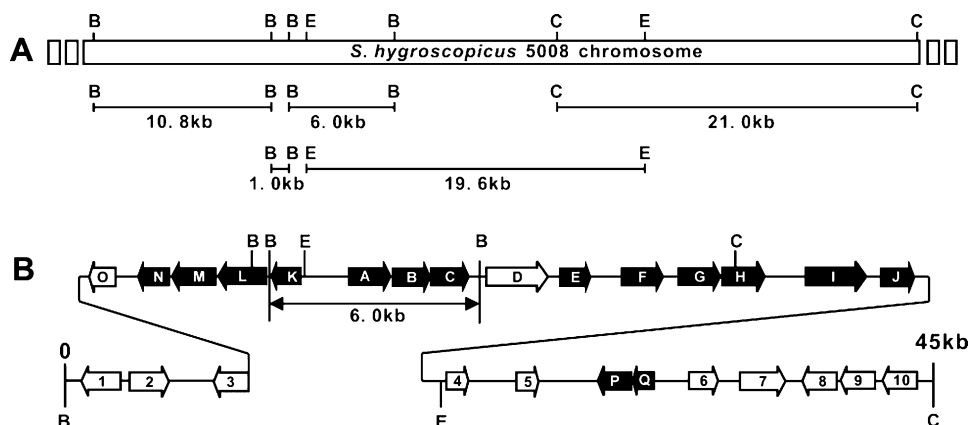


Figure 2. Cloning Strategy for Sequencing and Genetic Organization of *val* Gene Cluster

(A) Cloning strategy for sequencing of the 45 kb region. B, BamHI; C, ClaI; E, EcoRI.

(B) Genetic organization of the *val* gene cluster. Genes proposed to be involved in validamycin biosynthesis are represented as black arrows. The previously sequenced 6.0 kb BamHI fragment was indicated by double-headed line.

Table 1. Deduced Functions of the Open Reading Frames in the *val* Gene Cluster

Protein	aa	Proposed Function	Homologs, Origin	Identity, Similarity (%)	Accession No.
ORF1	400	Integrase	Urd-Int, <i>Streptomyces fradiae</i>	88, 94	AAF00216
ORF2	391	Transposase	SAV1204, <i>Streptomyces avermitilis</i>	90, 92	BAC68914
ORF3	276	Transposase	SCO7080, <i>Streptomyces coelicolor</i>	91, 93	CAB89752
ValO	221	Phosphatase/phosphohexomutase	COG0637, <i>Thermobifida fusca</i>	39, 58	XP_00293893
ValN	331	Cyclitol reductase	GlR0369, <i>Gloeobacter violaceus</i>	29, 39	BAC88310
ValM	424	Aminotransferase	DRA0029, <i>Deinococcus radiodurans</i>	39, 54	NP_285353
ValL	492	Validoxylamine A 7'-phosphate synthase	Rxy102001033, <i>Rubrobacter xylanophilus</i>	30, 47	XP_00187712
ValK	324	Epimerase/dehydratase	APE1179, <i>Aeropyrum pernix</i>	31, 51	BAA80165
ValA	414	2- <i>epi</i> -5- <i>epi</i> -valiolone synthase <sup>a</sup>	AcB <sup>a</sup> , <i>Actinoplanes</i> sp. 50/110	48, 65	CAA77208
ValB	373	Valienol-1-phosphate nucleotidyltransferase	GlgC, <i>Bacillus halodurans</i>	33, 50	BAB04806
ValC	351	Cyclitol kinase	GlcK, <i>Bacillus cereus</i>	29, 44	XP_00238533
ValD	451	Glyoxalase	NocaDRAFT_3188, <i>Nocardioides</i> sp. JS614	40, 49	XP_00657213
ValE	331	Oxidoreductase	SAV4805, <i>S. avermitilis</i>	45, 56	BAC72517
ValF	401	Oxidoreductase	SAV1825, <i>S. avermitilis</i>	80, 87	BAC69536
ValG	422	Glycosyltransferase <sup>a</sup>	KfoC <sup>a</sup> , <i>Escherichia coli</i>	35, 49	BAC00523
ValH	398	Transport protein	SAV6900, <i>S. avermitilis</i>	52, 66	BAC74611
ValI	611	Glycosyl hydrolase	SAV6503, <i>S. avermitilis</i>	69, 81	BAC74214
ValJ	305	Oxidoreductase	SAV4805, <i>S. avermitilis</i>	43, 55	BAC72517
ORF4	204	Unknown	MII4476, <i>Mesorhizobium loti</i>	39, 55	NP_105339
ORF5	347	Unknown	SMc00549, <i>Sinorhizobium meliloti</i>	28, 47	NP_385900
ValP	293	Regulatory protein	SCO3723, <i>S. coelicolor</i>	49, 63	CAD55330
ValQ	163	Regulatory protein	SCO3723, <i>S. coelicolor</i>	45, 59	CAD55330
ORF6	191	Tellurium resistance protein	SAV896, <i>S. avermitilis</i>	92, 97	BAC68606
ORF7	451	Unknown	SCO3725, <i>S. coelicolor</i>	72, 81	CAB76972
ORF8	290	Unknown	SAV7268, <i>S. avermitilis</i>	52, 68	BAC74979
ORF9	308	Esterase	Bcspa03001004, <i>Burkholderia cepacia</i>	47, 59	XP_00218391
ORF10	349	Cytochrome P450 hydroxylase	P450-SU2, <i>Streptomyces griseolus</i>	51, 61	AAA26825

<sup>a</sup> With confirmed function.

*coli*, and third by heterologous expression of the *val* gene cluster, with or without *valG*, to yield VAL-A and validoxylamine A, respectively, in *S. lividans* 1326.

## Results

### Overall Architecture of the Sequenced Region

A total of 45,168 bp of DNA sequence (deposited in GenBank under accession number [DQ164098](#)), including the previously reported 5877 bp of DNA ([AY753181](#)), was determined using overlapping or linking fragments ([Figure 2A](#)). Overall GC content of this region is 68.7%—somewhat lower than the typical GC content as exemplified by *S. coelicolor* A3(2) [16] and *S. avermitilis* [17] (~72%). In silico analysis of the sequence revealed the presence of 27 complete open reading frames (ORFs), whose organization is shown in [Figure 2B](#) and proposed functions are listed in [Table 1](#). Structural and regulatory genes likely involved in VAL-A biosynthesis, designated as *valA*-Q and indicated as black arrows in [Figure 2B](#), are distributed in the middle part of the sequenced region, among which subcluster *valA*-*valJ* and *valK*-*valN* seem to be transcribed in opposite directions.

### The *valA*-*valJ* Subcluster

*valA*, whose deduced protein sequence shows significant homology to AcbC from the acarbose biosynthetic gene cluster, was confirmed to encode ValA [15]. The deduced amino acid sequence of ValB shows 33% identity with GlgC, a glucose-1-phosphate adenyltransferase from *Bacillus halodurans* C-125 [18]. ValC shows 30% identity with AcbM, the 2-*epi*-5-*epi*-valiolone

7-kinase from the acarbose cluster, whose demonstrated function is to phosphorylate 2-*epi*-5-*epi*-valiolone to form 2-*epi*-5-*epi*-valiolone 7-phosphate [19]. Additionally, ValC shows a weaker homology to glucokinase from a *Bacillus* species (29% identity/44% similarity).

Full-length ValD (451 aa) is homologous to a putative glyoxalase/bleomycin resistance protein/dioxygenase domain from *Nocardioides* sp. JS614. Two domains of ValD, one at Gly35-Glu194 and the other at Gly220-Ala382, showed 35.5% identity/50% similarity to each other.

Both ValE (331 aa) and ValJ (305 aa) show homology to 2-oxoglutarate and Fe(II)-dependent oxygenases, whose enzymatic functions are known to catalyze a variety of reactions typically involving the oxidation of an organic substrate with an oxygen molecule, such as hydroxylation, desaturation, or oxidative ring closure [20]. Noticeably, ValE and ValJ showed significant homology (79.8% identity at nucleotide [nt] level and 66.3% identity at the protein level), and mainly differed from each other at the C terminus.

ValF is highly homologous with putative oxidoreductases from *S. avermitilis* and *S. coelicolor* (76% identity/82% similarity). In addition, a ValF-related glucose-fructose oxidoreductase from *Zymomonas mobilis* is known to catalyze the reductive conversion of fructose to sorbitol [21], suggesting that ValF might be involved in the keto reduction of intermediate compounds in VAL-A biosynthesis.

ValG exhibits considerable homology to chondroitin polymerase KfoC from *E. coli* K4, which is known to

catalyze the polymerization of the chondroitin backbone in the assembly of capsule polysaccharide K4 antigen [22]. However, ValG (422 aa) is 264 aa shorter than KfoC (686 aa), but 250 aa residues of its N-terminal domain, which is required for sugar donor binding, are very similar to known glycosyltransferases. In addition, one conserved short motif of DGS and two DXD motifs, which have been proposed to facilitate NDP-sugar binding and  $Mn^{2+}$  ion binding [23], respectively, are present in ValG.

The deduced product of *valH* is homologous to putative transport protein SAV6900 from *S. avermitilis* and the fucose permease FucP from *Mesorhizobium* sp. BNC1 (33% identity and 50% similarity), whose functions are known to be cotransport of fucose, glucose, or galactose with  $H^+$  [24]. Analysis of ValH by using the TMHMM Server v. 2.0 [25] suggests that 11 transmembrane helices span the length of the protein.

ValI seems to be closely related to the putative glucoamylase SAV6503 from *S. avermitilis*, which belongs to the glycosyl hydrolase family 15 [26]. Glucoamylases are known to catalyze the hydrolysis of  $\alpha$ -1,4 and  $\alpha$ -1,6 glucosidic linkages to release  $\beta$ -D-glucose from the nonreducing ends of starch and related poly- and oligosaccharides [27]. However, no obvious function could be assigned in the pathway for VAL-A biosynthesis, except that D-glucose could be a precursor for VAL-A.

#### The *valK-valN* Subcluster

The *valK-valN* genes are predicted to be transcribed in the opposite direction to the *valA-valJ* genes. BLAST analysis of ValK showed significant homology with dTDP-4-dehydrorhamnose reductase (RmlD) from *Aeropyrum pernix*, a member of the reductase/epimerase/dehydrogenase/dehydratase protein superfamily putatively involved in the biosynthesis of L-rhamnose from glucose-1-phosphate [28]. ValK contains a strictly conserved (P-X<sub>3</sub>-Y-X<sub>3</sub>-K-X<sub>3</sub>-E) motif that is characteristic for this superfamily and another conserved (STDNVF) motif that is unique for RmlD-type proteins [29]. Interestingly, the conserved NAD(P) binding Wierenga motif (G-X<sub>2</sub>-G-X<sub>2</sub>-G) [30] is missing from the N terminus of ValK. NAD(P)<sup>+</sup>-independent sugar epimerases from *E. coli* and other microorganisms have been reported previously [29, 31]. The reaction mechanism involves substrate to metal binding and a *cis* enediolate-stabilized intermediate [32, 33]. However, no significant sequence similarity was found between ValK and the NAD(P)<sup>+</sup>-independent sugar epimerases.

ValL shows significant homology to the putative trehalose 6-phosphate synthase Rxyl021033 from *Rubrobacter xylanophilus* and lower homology to its counterpart OtsA in *E. coli*. In *E. coli*, the trehalose 6-phosphate synthase (OtsA) and the trehalose 6-phosphatase (OtsB) are responsible for the biosynthesis of trehalose from glucose-6-phosphate and UDP-glucose, with net retention of the anomeric configuration of the donor sugar [28].

ValM is closely related to pyridoxal 5'-phosphate-dependent (PLP) aminotransferases, such as the putative 4-aminobutyrate aminotransferase from *Deinococcus radiodurans* and the acetylornithine aminotransferase from *Methanosarcina mazei*. Consistent with other aminotransferases, the conserved Lys268 of ValM is be-

lieved to be the PLP binding residue for the Schiff base linkage. In addition, other conserved residues important for PLP binding (Asp239, Gln242, and Thr322) and catalysis (Glu206 and Arg349) are also present [34].

ValN shows significant homology to the zinc-dependent threonine dehydrogenase, Krad06004313, from *Kineococcus radiotolerans* and AcbL from the acarbose pathway of *Actinoplanes* sp. SE50/110 (29% identity/39% similarity). AcbL, a putative cyclitol dehydrogenase was proposed to be involved in the reduction of the C-1 keto group of valienone 7-phosphate to give 1-*epi*-valienol 7-phosphate [19].

#### Putative Regulatory Genes

*ValP* and *valQ* are predicted to encode a two-component regulatory system consisting of a separate histidine kinase (sensor) and a Sigma B PP2C-like phosphatase (response regulator) [35], as ValP showed the highest homology with the response regulator phosphatase domain and ValQ with the sensor kinase domain of SCO3723 from *S. coelicolor* A3(2). The existence of *valP* and *valQ* as two independent genes separated by only 28 bp seems unique in the *val* gene cluster, since most of the counterpart proteins, exemplified by SCO3723, are known to exist as single peptide with two domains conforming to ValP and ValQ, respectively. It is not known yet whether the independent organization of ValP and ValQ could have anything to do with the induction of Sigma B activity and subsequent validamycin production in response to environmental stress, such as increased temperature and/or high salt concentration. However, we have consistently observed an apparent increase in production of VAL-A and enhanced transcriptions of both *valP* and *valQ* in parallel when the producer was grown at 37°C rather than at 30°C (Lei Li and Linquan Bai, unpublished data).

#### Boundaries of the Cluster and Other Most Likely Unrelated Genes

Beyond *valN*, three transposase-like or integrase-like ORFs (ORF1–3) and *valO*, which exhibits homology to a predicted phosphatase/phosphohexomutase from *Thermobifida fusca*, were detected. To determine the left boundary of the biosynthetic gene cluster, *valO*, which is located 764 bp downstream of *valN* and most likely monocistronically transcribed, was inactivated. Through double crossover recombination mediated by the 2.5 kb left flanking sequence and 4.0 kb right flanking sequence, an internal 588 bp fragment was replaced by *aac(3)IV* via ReDirect methodology [36]. The mutant ZYR-5, confirmed by PCR amplification, was found to produce a comparable amount of VAL-A as the wild-type, suggesting that *valO* is not involved in the biosynthesis of VAL-A. Thus, the left-border of the *val* gene cluster was defined.

Of the seven ORFs flanking the putative two-component regulatory genes, *valP* and *valQ*, ORF 4, 5, 7, and 8 seem to encode proteins with unknown biological functions. Moreover, none of the detected homologies of ORF 6, 9, and 10, with a tellurium resistance protein, a group of esterases, and a cytochrome P450, seem to be relevant to validamycin biosynthesis. It is thus likely that all of the genes required for VAL-A biosynthesis



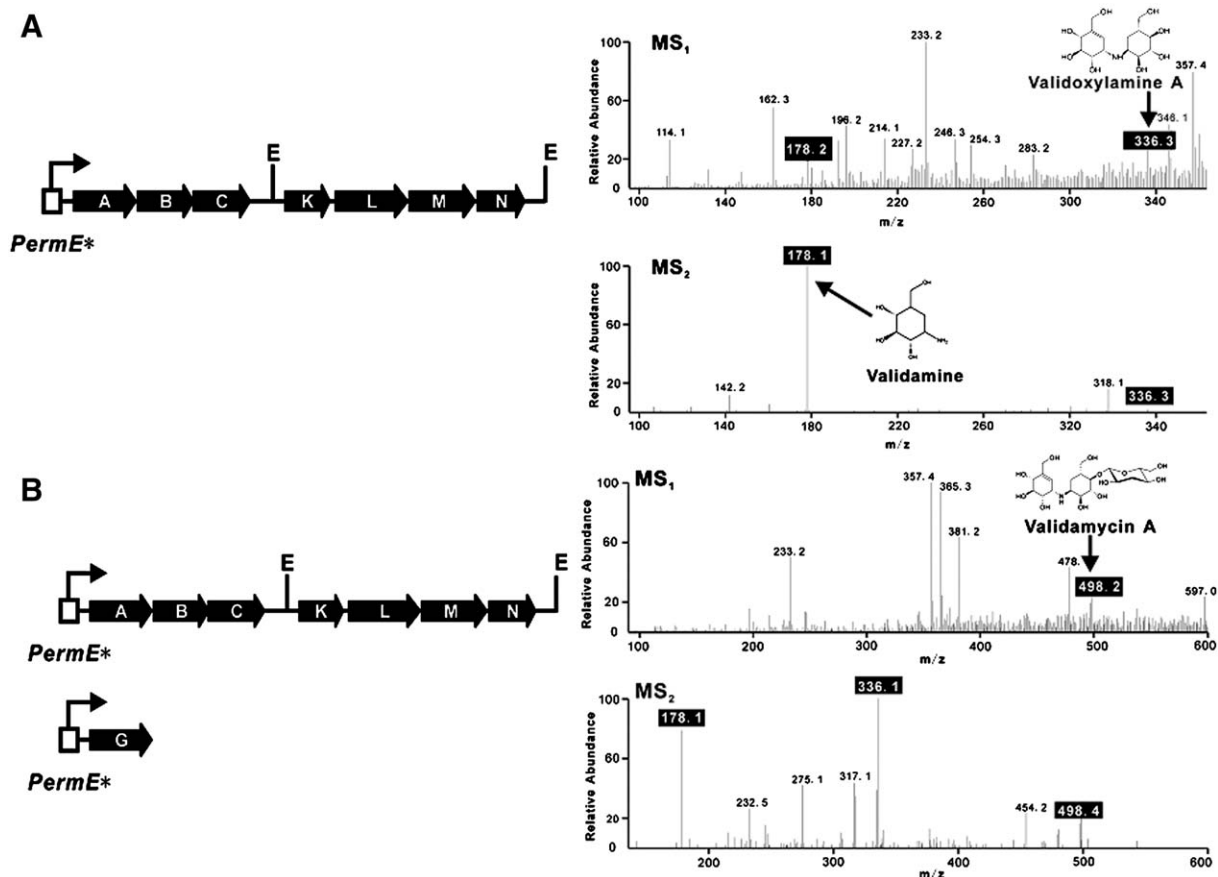


Figure 3. Heterologous Production of Validoxylamine A and VAL-A

(A) Reassembly of *valABCKLMN* and MS analysis (MS<sub>1</sub>). *PerME\**, the upmutated promoter of the erythromycin resistance gene; E, EcoRI. Fragmentation of validoxylamine A (m/z 336.1) yields validamine (m/z 178.1) (MS<sub>2</sub>). (B) Coexpression of *valABCKLMN* and *valG* and MS detection (MS<sub>1</sub>). Fragmentation of VAL-A (m/z 498.2) generates validoxylamine A (m/z 336.1) and validamine (m/z 178.1) (MS<sub>2</sub>).

are located between *valN* and *valQ* in a ~30 kb region (Figure 2B).

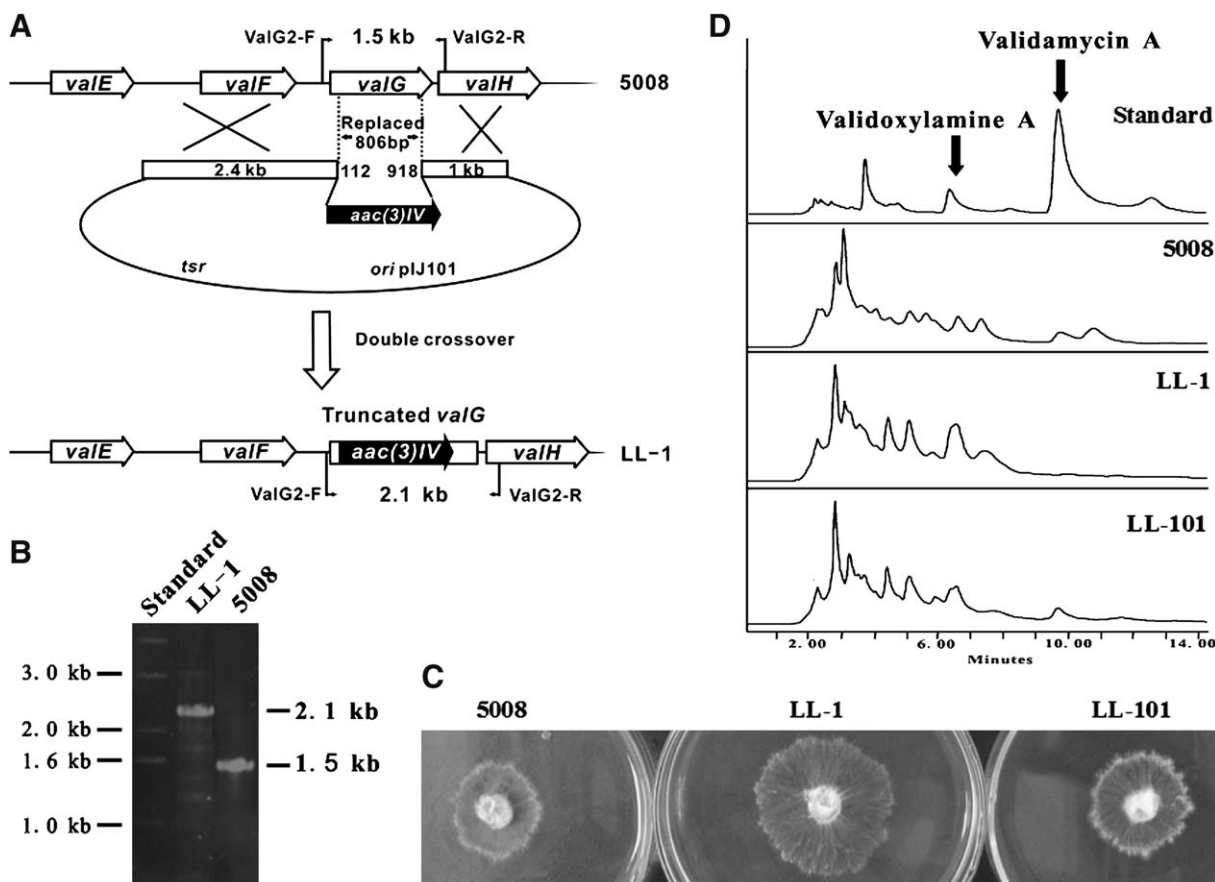
#### Heterologous Production of Validoxylamine A and VAL-A in *S. lividans* 1326

*valABC* seems to transcribe as one operon and *valKLMN* as another operon transcribed in the opposite direction (Figure 2B). The two operons were reassembled to have seven genes in the same orientation under the control of a constitutive *PerME\** promoter (Figure 3A), and a construct (pJTU757) was introduced into a heterologous host, *S. lividans* 1326, to see if the biosynthetic intermediate, validoxylamine A, can be produced. Since the 251 bp fragment to the N terminus of *valK* may still contain its own promoter, the recombinant, XH-6, was thus initially fermented at 37°C instead of at 30°C, as commonly set for *S. lividans*.

Even though bioassay with chloroform-extracted fermentation broth of XH-6 showed almost no inhibitory activity against *P. sasakii*, and HPLC analysis gave no obvious peak corresponding to validoxylamine A because of the lower yield, mass spectrometry analysis through direct injection of the sample did reveal the presence of validoxylamine A (Figure 3A). The typical fragmentation of validoxylamine A (m/z 336.1) under

positive-ion mode gives validamine (m/z 178.1). The isolated compound with m/z of 336.3 (MS<sub>1</sub>) from XH-6 was indeed fragmented to give a peak with m/z of 178.1 (Figure 3A, MS<sub>2</sub>). These data demonstrate that only seven genes, *valABCKLMN*, from the two original operons are required for the synthesis of validoxylamine A. Fermentation of XH-6 at 30°C, followed by MS analysis, also revealed the production of validoxylamine A at low yield (data not shown).

Since *valG* is the only gene found in the sequenced region encoding a glycosyltransferase, it was tempting to coexpress this gene with *valABCKLMN* for the possible production of VAL-A in *S. lividans*. The new recombinant XH-9 harboring *valABCKLMNG* showed inhibitory activity in bioassay after fermentation at 37°C for 7 days, or at 30°C for 11 days (data not shown), suggesting the accumulation of VAL-A. The typical fragmentation pattern of VAL-A (m/z 498.2) under positive-ion mode is the loss of glucose, to give validoxylamine A (m/z 336.1), which is further fragmented to validamine (m/z 178.1). Detailed analysis by mass spectrometry of a 7-day fermented broth of XH-9 at 37°C clearly indicated the presence of VAL-A with m/z of 498.2, which was sequentially fragmented into validoxylamine A (m/z 336.1) and validamine (m/z 178.1) in the MS/MS experiment (Figure 3B).



**Figure 4. Inactivation and Complementation of Glycosyltransferase *valG***

(A) Schematic representation of the replacement of an 806 bp internal fragment of *valG* with the 1.4 kb *aac(3)/IV*. In shuttle plasmid pJTU609, *aac(3)/IV* was inserted between the 2.4 kb and 1.0 kb genomic fragments originally flanking the deleted 806 bp region. While wild-type *S. hygroscopicus* should give a 2.1 kb PCR-amplified product, mutant LL-1 should yield a 1.5 kb product by using a pair of primers, ValG2-F and ValG2-R.

(B) PCR analysis of wild-type *S. hygroscopicus* and mutant LL-1.

(C) Bioassay comparison between the wild-type (left), LL-1 (middle), and LL-101 (right). LL-101 is the derivative of LL-1 harboring shuttle plasmid pJTU612 with *valG*.

(D) HPLC profiles of the standards, wild-type, LL-1, and LL-101. The retention time of VAL-A is 9.7 min, and that of validoxylamine A is 6.5 min.

The 11 day fermentation broth of XH-9 at 30°C was also analyzed by MS/MS, and an obviously similar fragmentation pattern for VAL-A could be observed (data not shown).

The successful heterologous production of validoxylamine A and VAL-A in *S. lividans* 1326 not only confirmed the necessity for VAL-A biosynthesis of the 8 structural genes from the 27 ORFs found in the sequenced region, but also revealed the identity of *valG* as a glycosyltransferase that transfers activated glucose to validoxylamine A as the final step in VAL-A biosynthesis.

## Targeted Replacement of *valG* Abolishes VAL-A Biosynthesis

LL-1 (Figure 4B) was one of the confirmed mutants, with an internal 806 bp region of *valG* (from nt number 112 to 918) replaced by an *aac(3)IV* from the strain 5008 (Figure 4A). Bioassay indicated that the LL-1 fermentation broth had retained inhibitory activity, albeit at a reduced capacity (Figure 4C), but the presence of VAL-A (retention time of 9.7 min) could not be detected by HPLC analysis. Despite the fact that, under normal conditions,

the wild-type strains also produced validoxylamine A (retention time of 6.5 min), a significantly increased accumulation of validoxylamine A was observed in LL-1 (Figure 4D). The abolished production of VAL-A and the increased accumulation of validoxylamine A in LL-1 strongly suggest that ValG catalyzes the glycosylation of validoxylamine A to yield VAL-A. The observed weak inhibitory activity agrees with the previous finding that validoxylamine A has much lower *in vivo* inhibitory activity than VAL-A [4].

### Complementation by Cloned *valG* Restores VAL-A Production

When pJTU612, a pIJ101-derived plasmid with an intact *valG*, was introduced into mutant LL-1, the culture broth of the thiostrepton-resistant exconjugant (LL-101) was found to have regained the inhibitory activity (Figure 4C) in bioassay. HPLC analysis unambiguously demonstrated the presence of VAL-A with the retention time of 9.7 min. However, the presence of *valG* under the strong constitutively expressed *PerME*<sup>\*</sup> promoter in LL-101 was not sufficient for the conversion of all

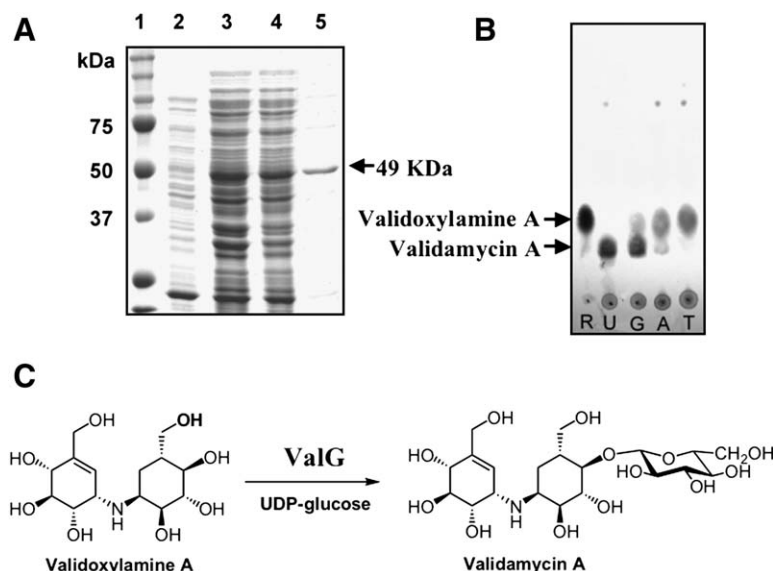


Figure 5. Heterologous Expression of *valG* and Characterization of the Glycosyltransferase

(A) SDS-PAGE analysis of ValG protein: (1) molecular weight marker; (2) soluble protein of extract of *E. coli* BL21Gold(DE3)pLysS/*valG* before induction; (3) total protein of extract of *E. coli* BL21Gold(DE3)pLysS/*valG* after induction with IPTG; (4) soluble protein of line 3; (5) purified his-tagged ValG. (B) TLC analysis of ValG reaction (silica gel, solvent system: nPrOH:AcOH:H<sub>2</sub>O = 4:1:1); R, validoxylamine A standard; U, ValG reaction by using UDP-glucose as glucosyl donor, G, ValG reaction by using GDP-glucose as glucosyl donor; A, ValG reaction by using ADP-glucose as glucosyl donor; T, ValG reaction using TDP-glucose as glucosyl donor. (C) Conversion scheme of validoxylamine A to VAL-A catalyzed by ValG.

validoxylamine A to VAL-A, as a peak with the retention time of 6.5 min corresponding to validoxylamine A was also observed (Figure 4D).

#### In Vitro Glycosylation of Validoxylamine A to VAL-A Using Recombinant ValG

Further proof to support the glycosyltransferase function of *valG* was obtained from its heterologous expression in *E. coli*, which gave rise to a 49 kDa soluble polyhistidine-tagged protein. Affinity purification on an Ni-NTA spin column (QIAGEN) or a BD TALON column gave a protein that was >75% pure as judged by SDS-polyacrylamide gel electrophoresis (Figure 5A). To test the catalytic activity of ValG, the enzymatic reaction was carried out using validoxylamine A and UDP-, GDP-, ADP-, or TDP-glucose as substrates. The reactions gave rise to a product that has the same Rf value as authentic VAL-A on TLC (Figure 5B). On the other hand, incubation with cell-free extracts of *E. coli* harboring empty pRSET-B vector gave no product (data not shown). The product was confirmed to be VAL-A by ESI-MS and NMR. UDP-glucose was the most efficient glucosyl donor for the ValG reaction, as judged by TLC (Figure 5B), whereas GDP-glucose was less efficient and ADP-glucose was much less efficient. To our surprise, TDP-glucose was the least efficient donor for the reaction among the activated sugars tested.

#### Discussion

Earlier feeding experiments with isotopically labeled potential intermediates led to a proposal that the biosynthesis of VAL-A is initiated by the cyclization of D-sedoheptulose 7-phosphate to form 2-*epi*-5-*epi*-valiolone, which subsequently epimerizes at C-2 to give 5-*epi*-valiolone, followed by dehydration between C-5 and C-6 to generate valienone. The conversion of valienone to validone takes place by *anti*-addition of hydrogen followed by transamination, with the  $\alpha$ -nitrogen of glutamate as the probable nitrogen source, to form validamine. Analysis of an ~45 kb region of the *S. hygroscopicus* 5008 chromosome revealed 16 potential struc-

tural genes, which include genes deduced or proven to encode ValA, epimerase/dehydratase (ValK), reductase (ValN), transaminase (ValM), and glycosyltransferase (ValG) (Table1). These enzymes fit nicely with the proposed scheme of VAL-A biosynthesis resulting from the feeding experiments. Despite some discrepancies between the biosynthesis of acarbose and VAL-A, cloning of the acarbose biosynthetic gene cluster and subsequent biochemical characterization of several enzymes involved in the biosynthesis has provided additional information for the elucidation of VAL-A biosynthesis [13]. In the case of acarbose, 2-*epi*-5-*epi*-valiolone, the cyclization product of AcbC is phosphorylated by the kinase AcbM to give 2-*epi*-5-*epi*-valiolone 7-phosphate. Further conversion to 5-*epi*-valiolone 7-phosphate is catalyzed by the epimerase AcbO. The presence of ValC, a homolog of AcbM, in the *val* gene cluster suggests the involvement of phosphorylated intermediates in VAL-A biosynthesis as well. However, since 2-*epi*-5-*epi*-valiolone, 5-*epi*-valiolone, valienone, and validone were efficiently incorporated into VAL-A, the kinase (ValC) may utilize valienone and/or validone as substrate. Alternatively, the kinase might have broad substrate specificity, activating 2-*epi*-5-*epi*-valiolone, 5-*epi*-valiolone, valienone, and validone, and making it possible for them to be incorporated into the biosynthetic pathway.

In analogy to the proposed function of the dehydrogenase AcbL in acarbose biosynthesis, ValN is predicted to reduce the C-1 keto group of valienone-7-phosphate to give either valienol 7-phosphate or its epimer, 1-*epi*-valienol 7-phosphate (Figure 6). If valienol 7-phosphate is formed, a downstream coupling reaction, most likely catalyzed by the trehalose 6-phosphate synthase homolog (ValL), would occur with net retention of the C-1 configuration of the unsaturated cyclitol. However, if 1-*epi*-valienol 7-phosphate is formed, the coupling reaction must take place with inversion of the C-1 configuration. In acarbose biosynthesis, it was proposed that reduction of valienone 7-phosphate yields 1-*epi*-valienol 7-phosphate, and further phosphorylation of the latter compound by an unidentified kinase gives 1-*epi*-valienol

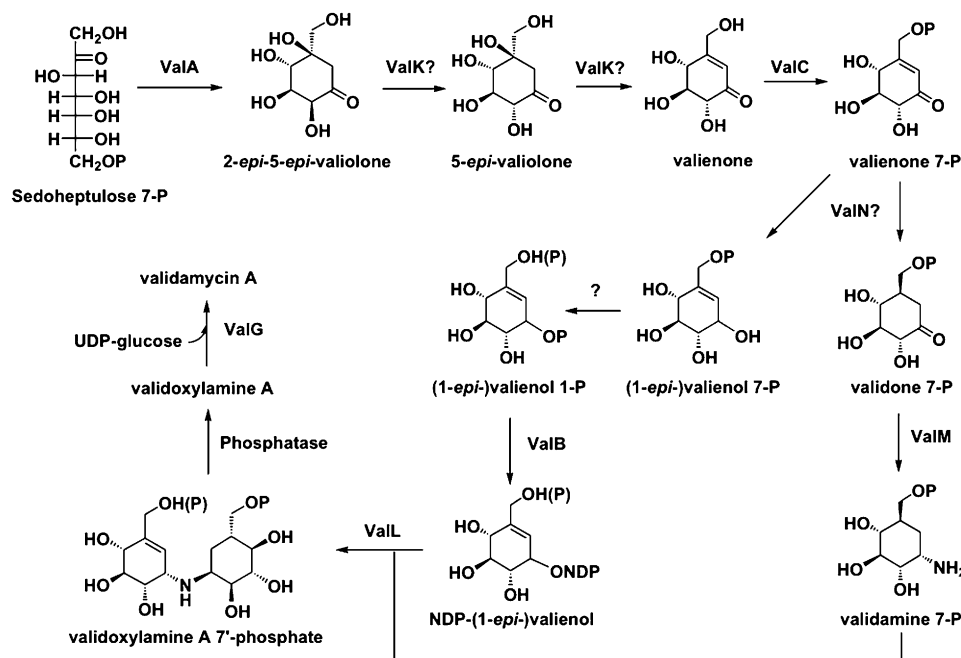


Figure 6. Proposed Biosynthetic Pathway to VAL-A

1,7-diphosphate. However, while this hypothesis is attractive, there is no evidence for the presence of a second kinase in the validamycin cluster. The presence of *valO*, which encodes a protein homologous to phosphatase/phosphohexomutase, in the validamycin cluster conversely suggests the involvement of a phosphomutation step from the C-7 to the C-1 hydroxyl group, as commonly observed in sugar biosyntheses. However, inactivation of *valO* did not abolish the production of VAL-A, which leaves this part of the pathway unclear.

Nevertheless, the availability of the genetic information reported here, as well as results from our previous feeding experiments and biochemical analysis of several related enzymes, allows us to propose a more comprehensive pathway to VAL-A (Figure 6). In vivo inactivation and in vitro characterization of ValA confirmed the initiation step from D-sedoheptulose 7-phosphate to the cyclic product, 2-epi-5-epi-valienone [15]. Subsequently, epimerization at C-2 and dehydration at C-5/C-6 of 2-epi-5-epi-valienone by the proposed bifunctional ValK would give valienone. ValK has significant homology with the sugar epimerase/dehydratase, and it appears to be the only plausible candidate enzyme for the epimerization and dehydration reactions among the eight proteins known to be involved in VAL-A biosynthesis. In addition, ValK may not require NAD(P)<sup>+</sup> for catalytic activity, which is similar to the L-ribulose-5-phosphate 4-epimerase and the dTDP-4-dehydrorhamnose 3,5-epimerase [29, 31]. However, no significant homology was found between the primary structures of ValK and the NAD<sup>+</sup>-independent sugar epimerases. This raises questions regarding the function and reaction mechanism of ValK, which prompts further investigations.

Prior to any further processing, valienone may be phosphorylated, presumably by the kinase, ValC, to give valienone 7-phosphate. Valienone 7-phosphate may serve as the branching point to be reduced to val-

idone 7-phosphate on the one hand, and to valienol 7-phosphate or 1-*epi*-valienol 7-phosphate on the other, by the reductase ValN and/or other reductases/dehydrogenases. There are three oxidoreductase genes present in the cluster: *valE*, *valF*, and *valJ*. However, the results of the heterologous expression experiments in *S. lividans* suggest that these genes are not necessary for the biosynthesis of VAL-A. Even so, it may be that they are involved in the biosynthesis of validamycins B and/or G, the hydroxylated analogs of VAL-A. Transamination of the keto group of validone 7-phosphate catalyzed by the aminotransferase ValM would give validamine 7-phosphate. On the other hand, valienol 7-phosphate or its 1-epimer is either phosphorylated at C-1 by an unidentified kinase to generate the 1,7-diphosphate derivative or transformed into valienol 1-phosphate through catalysis by a phosphoglucomutase. The product is subsequently converted to an NDP-valienol derivative by the nucleotidyl transferase ValB. Condensation of validamine 7-phosphate with the NDP-valienol derivative would give phosphorylated validoxylamine A, whose conversion to validoxylamine A requires a dephosphorylation activity. The corresponding phosphatase has not been located in the sequenced region. However, an internal trehalose 7-phosphate phosphatase or a nonspecific sugar phosphatase may be able to carry out this hydrolytic reaction. Finally, validoxylamine A is converted to VAL-A biosynthesis through catalysis by the glycosyltransferase ValG by using UDP-glucose as the sugar donor.

## Significance

VAL-A is a strong crop protectant produced by *Streptomyces hygroscopicus* and is widely used for the treatment of the sheath blight disease of rice plants and the vegetable dumping-off caused by the fungus



*Pellicularia sasakii*. In addition, its degradation product, valienamine, is a pharmaceutically important precursor for the production of the antidiabetic agent, voglibose. Despite its significant and wide applications in the biomedical and agricultural sectors, little is known about its biosynthesis. Sequencing and bioinformatics analysis of the validamycin biosynthetic gene cluster, as well as in vivo and in vitro characterization of a number of enzymes involved in VAL-A biosynthesis, provide important insights into the biosynthesis of this antibiotic. The data presented will also aid in the elucidation of the biosynthesis of other related aminocyclitol-containing natural products. In combination with knowledge about the acarbose biosynthetic gene cluster, it may now be possible to design unique primers for the screening and cloning of other aminocyclitol biosynthetic gene clusters from various sources. Exemplified by the significant accumulation of validoxylamine A in the glycosyltransferase mutants, in-depth knowledge about VAL-A biosynthesis could potentially be used to establish a platform for the production of drug precursors (e.g., valienamine, validamine, valioline, and hydroxyvalidamine). Functional analysis of validamycin biosynthesis also sets the stage for creating novel analogs of VAL-A via combinatorial biosynthesis and yield improvement through metabolic engineering. Finally, with all the genes necessary for the synthesis of this important crop protectant at hand, one can contemplate the possibility of expressing them in plant systems to endow crop plants with their built-in protection against pathogens such as *P. sasakii*.

#### Experimental Procedures

##### Bacterial Strains, Plasmids, Culture Techniques, and Media

*S. hygroscopicus* 5008 [37] and its derivatives were grown either on solid SFM medium or in modified liquid TSB medium [38] (supplemented with 1% yeast extract and 10.3% sucrose). Fermentation was carried out in FM-II medium at 37°C for 6 days [15]. DH10B (GIBCO-BRL), ET12567(pUZ8002) [39], and BL21Gold(DE3)pLysS [40] were used as *E. coli* hosts for plasmid construction, *E. coli-Streptomyces* conjugation, and protein overexpression, respectively. *P. sasakii* was used as indicator strain for VAL-A bioassay. pHZ1358 was the cosmid vector used for constructing the 5008 genomic library and for gene inactivation [41]. pJTU472 contains *aac(3)/IV* with HindIII sites on both sides, which was originally cloned from pHZ1070 [42]. pJ2925 [43] was *E. coli* vector for plasmid construction and pUC18 [44] was *E. coli* vector for DNA sequencing. pRSET-B (Invitrogen) was used as vector for protein overexpression in *E. coli*.

Shuttle vector pJTU695 for our complementation study was constructed as follows. The 1628 bp region between SphI and NruI of the pET15b (Novagen) was replaced first by an 868 bp SphI-BalI fragment containing *oriT* from pSET152 [45]. A 163 bp NdeI-BglII region of the resulting plasmid was further replaced by a 4365 bp fragment containing the replication origin of pJ101 and *tsr*, creating pJTU676. A region of pJTU676, with the *Perme*<sup>+</sup> promoter, ribosomal binding site (RBS), and NdeI site, was replaced with a corresponding fragment from pB139 [46] after digestion with EcoRI and HindIII. The distance between the RBS and the start codon of the downstream gene is 7 bp in the final construct, pJTU695, while that in pJTU676 is 23 bp.

##### DNA Sequencing and Analysis

DNA sequencing was performed using a set of five fragments directly isolated from cosmids 4G8 and 17F2 [15], or cloned onto pBluescript SK(+) (Stratagene) or pSL301 (Invitrogen), which were originally connected or overlapping to cover an ~45 kb region with

the validamycin biosynthetic gene cluster (Figure 2A). pHZ2232 with the 1.0 kb BamHI fragment was directly sequenced with T3 and T7 primers. Sequencing analysis of the 21 kb ClaI fragment, a 19.6 kb EcoRI fragment (pHZ2239), a 6.0 kb BamHI fragment (pHZ2229), and 7.0 and 3.8 kb BamHI fragments (pJTU464) were as described in [47].

##### Inactivation of *valG*

A 2.4 kb EcoRI-HindIII fragment flanking the left of the nt 112 of *valG* and a 1.0 kb HindIII-EcoRI fragment flanking the right of the nt 918 of *valG* were coligated into EcoRI-digested pJ2925 to generate pJTU600. A HindIII fragment carrying the 1.4 kb *aac(3)/IV* cassette from pJTU472 was inserted into the HindIII site between the 2.4 and 1.0 kb fragments to generate pJTU607, from which a 4.8 kb BglII fragment was excised for insertion into the BamHI-digested pHZ1358 to construct a vector (pJTU609). pJTU609 was then used for mediating gene replacement (Figure 4) to obtain a *valG* mutant (LL-1). The mutation was confirmed by PCR amplification using forward primer *valG*2-F (5'-AGAGCCGATCTGGTGGTGGAG-3') and reverse primer *valG*2-R (5'-GGTGATGATTAGCCCTTCTCG-3').

##### Complementation of Mutant LL-1 with Cloned *valG*

PCR-amplified *valG* was cleaved as a 1.2 kb NdeI-EcoRI fragment and cloned into pJTU695 digested with the same restriction enzymes. The new plasmid, pJTU612, was introduced from *E. coli* into LL-1 through conjugation as previously described [15], and exconjugants were selected with 25 µg/ml thiostrepton. Further confirmation was carried out through plasmid isolation from 2-day-old mycelia, transformation into *E. coli*, and comparison between the newly purified pJTU612 and the original one by restriction enzymes digestions. An exconjugant having the correct plasmid, named LL-101, was investigated through fermentation, bioassay, and HPLC analysis, as previously described [15].

##### Targeted Disruption of *valO*

The *valO* gene was disrupted with ReDirect Technology. The *aac(3)/IV* and *oriT* cassette was amplified from the pJ1773 disruption cassette [36] with the primers *ValO*-PCR-F (5'-ATGTATAAGGTCGCACTTTTCGATCTGGACGGCACGTTAattccggggatccgtcgacc-3') and *ValO*-PCR-R (5'-TCACGTGACGTCGTGATTGCAATCGGAGGGAGG AACGATgttaggctggagctgcttc-3') (pJ1773 homologous sequence is in lowercase letters). The resulting PCR product was used to replace *valO* first in pJTU714, a pHZ1358-derived plasmid carrying the 7.0 kb BamHI fragment containing *valO*, and then in strain 5008, as previously described [15]. Allelic replacement of *valO* in the ZYR-5 mutant was confirmed by PCR amplification with the primers of *ValO*-Det-F (5'-CGACGACCCCTCAACCTC-3') and *ValO*-Det-R (5'-CTGGCATCAAGCGACAC-3'). Fermentation of ZYR-5 and activity comparison between the wild-type and ZYR-5 through bioassay and HPLC were performed.

##### Cloning and Heterologous Overexpression of Recombinant His<sub>6</sub>-Tagged *ValG*

The *valG* gene was amplified by PCR with Platinum Pfx DNA polymerase (Invitrogen) with the cosmid 3G8 as template and primers *ValG*F3 (5'-GAAGATCTGCATATGCCCGGTGCGACATCCCATG-3'; engineered BglII and NdeI sites are in italics) and *ValG*R2 (5'-GGAATTCTCAGTACCGCGAAGAGACGGCTCG-3'; engineered EcoRI site is in italics). The PCR amplification was done in a thermocycler (Eppendorf, Mastercycler gradient) under the following conditions: 33 cycles of 90 s at 95°C, 45 s at 60°C, and 45 s at 72°C. The PCR products were digested with BglII and EcoRI, and subsequently ligated into BamHI/EcoRI-digested pRSET-B. The constructs were transformed into *E. coli* XL-1 Blue and plated on LB agar plates containing 100 µg/ml ampicillin. The plasmid DNA was isolated and introduced by heat-pulse transformation into *E. coli* BL21 Gold(DE3)pLysS (Stratagene), which was then plated onto LB agar plates containing 100 µg/ml ampicillin and 25 µg/ml chloramphenicol. The transformants were grown in 20 ml LB medium containing ampicillin and chloramphenicol at 37°C to an OD<sub>600</sub> of 0.6. IPTG was added to a final concentration of 0.2 mM, and the incubation was continued at 28°C for 24 hr. The cells were harvested by centrifugation at 3500 rpm for 15 min and stored frozen at -80°C until further use.

### Preparation of Cell-free Extracts and Purification of His<sub>6</sub>-Tagged ValG

Cells were thawed and resuspended in disruption buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole [pH 8.0]). The suspension was sonicated three times for 25 s each, and cell debris was removed by centrifugation at 10,000 rpm for 10 min. The protein solution was applied to a BD TALON spin column (BD Biosciences) and centrifuged at 500 rpm for 2 min. The column was washed with washing buffer (twice with 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl [pH 8.0], and once with 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 100 mM imidazole [pH 8.0]). The His<sub>6</sub>-tagged protein was eluted with elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 500 mM imidazole [pH 8.0]) and dialyzed for 24 hr against 1 liter of dialysis buffer (25 mM Tris-HCl [pH 7.6], 10 mM MgCl<sub>2</sub>, 20 mM NH<sub>4</sub>Cl, and 0.5 mM DTT). Protein concentration was measured by the Bradford protein microassay with bovine serum albumin as standard.

### Enzyme Assay

The enzyme assay was carried out at 30°C for 3–12 hr in a 100 µl volume of 25 mM Tris-HCl (pH 7.6), 10 mM MgCl<sub>2</sub>, 20 mM NH<sub>4</sub>Cl, 15 mM NDP-glucose, 10 mM validoxylamine A, and 50 µl of protein solution (2.4 mg/ml protein). The reaction progress was monitored by TLC analysis. For a scale-up experiment, 3.35 mg of validoxylamine A was used. The reaction mixture was applied to a Microcon YM-10 (Amicon) spin column. The flow through was subjected to a cation-exchange (Dowex 50Wx8-200 [H<sup>+</sup>]) chromatography column. The column was washed with 20 ml of water, and the reaction product was eluted with 0.5 M aqueous NH<sub>4</sub>OH. Fractions containing the desired product were pooled and lyophilized to give 4.1 mg of VAL-A as a white powder.

### Heterologous Production of Validoxylamine A and VAL-A

The *valABC* operon was first cloned downstream of the *PermE*<sup>+</sup> promoter by inserting a 623 bp NdeI-FspI DNA fragment of *valA* and a 3280 bp FspI-BamHI fragment from pHZ2229 into pSL301 digested with NdeI and BamHI. Then, a 3964 bp NdeI-EcoRI fragment containing the complete *valABC* was cleaved from the resulting plasmid, pJTU755, and ligated into pIB139 digested with the corresponding enzymes, generating pJTU756. Subsequently, a 5611 bp EcoRI fragment, containing *valKLMN* from cosmid 17F2, was inserted into EcoRI-digested pJTU756. Clones with *valABCKLMN* transcribed in the same orientation were selected and named pJTU757.

Plasmid pJTU757 was integrated into the chromosome of *S. lividans* 1326 by protoplast transformation and selection with apramycin to create a recombinant XH-6. Introduction of pJTU612 carrying *valG* into XH-6 by conjugation resulted in another recombinant XH-9, which was selected with thiostrepton and apramycin. Fermentation and bioassay analysis of XH-6 and XH-9 were done as previously described [15]. Tandem MS analysis was performed with fermentation broth extracted once with chloroform through directed injection on an Agilent 1100 series LC/MSD Trap system. The iontrap mass spectrometer was operated with the electrospray ionization source in the positive-ion mode. Drying gas flow was 10 liters/min, and nebulizer pressure was 50 psi. Drying gas temperature was 325°C. The fragmentation amplitude was varied between 1.0 and 1.8 V.

### Acknowledgments

This work received support from the Ministry of Science and Technology (2003CB114205), the National Science Foundation of China, the Ph.D. Training Fund of the Ministry of Education, and the Shanghai Municipal Council of Science and Technology. Work at Oregon State University was supported by NIH grant RAI061528A. K.M. received postdoctoral training funds from Shionogi & Co., LTD, Osaka, Japan. The authors thank Dr. Patricia M. Flatt for helpful comments.

Received: November 4, 2005

Revised: January 25, 2006

Accepted: February 1, 2006

Published: April 21, 2006

### References

- Iwasa, T., Yamamoto, H., and Shibata, M. (1970). Studies on validamycins, new antibiotics. I: *Streptomyces hygroscopicus* var. *limoneus* nov. var., validamycin-producing organism. Jpn. J. Antibiot. 23, 595–602.
- Iwasa, T., Higashide, E., Yamamoto, H., and Shibata, M. (1971). Studies on validamycins, new antibiotics. II: Production and biological properties of validamycins A and B. J. Antibiot. (Tokyo) 24, 107–113.
- Kameda, Y., Asano, N., Yamaguchi, T., and Matsui, K. (1987). Validoxylamines as trehalase inhibitors. J. Antibiot. (Tokyo) 40, 563–565.
- Asano, N., Tanaka, K., Kameda, Y., and Matsui, K. (1991). All eight possible mono-beta-D-glucosides of validoxylamine A. II: Biological activities. J. Antibiot. (Tokyo) 44, 1417–1421.
- Stratmann, A., Mahmud, T., Lee, S., Distler, J., Floss, H.G., and Piepersberg, W. (1999). The AcbC protein from *Actinoplanes* species is a C7-cyclitol synthase related to 3-dehydroquinase synthases and is involved in the biosynthesis of the alpha-glucosidase inhibitor acarbose. J. Biol. Chem. 274, 10889–10896.
- Mahmud, T., Tornus, I., Egelkrout, E., Wolf, E., Uy, C., Floss, H.G., and Lee, S. (1999). Biosynthetic studies on the alpha-glucosidase inhibitor acarbose in *Actinoplanes* sp.: 2-*epi*-5-*epi*-valiolone is the direct precursor of the valienamine moiety. J. Am. Chem. Soc. 121, 6973–6983.
- Arakawa, K., Bowers, S.G., Michels, B., Trin, V., and Mahmud, T. (2003). Biosynthetic studies on the alpha-glucosidase inhibitor acarbose: the chemical synthesis of isotopically labeled 2-*epi*-5-*epi*-valiolone analogs. Carbohydr. Res. 338, 2075–2082.
- Dong, H., Mahmud, T., Tornus, I., Lee, S., and Floss, H.G. (2001). Biosynthesis of the validamycins: identification of intermediates in the biosynthesis of validamycin A by *Streptomyces hygroscopicus* var. *limoneus*. J. Am. Chem. Soc. 123, 2733–2742.
- Mahmud, T., Lee, S., and Floss, H.G. (2001). The biosynthesis of acarbose and validamycin. Chem. Rec. 1, 300–310.
- Mahmud, T., Xu, J., and Choi, Y.U. (2001). Synthesis of 5-*epi*-[6-(2)H(2)]valiolone and stereospecifically monodeuterated 5-*epi*-valiolones: exploring the steric course of 5-*epi*-valiolone dehydratase in validamycin A biosynthesis. J. Org. Chem. 66, 5066–5073.
- Naganawa, H., Hashizume, H., Kubota, Y., Sawa, R., Takahashi, Y., Arakawa, K., Bowers, S.G., and Mahmud, T. (2002). Biosynthesis of the cyclitol moiety of pyralomicin 1a in *Nonomuraea spiralis* MI178–34F18. J. Antibiot. (Tokyo) 55, 578–584.
- Kameda, Y., Horii, S., and Yamano, T. (1975). Microbial transformation of validamycins. J. Antibiot. (Tokyo) 28, 298–306.
- Wehmeier, U.F., and Piepersberg, W. (2004). Biotechnology and molecular biology of the alpha-glucosidase inhibitor acarbose. Appl. Microbiol. Biotechnol. 63, 613–625.
- Zhang, C.S., Podeschwa, M., Altenbach, H.J., Piepersberg, W., and Wehmeier, U.F. (2003). The acarbose-biosynthetic enzyme AcbO from *Actinoplanes* sp. SE 50/110 is a 2-*epi*-5-*epi*-valiolone-7-phosphate 2-*epimerase*. FEBS Lett. 540, 47–52.
- Yu, Y., Bai, L., Minagawa, K., Jian, X., Li, L., Li, J., Chen, S., Cao, E., Mahmud, T., Floss, H.G., et al. (2005). Gene Cluster Responsible for Validamycin Biosynthesis in *Streptomyces hygroscopicus* subsp. *jinggangensis* 5008. Appl. Environ. Microbiol. 71, 5066–5076.
- Bentley, S.D., Chater, K.F., Cerdano-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417, 141–147.
- Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osone, T., et al. (2001). Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. Proc. Natl. Acad. Sci. USA 98, 12215–12220.
- Henner, D.J., Band, L., and Shimotsu, H. (1985). Nucleotide sequence of the *Bacillus subtilis* tryptophan operon. Gene 34, 169–177.

19. Zhang, C.S., Stratmann, A., Block, O., Bruckner, R., Podeschwa, M., Altenbach, H.J., Wehmeier, U.F., and Piepersberg, W. (2002). Biosynthesis of the C(7)-cyclitol moiety of acarbose in *Actinoplanes* species SE50/110: 7-O-phosphorylation of the initial cyclitol precursor leads to proposal of a new biosynthetic pathway. *J. Biol. Chem.* 277, 22853–22862.
20. Prescott, A.G., and Lloyd, M.D. (2000). The iron(II) and 2-oxo-acid-dependent dioxygenases and their role in metabolism. *Nat. Prod. Rep.* 17, 367–383.
21. Kingston, R.L., Scopes, R.K., and Baker, E.N. (1996). The structure of glucose-fructose oxidoreductase from *Zymomonas mobilis*: an osmoprotective periplasmic enzyme containing non-dissociable NADP. *Structure* 4, 1413–1428.
22. Ninomiya, T., Sugiura, N., Tawada, A., Sugimoto, K., Watanabe, H., and Kimata, K. (2002). Molecular cloning and characterization of chondroitin polymerase from *Escherichia coli* strain K4. *J. Biol. Chem.* 277, 21567–21575.
23. Hu, Y., and Walker, S. (2002). Remarkable structural similarities between diverse glycosyltransferases. *Chem. Biol.* 9, 1287–1296.
24. Gunn, F.J., Tate, C.G., Sansom, C.E., and Henderson, P.J. (1995). Topological analyses of the L-fucose-H<sup>+</sup> symport protein, FucP, from *Escherichia coli*. *Mol. Microbiol.* 15, 771–783.
25. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.
26. Ullman, C.G., and Perkins, S.J. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* 326, 929–942.
27. Chiba, S. (1997). Molecular mechanism in  $\alpha$ -glucosidase and glucoamylase. *Biosci. Biotechnol. Biochem.* 61, 1233–1239.
28. Rocchetta, H.L., Burrows, L.L., and Lam, J.S. (1999). Genetics of O-antigen biosynthesis in *Pseudomonas aeruginosa*. *Microbiol. Mol. Biol. Rev.* 63, 523–553.
29. Graninger, M., Nidetzky, B., Heinrichs, D.E., Whitfield, C., and Messner, P. (1999). Characterization of dTDP-4-dehydrorhamnose 3,5-epimerase and dTDP-4-dehydrorhamnose reductase, required for dTDP-L-rhamnose biosynthesis in *Salmonella enterica* serovar Typhimurium LT2. *J. Biol. Chem.* 274, 25069–25077.
30. Wierenga, R., De Maeyer, M., and Hol, W. (1985). Interaction of pyrophosphate moieties with  $\alpha$ -helices in dinucleotide binding proteins. *Biochemistry* 24, 1346–1357.
31. Johnson, A.E., and Tanner, M.E. (1998). Epimerization via carbon-carbon bond cleavage: L-ribulose-5-phosphate 4-epimerase as a masked class II aldolase. *Biochemistry* 37, 5746–5754.
32. Samuel, J., Luo, Y., Morgan, P.M., Strynadka, N.C., and Tanner, M.E. (2001). Catalysis and binding in L-ribulose-5-phosphate 4-epimerase: a comparison with L-fuculose-1-phosphate aldolase. *Biochemistry* 40, 14772–14780.
33. Lee, L.V., Poyner, R.R., Vu, M.V., and Cleland, W.W. (2000). Role of metal ions in the reaction catalyzed by L-ribulose-5-phosphate 4-epimerase. *Biochemistry* 39, 4821–4830.
34. Storici, P., Capitani, G., De Biase, D., Moser, M., John, R.A., Jansonius, J.N., and Schirmer, T. (1999). Crystal structure of GABA-aminotransferase, a target for antiepileptic drug therapy. *Biochemistry* 38, 8628–8634.
35. Delumeau, O., Dutta, S., Brigulla, M., Kuhnke, G., Hardwick, S.W., Volker, U., Yudkin, M.D., and Lewis, R.J. (2004). Functional and structural characterization of RsbU, a stress signaling protein phosphatase 2C. *J. Biol. Chem.* 279, 40927–40937.
36. Gust, B., Challis, G.L., Fowler, K., Kieser, T., and Chater, K.F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. USA* 100, 1541–1546.
37. Xia, T.H., and Jiao, R.S. (1986). Studies on glutamine synthetase from *Streptomyces hygroscopicus* var. *jinggangensis*. *Sci. Sin. [B]* 29, 379–388.
38. Kieser, T., Bibb, M.J., Chater, K.F., Butter, M.J., and Hopwood, D.A. (2000). Practical *Streptomyces* Genetics: A Laboratory Manual (Norwich, UK: John Innes Foundation).
39. Paget, M.S., Chamberlin, L., Atri, A., Foster, S.J., and Buttner, M.J. (1999). Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* 181, 204–211.
40. Moffatt, B.A., and Studier, F.W. (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* 49, 221–227.
41. Sun, Y., Zhou, X., Liu, J., Bao, K., Zhang, G., Tu, G., Kieser, T., and Deng, Z. (2002). 'Streptomyces nanchangensis', a producer of the insecticidal polyether antibiotic nanchangmycin and the antiparasitic macrolide meilingmycin, contains multiple polyketide gene clusters. *Microbiology* 148, 361–371.
42. Chen, S., Huang, X., Zhou, X., Bai, L., He, J., Jeong, K.J., Lee, S.Y., and Deng, Z. (2003). Organizational and mutational analysis of a complete FR-008/candicidin gene cluster encoding a structurally related polyene complex. *Chem. Biol.* 10, 1065–1076.
43. Janssen, G.R., and Bibb, M.J. (1993). Derivatives of pUC18 that have BglII sites flanking a modified multiple cloning site and that retain the ability to identify recombinant clones by visual screening of *Escherichia coli* colonies. *Gene* 124, 133–134.
44. 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.
45. Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N., and Schoner, B.E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* 116, 43–49.
46. Oliynyk, M., Stark, C.B., Bhatt, A., Jones, M.A., Hughes-Thomas, Z.A., Wilkinson, C., Oliynyk, Z., Demydchuk, Y., Staunton, J., and Leadlay, P.F. (2003). Analysis of the biosynthetic gene cluster for the polyether antibiotic monensin in *Streptomyces cinnamonensis* and evidence for the role of *monB* and *monC* genes in oxidative cyclization. *Mol. Microbiol.* 49, 1179–1190.
47. Ewing, B., Hillier, L., Wendl, M.C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I: Accuracy assessment. *Genome Res.* 8, 175–185.

#### Accession Numbers

The complete DNA and deduced protein sequences of the *val* gene cluster reported in this paper have been deposited in GenBank under the accession number [DQ164098](#). Previously reported 5877 bp of DNA within the 45 kb sequenced region was deposited in GenBank under accession number [AY753181](#).