INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Chlorination Genes of Chloramphenicol Biosynthesis in *Streptomyces venezuelae*

by

Mahmood Piraee

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
February 2002

© Copyright by Mahmood Piraee, 2002

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-67651-X

Canada

DALHOUSIE UNIVERSITY

FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "Chlorination Genes of Chloramphenicol Biosynthesis in *Streptomyces venezuelae*", by Mahmood Piraee, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: February 28, 2002

External Examiner:

Research Supervisor:

Examining Committee:

DALHOUSIE UNIVERSITY

DATE: **February 28, 2002**

AUTHOR: **Mahmood Piraee**

TITLE: **Chlorination Genes of Chloramphenicol Biosynthesis in *Streptomyces venezuelae***

DEPARTMENT OR SCHOOL: Biology

DEGREE: Ph. D. CONVOCATION: May YEAR: 2002

Permission is herewith granted to Dalhousie University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

[redacted]

Signature of Author

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

The author attests that permission has been obtained for the use of any copyright material appearing in the thesis (other than the brief excerpts requiring only proper acknowledgment in scholarly writing), and that all such use is clearly acknowledged.

To God be the glory

and

in memory of my beloved father

Table of Contents

List of figures ix

List of tables xi

Abstract xii

List of abbreviations xiii

Acknowledgements xv

Introduction 1

Literature Review 6

I. Streptomycetes and secondary metabolism 6

II. Antibiotics 10
- a. Discovery of novel antibiotics 11
  - i. Search for novel microorganisms 12
  - ii. Antibiotic-resistance markers 12
  - iii. Target-directed screening 12
  - iv. Genetic approaches 13
  - v. Chemical screening 13
- b. Discovery of antibiotic biosynthesis genes 13
  - i. Complementation of blocked mutants 15
  - ii. Heterologous hybridization probes for searching genomic libraries 15
  - iii. Analysis of DNA adjacent to antibiotic resistance genes ... 16
  - iv. Reverse genetics 16

III. Chloramphenicol 17
- a. Structure activity relationships 17
- b. The Cm biosynthesis pathway revealed by isotopic labeling 18
- c. Origin of the dichloroacetyl group 22
- d. Cm resistance in the producer 24
- e. Regulatory mechanisms in Cm biosynthesis 27
- f. Chromosomal location of Cm biosynthesis genes 30
- g. Identification of genes associated with Cm biosynthesis 30

IV. Biohalogenation 32
- a. Halogenating enzymes 32
  - i. Haloperoxidases 33
    - 1. Heme-dependent haloperoxidases 33
    - 2. Metal-dependent non-heme haloperoxidases 34
  - ii. Perhydrolases 34
  - iii. Halogenases 35

Materials and Methods

I. Bacterial strains and plasmids 39

II. Chemicals and biochemicals 39

III. Media 43

IV. Stock solutions and buffers 48

V. Culture conditions 53

a. Cultures for glycerol stock suspensions 53

i. *E. coli* 53

ii. *Streptomyces* spores 54

b. Cultures for genomic DNA extraction from streptomycetes 54

c. Cultures for plasmid DNA isolation 54

d. Propagation of *S. venezuelae* phage library 54

VI. Feeding of labeled substrates and purification of Cm 55

VII. Measurement of Cm concentration in cultures 55

VIII. Acetylation of Cm 56

IX. Transformation 57

a. Preparation of competent cells 57

b. Transformation of competent cells 57

c. Conjugal DNA transfer between *E. coli* and *Streptomyces* 58

d. Selection and investigation of transconjugants 59

X. DNA isolation procedures 60

a. Isolation of plasmid DNA from *E. coli* cells 60

b. Isolation of plasmid DNA from streptomycetes 60

c. Isolation of genomic DNA from streptomycetes 61

XI. DNA manipulations 62

a. Electrophoresis of DNA 63

b. Directed nested deletions in DNA inserts 64

c. Sequence determination and analysis 64

d. Purification of phage DNA 64

XII. Southern hybridization 65

a. Transfer of phage from plaques to a nylon membrane 65
b. Transfer of DNA from an agarose gel to a nylon membrane 66
c. Preparation of probes 66
d. Hybridization 67
XIII. PCR techniques and TOPO cloning 67
XIV. Phylogenetic trees 69
XV. Introducing ORFs 11, 12 and 13 into mutant *cml-2* 69

Results

Section 1. Feeding experiments 71
Section 2. Isolation of 3'-*O*-acetyl chloramphenicol 73
Section 3. Investigation of acyltransferase activity 76
Section 4. PCR amplification of a halogenase gene from *S. venezuelae* 77
I. Screening a genomic-library with the PCR probe 88
a. Sequence of pJV504 and pJV506 95
b. Analysis of pJV506 sequence 95
II. Identification of ORFs 100
a. ORF11 (encoding CmlK, an AMP-ligase) 106
i. Sequence analysis 106
ii. Functional analysis of CmlK 110
b. ORF12 (encoding CmlS, a halogenase) 114
i. Sequence analysis 114
ii. Phylogenetic comparison of halogenases and monooxygenases 123
iii. Functional analysis of CmlS 123
iv. Analysis of disrupted mutants VS1101 and VS1102 129
c. ORF13 (encoding CmlT, an aldo/keto reductase) 132
i. Sequence analysis 132
ii. Functional analysis of CmlT 132
III. Extension of sequence beyond ORF13 138
IV. Introduction of ORFs 11, 12 and 13 into mutant *cml-2* 145

Discussion

Section 1. Isotopic incorporation experiments 147

Section 2. Possible role for 3'-O-acetylchloramphenicol in Cm biosynthesis 149

Section 3. Search for acyltransferase activity associated with chloramphenicol biosynthesis 150

Section 4. PCR amplification of the halogenase gene fragment 151

I. Functions of gene products 155

a. CmlK: the AMP-ligase 155

b. CmlS: the halogenase 158

c. CmlT: the aldo/keto reductase 160

II. Analysis of the sequence downstream of ORF13 163

III. Complementation of *cml-2* mutant with cloned genes 164

Section 5. The Cm biosynthesis gene cluster 164

Appendix 166

References 168

List of figures

Figure	Description	Page
1.	The chemical structure of chloramphenicol	2
2.	Major pathways leading to biosynthesis of secondary metabolites	8
3.	Incorporation of isotopically labeled carbons in Cm from phenylalanine and CO_2	21
4.	The overall Cm biosynthesis pathway	25
5.	Degradation and modification of chloramphenicol	28
6.	chloramphenicol biosynthesis pathway based on identified genes	31
7.	Mechanism of chlorination of tryptophan catalyzed by a halogenase	38
8.	Alignment of halogenases used to design PCR primers MPF1 and MPR2	79
9.	Hybridization of labeled-PCR product with restriction digests of *S. venezuelae* ISP5230 genomic DNA	81
10.	The putative halogenase PCR-product from *S. venezuelae*	83
11.	Alignment of halogenase sequences with the deduced amino acid sequence of the PCR product cloned in pJV501	84
12.	Agarose gel electrophoresis of products from genomic DNA of six streptomycete strains amplified by PCR with primers MPF1 and MPR2	87
13.	Hybridization signals from recombinant phages probed with the labeled PCR product	89
14.	*Sac*I fragments obtained by digesting DNA from recombinant phages	91
15.	Construction of pJV502 by ligating the 6.2-kb *Sac*I-*Sac*I obtained from recombinant phage MP30 DNA with pBluescript SK+	92
16.	Restriction fragments of pJV502 and pJV503 probed with the labeled-PCR product	93
17.	Construction of pJV506 by digesting pJV502 with *Not*I and religating the remaining 6.8-kb fragment	94
18.	Constructing pJV504 by ligating the 1.5-kb *Sal*I-*Sal*I fragment from pJV502 with pUC18	96
19.	Restriction fragments from *Sal*I digestion of pJV506	97

20. Frameplot analysis of the 3.8-kb *Not*I-*Sac*I fragment (insert of pJV506) containing ORF11 and ORF12 98

21. Relationships between pJV502, pJV506 and the overlapping 2.4-kb *Sac*I-*Bam*HI segment of DNA in the 7.5-kb *Bam*HI-*Bam*HI fragment of He *et al* (2001) 99

22. Restriction map of the 3.8-kb *Not*I-*Sac*I DNA fragment cloned in pJV506 101

23. Nucleotide sequence of the 3.8-kb *Not*I-*Sac*I fragment subcloned in pJV506 102

24. The deduced amino acid sequence (CmlK) of ORF11 107

25. Alignment of the aa sequence of CmlK with sequences of similar proteins with AMP-ligase activity 108

26. Construction of pJV526 by ligating the 6.2-kb *Sac*I-*Sac*I fragment from pJV502 into pJV326 111

27. Construction of disruption plasmid pJV527 by ligating a *Not*I-apramycin cassette into the unique *Not*I site inside ORF11 112

28. Evidence from Southern hybridization that the genome of strain VS1101 contains the Am^R gene inserted by double crossover 113

29. The deduced amino acid sequence of CmlS and its putative NAD(P)H and FAD binding sites 115

30. Alignment of the CmlS sequence with the sequences of halogenases reported from other microorganisms 117

31. Comparison of the putative NADH binding site of CmlS with similar sequences in other halogenases 121

32. Comparison of putative FAD-binding sites in CmlS and other halogenases with the flavin-binding sites identified in PobA (hydroxylase) proteins 122

33. A dendrogram depicting the relatedness of halogenases to certain monooxygenases 124

34. Construction of pJV507 by ligating the 3.8-kb *Sac*I-*Xba*I fragment from pJV506 into the multiple cloning site of pJV326 126

35. Construction of disruption plasmid pJV508 by ligating the *Eco*RV-apramycin cassette into *Stu*I site of ORF12 127

36. Evidence from Southern hybridization for a double cross-over in mutant strain VS1102 128

37. Chromatograms from HPLC analyses of the extract from mutant strain VS1102 130
38. The deduced amino acid sequence (CmlT) of ORF13 133
39. Alignment of the CmlT sequence with other aldo/keto reductases 134
40. Construction of disruption vector pJV518 containing disrupted ORF13 136
41. Evidence from Southern hybridization for a double cross-over in mutant strain VS1103 137
42. Various restriction fragments of MP30 DNA probed with an 850-bp *Sac*I-*Sal*I fragment from the 3'-end of the 3.8-kb *Not*I-*Sac*I fragment 139
43. Hybridization of Southern blot from the *Sac*I digest of MP30 DNA with the 7.5-kb *Pst*I probe from MP30 140
44. Restriction map of the *Sac*I-*Sac*I fragment of the *S. venezuelae* chromosome extending downstream of ORF13 141
45. Nucleotide sequence of the 1.9-kb *Sac*I-*Sac*I fragment cloned in pJV513 that lies beyond the 3.8-kb *Not*I-*Sac*I fragment cloned in pJV506 142
46. Construction of pJV528 by deleting a 2.7-kb *Stu*I-*Eco*RV from pJV526 146
47. ^{1}H-NMR spectra of the compound purified from cultures of mutant VS1102, and an authentic sample of corynecin II 167

List of Tables

1. Bacterial strains, plasmids, and phages 40

2. Enrichment of ^{13}C from labeled compounds in carbons of the dichloroacetyl group of chloramphenicol 72

3. Evidence for esterase-catalyzed hydrolysis of acetyl-Cm by the enzyme located within the insoluble fraction of cell membrane 75

4. Halogenase amino acid sequences and specifications of primers 80

5. Retention times of Cm and corynecins using RP-HPLC 131

6. Aldo/keto-reductase motifs identified in CmlT 162

Abstract

The highly conserved amino acids of FADH-dependent bacterial halogenases were selected as PCR primer target regions. Degenerate primers (MPF1 and MPR2) were designed taking account of codon usage in *Streptomyces*, and were used successfully with a touch-down (TD) PCR procedure to amplify a halogenase gene fragment from the chloramphenicol (Cm)-producer *Streptomyces venezuelae* ISP5230. This was the first PCR-assisted cloning of a halogenase gene fragment. The [α-^{32}P]dCTP-labeled PCR product hybridized under high stringency conditions with a Southern blot of restriction-digested genomic DNA from *S. venezuelae* ISP5230, confirming that the amplicon originated from this organism.

By screening a genomic library of *S. venezuelae* with the [α-^{32}P]-labeled PCR product as a hybridization probe, a segment of DNA containing a cluster of three ORFs (11, 12 and 13) was isolated; the functions of three ORFs were investigated by gene disruption and HPLC analyses of the culture extracts. That ORF11 encoded an AMP-ligase (CmlK; 368 aa) involved in the chlorination reaction of Cm biosynthesis was determined from the results of gene disruption. CmlK lacks a thiolation site and is believed to be involved in activation of an unknown substrate for chlorination. ORF12 encoded a halogenase (CmlS; 535aa) responsible for chlorination of the antibiotic. Putative NAD(P)H-binding sites are conserved in most of known halogenases, and can be recognized in the N-terminal region of the deduced aa sequence of *cmlS*. A second motif, resembling the FAD-binding site of monooxygenases, is also present in CmlS. This is believed to be required for the chlorination reaction. Disruption of either *cmlK* in strain VS1101 or *cmlS* in strain VS1102 caused accumulation of non-chlorinated congeners of Cm. Culture extracts of the disrupted strains contained mainly corynecin II, which has a propionyl in place of the dichloroacetyl group of Cm. Corynecin II was purified from strain VS1102 and its structure was confirmed by ^{1}H-NMR spectroscopy. The deduced amino acid sequence of ORF13 (CmlT; 292 aa) resembled the sequences of aldo/keto reductases that catalyze reduction of a carbonyl group to the corresponding alcohol. Since disruption of *cmlT* had no effect on antibiotic production, CmlT does not have a role in Cm biosynthesis. Extending the cloned chromosome sequence into the region downstream of ORF13 located two more ORFs (14 and 15) , both of which encoded proteins with unknown functions, and thus could not be assigned a role in Cm biosynthesis.

List of Abbreviations

AKR	aldo/keto-reductase
A	alanine
Am	apramycin
Am^R	apramycin resistance
AMP	adenosine monophosphate
Ap	ampicillin
ASA	acetylsalicylic acid
D	aspartic acid
BSA	bovine serum albumin
CIAP	calf intestinal alkaline phosphatase
cDNA	complementary DNA (synthesized on an RNA template)
Cm	chloramphenicol
CoA	coenzyme A
Da	Dalton
DAHP	deoxy-*arabino*-heptulosonate-7-phosphate
dCTP	deoxycytidine 5'-triphosphate
DEAE-cellulose	diethylaminoethyl cellulose
DMAP	4-(dimethylamino)pyridine
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
EDTA	ethylenediamine tetraacetic acid
FAD	flavin adenine dinucleotide
H	histidine
HPLC	high-performance liquid chromatography
HPLC-MS	HPLC & mass spectroscopy
HPO	haloperoxidase
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase
Km	kanamycin

K	lysine
NAC	*N*-acetylcysteamine
NAD	nicotinamide adenine dinucleotide
NRPS	nonribosomal peptide synthetase
ORF	open reading frame
PABA	*p*-aminobenzoic acid
PAPA	*p*-aminophenylalanine
PAPS	*p*-aminophenylserine
PCP	peptidyl carrier protein
PCR	polymerase chain reaction
PEG	polyethylene glycol
PKS	polyketide synthase
PMSF	phenylmethylsulfonyl fluoride
qs	sufficient quantity to
RP-HPLC	reverse phase HPLC
rpm	revolutions per minute
Rt	retention time
SDS	sodium dodecyl sulfate
SSC	sodium chloride-sodium citrate solution
TD-PCR	touchdown-PCR
TLC	thin layer chromatography
Tn	transposon
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
Ts	thiostrepton
Y	tyrosine
UV	ultra violet
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Acknowledgements

First and foremost, I would like to express my sincere thanks and gratitude to my supervisor, Dr. Leo C. Vining, for his great guidance, supportive advice, and encouraging patience through out the preparation of this thesis. He taught me valuable lessons in science as well as in general life.

My special thanks goes to Dr. Robert L. White for his valuable discussions. His comments and encouragement came at times when they were most needed. I would like to extend my thanks to Dr. Karl-Heinz van Pée and Dr. Jonathan M. Wright for reading my thesis and for their valuable comments.

I would like to express my thanks to Dr. Stephen Armstrong whose conversations were a source of inspiration for staying happier and focused. I would also like to thank my lab mates Liru, Jian-Yong and Brad for conversations and sharing research ideas, and to all my friends outside the lab for cheerful times and soccer games.

I am appreciative of the Ministry of Health & Medical Education of Iran, Dalhousie Department of Graduate Studies, and the Patrick Lett Fund for their financial support during my studies.

And last but not least, I would like to thank my family whose love and support made my education more joyful; thank you for believing in me.

Introduction

Chloramphenicol (Cm) is produced as a secondary metabolite by *Streptomyces venezuelae* and certain other actinomycetes. It is a broad-spectrum antibacterial antibiotic, and is considered the drug of choice for treatment of persistent *Salmonella* infections (Mitscher, 1995). Among natural compounds it is unique in having not only a nitrobenzene moiety but also a dichloroacetyl side chain (Figure 1) formed in a halogenase-catalyzed chlorination reaction. Many attempts have been made to find the biosynthetic pathway for Cm, particularly the chlorination steps. A plausible series of reactions leading to the antibiotic has been deduced from analyses of blocked mutants and the incorporation patterns of isotopically labeled substrates (Doull *et al.*, 1985; McGrath *et al.*, 1968; Vining & Stuttard, 1995). Recent studies (Chang *et al.*, 2001) have shown that the biosynthesis of Cm and *p*-aminobenzoic acid share 4-amino-4-deoxychorismic acid as a common intermediate. The key metabolite leading to Cm is *p*-aminophenylalanine, which is produced from 4-amino-4-deoxyprephenic acid enzymatically by a postulated arylamine synthase complex (Jones & Westlake, 1974). This complex is now believed to include aminodeoxychorismate synthase (Chang *et al.*, 2001). The next step is the conversion of *p*-aminophenylalanine to *p*-aminophenylserine (*p*-APS), which then becomes the substrate accepting a dichloroacetyl group (McGrath *et al.*, 1968). The dichloroacetyl substituent is essential for the potent and specific antibacterial activity of Cm. The main focus of research presented in this thesis has been elucidation of the chlorination reaction in Cm biosynthesis. Biochemical and genetic tools, particularly PCR and gene function analysis have been used.

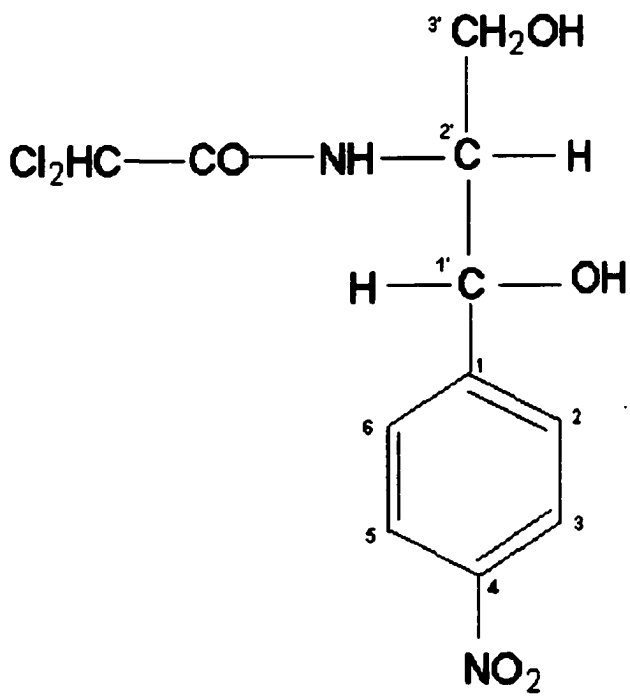


Figure 1. The chemical structure of chloramphenicol.

Based on the pathway postulated for Cm biosynthesis (Vining & Stuttard, 1995), an acyl transferase is expected to deliver the dichloroacetyl group from a suitable donor to *p*-aminophenylserine (Vining & Stuttard, 1995). Congeners of Cm with *N*-acyl groups lacking halogen atoms have been detected at low levels in cultures of *S. venezuelae* during Cm production. The same compounds were also produced by *S. venezuelae* wild-type cultures deprived of halogen ions, and in some variant strains occurring naturally or prepared by mutagenesis (Doull *et al.*, 1985). These findings implied the presence of an acyltransferase able to transfer halogen-free acyl groups (Simonsen *et al.*, 1978) in a reaction similar to that for transfer of the dichloroacetyl group, and it was proposed that the acyltransferase and dichloroacetyl transferase activities are catalyzed by a single enzyme with higher affinity for the dichloroacetyl group. Therefore, one component of the current study was to search for an enzyme that catalyzes transfer of an acetyl group to *p*-aminophenylserine.

The putative product of the dichloroacetylation reaction, *N*-dichloroacetyl *p*-aminophenylserine, undergoes a series of reduction and oxidation reactions to produce Cm. The probable reaction sequence has been deduced from the incorporation of isotopically labeled potential intermediates and examination of mutants (Vining & Westlake, 1964; McGrath *et al.*, 1968; Doull *et al.*, 1985). However, the origin of the dichloroacetyl component itself has remained uncertain. In studies carried out in this and other laboratories before the present work began, dichloroacetic acid labeled in various positions, had been tested as a precursor in the biosynthesis of Cm, and shown to have no direct role (Gottlieb *et al.*, 1956; Simonsen *et al.*, 1978). Nevertheless, results consistent with formation of the dichloroacetylating agent via acetyl-coenzyme A (CoA) had been

reported (Munro *et al.*, 1975), and some evidence suggested formation of the dichloroacetyl substituent from malonyl-coenzyme A (Simonsen *et al.*, 1978). Subsequent results (Ranade and Vining unpublished) confirming enrichment of the dichloroacetyl carbonyl group by $^{13}CO_2$, but showing no ^{13}C enrichment in the fatty acids of mycelial lipids excluded malonyl CoA as a precursor for the dichloroacetyl group of Cm. Since the cumulative results did not provide a clear picture of the reactions and substrates involved in the chlorination step of Cm biosynthesis, possible roles for CoA derivatives of acetic, acetoacetic and dichloroacetic acids in introducing the dichloroacetyl moiety into Cm were investigated using ^{13}C-labeled thioesters of *N*-acetylcysteamine (NAC). These thioesters are structural analogues of CoA esters, and are capable of mimicking the thiol terminus of the pantotheine unit of CoA. In the initial part of this study (see Results, Section 1) the ^{13}C-labeled analogues prepared in the Department of Chemistry, Dalhousie University (Lewis, 1998) were fed to Cm-producing cultures, and the purified antibiotic was analyzed by ^{13}C-NMR spectroscopy to determine the isotopic labeling pattern.

As a secondary-metabolite producer, *Streptomyces venezuelae* ISP5230 is equipped with mechanisms for protecting its biological systems from the inhibitory effects of Cm. Two such mechanisms, antibiotic degradation and antibiotic efflux (Mosher *et al.*, 1990; Mosher *et al.*, 1995) are the strategies identified so far in this organism. Degradation of the antibiotic occurs only if the organism is in the logarithmic growth phase and not producing the antibiotic. At this time it is susceptible to Cm entering the mycelium from the environment, and inactivates the antibiotic by two separate routes; one is initiated with a hydrolase removing the dichloroacetyl group, and the other uses an

O-phosphokinase to form the nontoxic Cm-3'-*O*-phosphate. In the second part of this study (see Results, Section 2) an additional protective mechanism integrated into the Cm biosynthetic pathway was discovered. Isolation of 3'-*O*-acetylchloramphenicol from cultures producing Cm, suggested a role for this biologically inactive intermediate in masking antibiotic activity during the assembly of Cm. The role of acetylation as a step in the biosynthetic pathway, the presence of an esterase in the mycelium to liberate Cm and its possible link with antibiotic efflux were explored.

The third part of this study (see Results, Section 3) consisted of a search for the proposed nonspecific transferase responsible for introducing the acyl side chain into Cm and its corynecin congeners.

In the fourth part of the study (see Results, Section 4), a halogenase gene encoding the enzyme catalyzing the chlorination reaction was sought by PCR cloning methods. A gene cluster containing genes that encode a halogenase and an AMP-ligase was cloned and the function of these two genes in Cm biosynthesis was confirmed by insertional inactivation.

Literature review

I. Streptomycetes and secondary metabolism

Bacteria in the genus *Streptomyces* are high G+C Gram-positive filamentous soil bacteria grouped phylogentically among the actinomycetes. They exhibit a complex life cycle during which the elongated and branching vegetative hyphae form aerial mycelium and spores that disperse to assist colonization of their soil habitat. *Streptomyces* strains produce many useful natural products such as antibiotics, and in recent times have been engineered by recombinant DNA techniques to make novel substances. Over 70% of the known naturally occurring antibiotics, representing classes such as macrolides (e.g., erythromycin), aminoglycosides (e.g., gentamycin), tetracyclines (e.g., aureomycin), ansamycins (e.g., rifamycin), the polyene antifungal agents (e.g., amphotericin) and antitumor agents (e.g., daunorubicin), have been isolated from this genus. The organisms have long been known also as a source of protein products with useful enzymatic activities. Among those commercially available are pronase from *Streptomyces griseus*, amylase from *Streptomyces amyloliquefaciens*, and many restriction endonucleases (Tomich, 1988; Vining, 1992; Lancini & Lorenzetti, 1993, Glazer & Nikaido, 1995; August *et al.*, 1999).

In contrast to the products of primary metabolism that are normally retained within the organism to support growth and reproduction, secondary metabolites are excreted to exert effects outside the producer, and thus mediate interactions between the organism and its environment. Antibiotics, mycotoxins, insecticides and herbicides, all products of secondary metabolism, have targets external to the producer (Vining, 1992; Glazer & Nikaido, 1995). Because secondary metabolites are not essential for growth of the

producer, but may be important for survival in the environment, they are often produced at the end of the growth phase. Elimination by mutagenesis of genes encoding proteins associated with secondary metabolite biosynthesis does not affect the growth of the producing organism.

The external effects of secondary metabolites probably provide selectional advantages to the producer in competitive environments such as soil, and the benefits justify devotion of the significant proportion of the genome and energy used for secondary metabolism. The diverse chemical structures of secondary metabolites are derived from primary metabolites after complex modifications (Figure 2) requiring usually more than 10 structural genes in addition to regulatory and self-defense genes. The timing and expression of secondary metabolite biosynthesis genes are controlled by regulatory genes through specific mechanisms, linked with the physiology of the producing organism. Through these mechanisms, secondary metabolism is initiated only when there is a slow down in growth rate, a deficiency in nutrients, or various other physiological stresses; small signaling molecules are known to be involved (Horinouchi & Beppu, 1995). Antibiotic biosynthesis in streptomycetes is coordinated with cell differentiation (e.g., sporulation). The evolution of these sophisticated regulatory systems (Vining & Stuttard, 1995) is an indication that secondary metabolism is advantageous to the organism (Vining, 1992).

Similarities between the genes of secondary metabolism, and certain genes of primary metabolism suggest that pathways for the biosynthesis of secondary metabolites have arisen from existing primary pathways by gene duplication and modification. The complexity of secondary metabolic pathways, as well as their closely

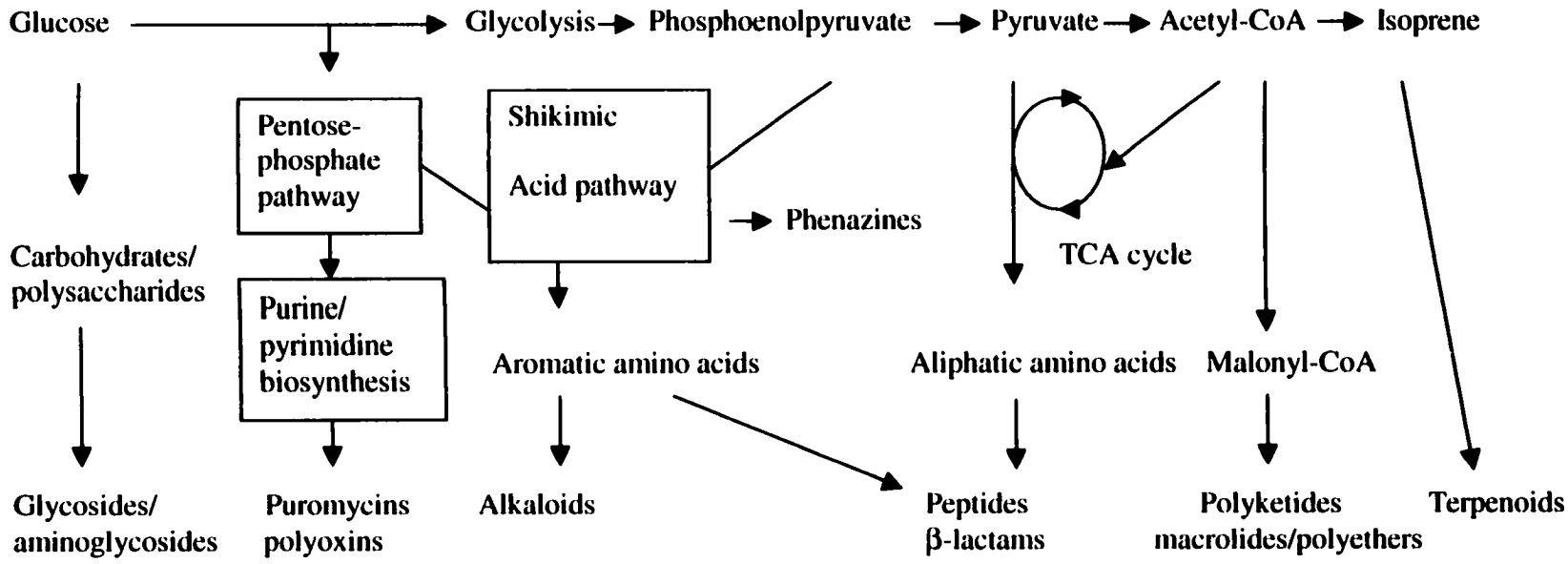


Figure 2. Major pathways leading to biosynthesis of secondary metabolites (modified from August *et al.*, 1999).

coordinated regulatory and self-defense mechanisms, and the fact that their genes are clustered, all imply that these activities have evolved together and persist because they provide advantages to the organism (Maplestone *et al.*, 1992; Vining, 1992).

Although the chemical structures of secondary metabolites are extremely diverse, they can be grouped according to the starting materials used for their biosynthesis: 1) those derived via shikimic acid, such as Cm, and plant phenolics; 2) those made from amino acids, such as alkaloids and peptide antibiotics (including penicillins and cephalosporins); and 3) metabolites derived from acetate, such as the macrolide and polyketide antibiotics (including erythromycin and tetracyclines). The ultimate source of the starting materials for each of the groups is, nevertheless, photosynthesis, intermediary metabolism, and the biosynthesis pathways of primary metabolism. Reactions involved in biosynthesis of secondary metabolites could also be classified into three types: 1)- Transformation of primary metabolites, such as amino acids, into biosynthetic intermediates. For example tryptophan and proline are transformed into intermediates for biosynthesis of the antibiotics pyrrolnitrin and pyoluteorin, respectively (Hammer *et al.*, 1997; Nowak-Thompson *et al.*, 1999). Transformation of chorismic acid to *p*-aminophenylalanine and its β-hydroxylation for use in Cm biosynthesis (Vining & Stuttard, 1995), or modification of sugars for aminoglycoside biosynthesis (Hotta *et al.*, 1995) are other instances. 2)- Polymerization of small metabolites to form intermediates or final products. Examples of this type of reactions are the condensation of acetate and malonate in the biosynthesis of polyketides (Vining & Hopwood, 1995), condensation of amino acids in the formation of peptide antibiotics (Vining & von Dohren, 1995), and

condensation of isoprenoids in terpenoid biosynthesis (Cane, 1995). 3)- Modification reactions such as halogenation and glycosylation that occur in the assembled structures.

II. Antibiotics

Like other secondary metabolites, antibiotics are produced abundantly and in the greatest structural diversity by actinomycetes. Antibiotics are widely used in agriculture, animal husbandry, and the food industry, as well as in the therapy of infectious diseases in humans. The era of antibiotic discovery owed much of its success to Selman Waksman's interest in the diversity and interactions of microorganisms in the soil. His research was focused on the antagonistic relationships among soil microorganisms, particularly actinomycetes, and the role of their products. The antagonistic substances were named "antibiotics", a term defined as "a chemical substance, derived from a microorganism which has the capacity to inhibit growth, and even destroy, other microorganisms in dilute solution" (Woodruff, 1996; Strohl, 1997). The effect of antibiotics on other organisms provides competitive advantage to the producer when resources and environmental conditions do not favor growth. If the resources are abundant, antibiotic production is suppressed since the producer can compete by rapid growth, but when the growth rate decreases, antibiotics can be used to suppress competition for scarce nutrients. Antibiotic producers control the timing and level of antibiotic production by complex regulatory mechanisms. Feedback regulation is rare, because the product is excreted and does not interact with the biosynthetic system. Secondary metabolites are usually produced only after cellular growth declines in response to deficiency of an essential nutrient. Where this leaves the culture with an excess of glucose, the biosynthesis of many antibiotics is repressed. Actinomycin,

cephamycin, kanamycin, streptomycin and penicillin are specific examples of antibiotics whose biosynthesis is sensitive to the repressive effect of glucose. Nitrogen and phosphate sources can also regulate the biosynthesis of antibiotics by repressive or inhibitory effects.

While these physiological factors exert broad control of the production of secondary metabolites, more specific regulation of antibiotic biosynthesis is usually due to pleiotropic effectors, which at very low concentrations selectively control a wide range of activities including antibiotic biosynthesis. One such substance is A-factor, a well-studied member of the butanolide family of autoregulators that is essential for induction of sporulation and streptomycin biosynthesis in *S. griseus* (Retzlaff *et al.*, 1993; Vining & Stuttard, 1995).

a. Discovery of novel antibiotics

The conventional approach to discovering antibiotics from natural sources has been based on screening for agents with the ability to inhibit pathogenic organisms. Novel products have been sought by protoplast fusion of two producing microorganisms, chemical modification of generic molecules such as β-lactams, or genetic modification of the biosynthetic gene clusters to obtain antibiotics with improved activity (White *et al.*, 1988; Omura, 1992; Strohl, 1997).

There is an increasing need to discover novel antimicrobial agents to replace current classes of antibiotics, the efficacy of which is limited due to emergence of resistance among pathogenic microorganisms. New lead structures with novel targets are not only needed against resistant Gram-positive bacteria, but also for use against tuberculosis and mycotic, protozoal or viral infections. The current tools of antibiotic discovery need

to be improved so that agents effective against strains with resistant mechanisms are available. Some methods that may be used are described briefly here.

i. Search for novel microorganisms

Since antibiotic production is usually strain specific, searching for new strains or exploring new environments such as deserts, oceans, and tropical rain forests (Glazer & Nikaido, 1995; McVeigh *et al.*, 1996; Strohl, 1997) may increase the chance of isolating strains producing novel products.

ii. Antibiotic-resistance markers

Antibiotic-producing organisms are resistant to their own antibiotic (Demain, 1974; Vining, 1979), and the resistance genes are usually clustered with antibiotic biosynthesis genes (Vining & Stuttard, 1995). By using specific probes for antibiotic-resistance genes in hybridization of Southern blots, biosynthetic gene sequences can be located among potential antibiotic producers.

iii. Target-directed screening

Some recent methods of screening for new antimicrobials are based on detecting a specific mechanism rather than inhibitory action on bacterial whole cell functions. These methods have high sensitivity; their ability to detect a low concentration of the active ingredient increases the probability of finding a new substance. Specificity of the screen can increase sharply by targeting sites in bacterial cells such as antibiotic efflux pumps and components of cell membrane biosynthesis that are absent from mammalian cells (Kuntz, 1992; Desnottes, 1996).

iv. Genetic approaches

Molecular genetic techniques can be used to activate undiscovered cryptic biosynthetic pathways by derepressing repressor genes (Strohl, 1997). Other possibilities include the construction of hybrid antibiotics with new and improved activities by combinatorial biosynthesis (Stemma, 1994; Hutchinson, 1999), and the designing of species-specific probes to search for new producers that are unculturable under routine laboratory conditions (Amann *et al.*, 1994).

v. Chemical screening

Promising chemical entities can be sought without first looking for biological activity. In such screens only compounds that show particular physico-chemical characteristics detected chromatographically by TLC, HPLC, or HPLC-MS are isolated. These lead structures are tested for their biological activity and antimicrobial spectrum. Using this approach tetracenomycins B_3 and D_3 were discovered in *Streptomyces olivaceus* TU2353 by HPLC. The compounds proved to be structurally similar to tetracenomycins A_2, B_1, B_2, D and have moderate antibacterial activity (Rohner *et al.*, 1988).

b. Discovery of antibiotic biosynthesis genes

The increased knowledge of antibiotic biosynthesis pathways provided by identification and characterization of the genes involved allows a rational approach to increasing metabolic flux into their biosynthetic pathways. Increasing the copy number of a strategic gene (e.g., encoding a rate limiting enzyme) has been used to improve cephamycin production significantly in recombinant strains of *Streptomyces*

clavuligerus with double the copy number of the gene encoding lysine ε-aminotransferase (Malmberg *et al.*, 1993). Many regulatory genes controlling antibiotic production have been discovered within antibiotic biosynthesis gene clusters, and it is possible to overproduce the metabolite by increasing the copy number of positively acting, pathway-specific regulators. *act*II-ORF4 (Arias *et al.*, 1999) and *dnr*R1-2 (Stutzman-Engwall *et al.*, 1992) are examples of positive regulators found in the biosynthetic gene clusters for actinorhodin and daunorubicin, respectively. Alternatively, manipulation of genes may assist searches for improved activities by altering the chemical structure of the antibiotic. Once gene clusters for biosynthesis of antibiotics are identified and their organization is uncovered, it is possible to produce hybrid metabolites by genetic engineering. For example, introducing genes from the actinorhodin biosynthetic cluster of *S. coelicolor* into *Streptomyces* sp. AM-7161 (a medermycin producer), and into *Streptomyces violaceoruber* Tu22 (a producer of granaticin and dihydrogranaticin) gave the first hybrid antibiotics mederrhodins A and B, and dihydrogranatirhodin, respectively (Hopwood *et al.*, 1985b). However, these novel approaches depend on detailed knowledge of biosynthetic gene clusters and the enzymology of the biosynthetic reactions. Antibiotic biosynthetic genes are in principal difficult to clone because antibiotics are synthesized by gene clusters containing 10-30 structural genes as well as regulatory and resistance genes. The entire cluster must be cloned in a suitable host to detect antibiotic production. However a variety of approaches have been developed and many antibiotic biosynthesis gene clusters have been identified; a summary of the methods is presented below.

i. Complementation of blocked mutants

Mutants blocked in antibiotic biosynthesis are used as hosts for transformation with fragments of wild-type genomic DNA cloned in appropriate vectors. If a reliable screening procedure to detect antibiotic production is available, complementation of the mutant allows detection of a particular gene or gene cluster in the wild-type strain (August *et al.*, 1999). Cloning one gene from an antibiotic biosynthesis cluster (Vining & Stuttard, 1995) provides the opportunity to isolate contiguous genes in a cluster.

The first aromatic polyketide synthase gene cluster was isolated from *S. coelicolor* by complementation of mutants blocked in actinorhodin biosynthesis (Malpartida & Hopwood, 1984). In a variation of this strategy, complementation can be used to clone a gene for secondary metabolism if it has a counterpart in primary metabolism. The *pabAB* gene required for Cm biosynthesis in *S. venezuelae* was identified by shotgun cloning and complementing the auxotrophic *S. lividans* strain JG10 that cannot synthesize PABA and is unable to grow on minimal medium containing sulphanilamide (Brown *et al.*, 1996).

ii. Heterologous hybridization probes for searching genomic libraries

The sequences of known gene clusters have been used as heterologous probes to identify homologous clusters from other species. For instance, *act*I from the actinorhodin biosynthetic gene cluster of *S. coelicolor* A3 (2) (Malpartida *et al.*, 1987) has been used to probe for aromatic PKS genes of other organisms. The PKS gene clusters for biosynthesis of granaticin from *Streptomyces violaceoruber* (Sherman *et al.*, 1989), and jadomycin B from *S. venezuelae* (Yang *et al.*, 1995) were identified by this

strategy. The method can potentially lead to identification of new pathways that have never been detected.

iii. Analysis of DNA adjacent to antibiotic resistance genes

As noted earlier self-defense genes are usually clustered with antibiotic biosynthesis genes, and are relatively easy to isolate by shotgun cloning in sensitive hosts and selecting antibiotic-resistant transformants. The structural genes for antibiotic biosynthesis are then located by examining the DNA surrounding the resistance genes. The gene cluster for biosynthesis of chlortetracycline was identified in this way (Dairi *et al.*, 1995).

iv. Reverse genetics

It is sometimes possible to purify an enzyme used in the biosynthesis of an antibiotic, when one of the pathway intermediates is known. Characterization of the purified protein gives information that can be used in many ways to isolate the encoding gene. Antibodies raised against the protein can be used to probe expressed genomic or cDNA libraries of the producer; amino acid sequence of the peptide fragments of the protein can be used to design degenerate primers for PCR amplification of a region of the encoding gene (August *et al.*, 1999). Moreover, if the sequences of protein homologues of the gene product are available in databases, degenerate primers can be designed from consensus core sequences to amplify the desired gene. Non-ribosomal peptide synthases have been cloned from *Pseudomonas syringae* and *Bacillus lichniformis* by this technique (Turgay & Marahiel, 1994). In another approach, the amino acid sequence of a purified enzyme can be used to design a degenerate

oligonucleotide for probing a genomic library directly, and cloning the corresponding gene. A bromoperoxidase-catalase was isolated from *S. venezuelae* ISP5230 by this procedure and shown not to be involved in Cm biosynthesis (Facey *et al.*, 1996).

III. Chloramphenicol

Chloramphenicol was first isolated in 1947 by Ehrlich and coworkers from *Streptomyces venezuelae* collected in Caracas, Venezuela, and later from several other organisms found in soil samples from widely separated locations. Because Cm has a simple chemical structure it was soon synthesized chemically, using *p*-nitroacetophenone as the starting compound (Long & Troutman, 1949). Cm was an extremely popular antibiotic due to its broad-spectrum of activity and effectiveness against some infectious diseases not amenable to treatment with other drugs. It is specifically used for infections caused by penicillin G- and ampicillin-resistant bacteria, such as *Haemophilus influenzae*, *Salmonella typhi*, *Streptococcus pneumonia*, and *Neiseria meningitidis*. It is also used to treat rickettsial infections such as Rocky Mountain spotted fever, but occasionally can cause a fatal anaemia (Martin, 1998). It binds reversibly to the 50S ribosomal subunit, and appears to inhibit the binding of the aminoacyl tRNA to the acceptor site on the ribosome. This prevents the interaction between peptidyltransferase and its amino acid substrate, causing inhibition of peptide bond formation (Mitscher, 1995).

a. Structure activity relationships

To correlate the structure of Cm with its antibacterial activity, two main groups of analogues have been synthesized and tested. One has modifications in the side chain, and

the other is modified in the para substituent of the ring. While, the inductive (electron withdrawal) effects of the side-chain acyl group are important for antibacterial activity, the hydrophobic properties of the side chain are less essential, and bulky substituents are undesirable. Analogues with $CHCl_2$ replaced by $CHBr_2$, CH_2Cl, CHF_2, CH_2Cl, $CHClCH_3$, or CCl_3 are less active than Cm, but replacement with CF_3 gave 1.7 times higher activity; groups such as C_6H_5 or $CH_2C_6H_5$ completely eliminated the antibacterial activity. Oxidation of the alcohol group on C-1' caused a significant loss of activity (Hansch *et al.*, 1973). Acetylation of the 3'-hydroxy group of Cm, which also occurs in certain strains of bacteria as a resistance mechanism catalyzed by Cm acetyltransferase (Shaw, 1983), renders the product inactive.

In contrast to its lack of importance in the side chain, the hydrophobicity of substituents in the *para* position of the ring is most important. For example, replacing the NO_2 group with C_6H_5 gave a derivative with three times more antibacterial activity than replacing it with NH_2, though the phenyl analog was still only 10% as active as Cm. Reduction of the nitro group, brought about also by the normal flora of the gut, gives an inactive metabolite. Substitution of NO_2 by H, or SCH_3 or $COCH_3$, OCH_3 lowers the activity, and moving the NO_2 group from the *para* to the *meta*-position on the ring causes an 80% decline in activity (Hansch *et al.*, 1973). Metabolites of Cm, such as the C-3' glucuronide and deamination products formed in the human body, are all inactive (Mitscher, 1995; Martin, 1998)

b. The Cm biosynthesis pathway revealed by isotopic labeling

Because of its phenylpropanoid structure (Figure 1), Cm was at first thought to share its biosynthesis pathway with those of the aromatic protein amino acids. Early

studies showed stimulation of Cm production when the cultures were supplemented with phenylalanine, but structurally unrelated amino acids, such as leucine and norleucine also increased production, and subsequent studies with radioactive supplements showed that these compounds were not biosynthetic intermediates (Gottlieb *et al.*, 1962; Vining & Westlake, 1964). *p*-Nitrophenylserinol also appeared to stimulate antibiotic production, but using the labeled compound showed that it was not converted to Cm; instead, it was N-acetylated (Gottlieb *et al.*, 1956).

Incorporation of label from [6-^{13}C]glucose into carbons 1', 2 and 6, and from [1-^{14}C]glucose, [2-^{14}C]glucose, and [2-^{14}C]glycine into the ring and side-chain carbon atoms of the *p*-nitrophenylserinol moiety of Cm, and also into cellular phenylalanine, implicated the shikimate pathway in the formation of Cm. Consistent with this was the selective incorporation of [U-^{14}C]shikimate into chloramphenicol and the phenylpropanoid amino acids of mycelium (Vining & Westlake, 1964). Supporting evidence was provided by using labeled acetate or pyruvate (O'Neil *et al.*, 1973), or [2-^{14}C]lactate (Vining & Westlake, 1964).

Later results showing L-[U-^{14}C]phenylalanine, L-[*carboxy*-^{14}C]phenylalanine, L-[*carboxy*-^{14}C]tyrosine, DL-[2-^{14}C]-*p*-hydroxyphenylserine, and [2-^{14}C]acetate to be poor precursors of the nitrophenylserinol moiety strengthened the conclusion that the phenylpropanoid moiety of Cm was not synthesized from aromatic amino acids. In fact, most of radioactivity from L-[*carboxy*-^{14}C]phenylalanine, and L-[U-^{14}C]phenylalanine was incorporated into the dichloroacetyl moiety (Vining & Westlake, 1964; Figure 3).

Experiments with cell-free systems provided evidence for conversion of chorismate to *p*-aminophenylalanine. Conversion of ^{14}C-labeled chorismate to *p*-

aminophenylalanine was catalyzed by arylamine synthase an enzyme activity absent from nonproducing strains of *S. venezuelae* (Jones & Westlake, 1974; Jones & Vining, 1976).

The next step after formation of *p*-aminophenylalanine is believed to be C-1' hydroxylation because DL-[*carboxy*-^{14}C]*threo*-*p*-aminophenylserine labeled C-3' of chloramphenicol efficiently, whereas [α-^{14}C]L-*p*-nitrophenylalanine did not. L-*p*-Nitrophenylalanine failed to label Cm, even when the N-acetyl derivative was used to improve permeability. Since radioactivity was also not incorporated into Cm from [3-^{14}C]*p*-nitrophenylserine or [3-^{14}C]*p*-aminophenylserinol, the next reaction was concluded to be either acylation or chloroacylation of aminophenylserine. Evidence from mutants suggested that chloroacylation is the most likely reaction (Doull *et al.*, 1985).

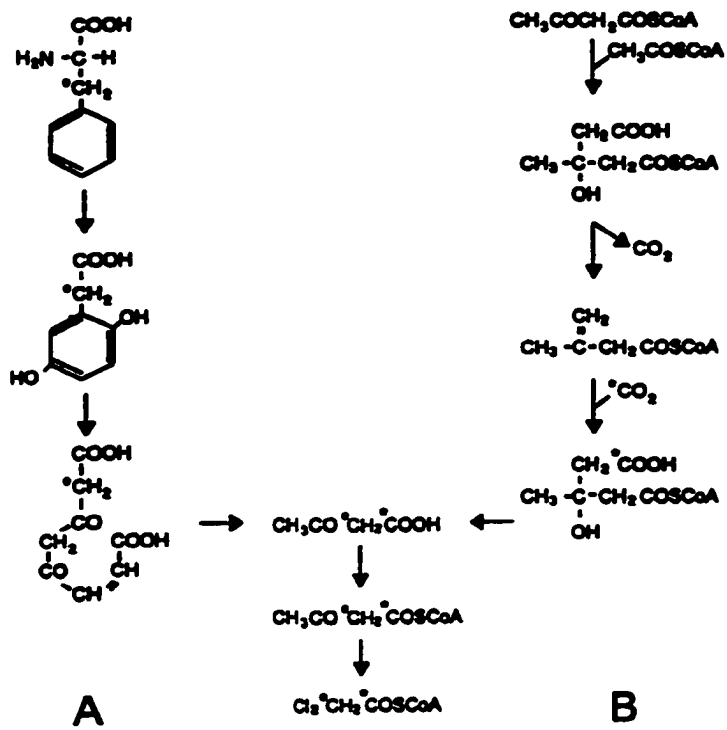


Figure 3. Possible pathways leading to incorporation of isotopically labeled carbon to Cm from phenylalanine (A) and CO_2 (B) (Vining & Stuttard, 1995).

c. Origin of the dichloroacetyl group

The origin of the dichloroacetyl group has been explored many times with isotopically-labeled precursors. Radioactivity from [1-^{14}C]acetate and [2-^{14}C]acetate was incorporated into the carbonyl and dichloromethine carbons of the Cm dichloroacetyl group, respectively. [2-^{14}C]Acetate also enriched the aromatic ring of Cm but at a lower intensity. The pattern of incorporation from [1,2-$^{13}C_2$]acetate showed that most of the isotope entered the dichloromethine carbon of the dichloroacetyl group after scission of the original ^{13}C-^{13}C bond (Simonsen *et al.*, 1978). Unequal incorporation of carbons from [1,2-$^{13}C_2$]acetate, and exclusive incorporation of label from [^{14}C]formate (readily converted to [^{14}C]carbon dioxide) and from [^{14}C]carbon dioxide itself into the dichloroacetyl carbonyl group suggested that the carbonyl group of acetate exchanged with carbon dioxide during formation of the dichloroacetyl moiety (Figure 3).

Incorporation of [2,3-$^{13}C_2$]succinate labeled only the carbonyl carbon of the dichloroacetyl group, suggesting that equilibration of the carbonyl group with carbon dioxide does not involve Krebs cycle intermediates such as oxaloacetate (Simonsen *et al.*, 1978).

Incorporation of labeled-carbon from [6-^{13}C]glucose into the dichloromethine carbon of Cm is consistent with a route via acetyl coenzyme A (Munro *et al.*, 1975). Subsequent results suggested that the dichloroacetyl moiety is formed before it is attached to the aromatic component, presumably as an activated dichloroacetyl intermediate such as an ester of coenzyme A (Doull *et al.*, 1985). Based on the assumption that dichloroacetyl coenzyme A acylates a phenylpropanoid intermediate,

the precursor that receives the chlorine atom was suggested to be either acetoacetyl coenzyme A or malonyl CoA. The evidence mentioned above that Krebs cycle intermediates are not directly involved ruled out oxaloacetate as the precursor. Earlier precursor feeding experiments with ^{14}C, ^{3}H, and ^{2}H labeled dichloroacetic acid had ruled out the free acid as an intermediate (Gottlieb *et al.*, 1956). Considering the evidence that CO_2 was incorporated into the dichloroacetyl group, a role for malonyl CoA formed from acetyl CoA and carbon dioxide would require its equilibration with free malonic acid. This being so, the malonyl-derived fatty acids of the mycelium should be labeled. However, when cultures were supplemented with labeled CO_2, no isotope enrichment was observed in fatty acids, while Cm showed appreciable enrichment. This ruled out malonyl CoA as an intermediate in generation of the dichloroacetyl group, and favored acetoacetyl CoA as the precursor for the chlorination reaction. Specific incorporation from C-3 of phenylalanine into the dichloromethine carbon of Cm strengthened this assumption and suggested a route via formation of acetoacetate. Phenylalanine can generate acetoacetate by metabolism via the homogentisate pathway. Since both carbons of the dichloroacetyl group in Cm derived from [U-^{14}C]phenylalanine had about the same specific activity, these two carbons may be derived via acetoacetate as a unit. Moreover, the incorporation of CO_2 into the carbonyl group could occur via β-hydroxy-β-methylglutaryl coenzyme A generating, eventually, acetoacetate (Figure 3). Consistent with this was the equal incorporation from [U-^{14}C]leucine into the dichloroacetyl carbons, implying that the two carbons are also derived as a unit during metabolism of leucine to acetoacetate (McGrath *et al.*, 1968; Vining & Stuttard, 1995). Nevertheless, the low incorporation from [1,2

$^{14}C_2$]acetate, and absence of free dichloroacetic acid incorporation indicate a more complex route (Simonsen *et al.*, 1978). After chloroacylation, the two reactions that complete the biosynthesis of the antibiotic are reduction of the carboxyl group and oxidation of the *p*-amino group.

During production of Cm small amounts of N-dichloroacetyl-*p*-aminophenylserinol have been detected in cultures, and it has been shown that the ^{14}C–labeled compound is converted to Cm without randomization of the label (McGrath *et al.*, 1968). Therefore, oxidation of the amino group was suggested (Doull *et al.*, 1985) to be the final step in Cm biosynthesis, giving the overall pathway shown in Figure 4.

d. Cm resistance in the producer

Cm inhibits protein synthesis in bacteria by binding reversibly to the 50S ribosomal subunit and inhibiting peptidyl transferase activity. Protein synthesis in *S. venezuelae* is also sensitive to this inhibitory activity, when tested *in vitro* (Mitscher, 1995, Vining & Stuttard, 1995), and is equally sensitive in extracts from Cm-producing and nonproducing mycelium (Malik & Vining, 1972). However, *S. venezuelae* mycelium is sensitive *in vivo* only when not producing the antibiotic, and acquires resistance inducibly upon exposure to Cm, whether produced endogenously or added exogenously (Malik & Vining, 1970). The non-producing cultures become resistant to Cm at the concentration to which they are exposed, but return to the sensitive state after one passage through Cm-free medium (Vining & Westlake, 1964). In cultures grown under Cm production conditions, resistance was correlated with Cm synthesis (Malik & Vining, 1970, 1972). The much higher resistance of antibiotic producing cultures than nonproducing cultures, was ascribed to an acquired resistance mechanism, probably a Cm efflux system, associated

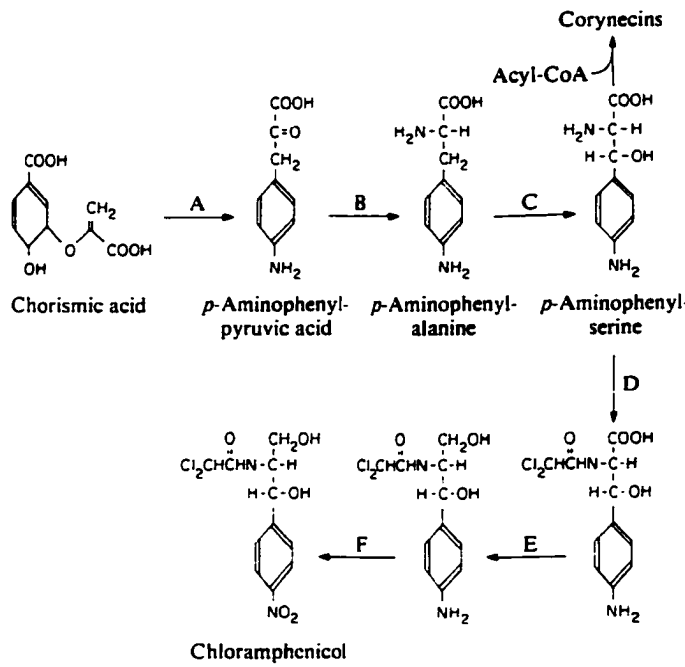


Figure 4. The overall Cm biosynthesis pathway (from Doull *et al.*, 1985). Reactions: A, arylamine synthase; B, aminotransferase; C, hydroxylase; D, dichloroacyl transferase; E, reductase; F, oxidase.

with Cm biosynthesis (Malik & Vining, 1972; Mosher *et al.*, 1995).

Acetylation of Cm catalyzed by chloramphenicol acetyltransferase (CAT) is the mechanism of resistance in most eubacteria (Shaw, 1983). The CAT product, 3'-O-acetyl-Cm, does not possess antibacterial activity. However, no CAT activity has been detected in *S. venezuelae* (Shaw & Hopwood, 1976, Nakano *et al.*, 1977), and thus, there should be another mechanism protecting a producing organism from inhibition by its own antibiotic (Vining & Westlake, 1984). Malik & Vining (1971) identified an intracellular Cm hydrolase that deactivates Cm by removing the dichloroacetyl group. Mycelia from nonproducing cultures converted Cm via *p*-nitrophenylserinol to N-acetyl-*p*-nitrophenylserinol, *p*-nitrobenzyl alcohol, and *p*-nitrobenzoic acid. The enzyme Cm-hydrolase has broad substrate specificity, and catalyzes removal of an acetyl or a dichloroacetyl group from the *p*-nitrophenylserinol moiety (Figure 5). Since the enzyme is expressed at equivalent levels in producing and nonproducing cultures, it was presumed not to be responsible for the resistance induced during Cm production (Malik & Vining, 1970; Vining & Stuttard, 1995).

Experiments in which the labeled Cm was added to cultures confirmed that metabolism of the supplement ceased when endogenous synthesis of the antibiotic reached 20-30 mg/liter. From this time the Cm added remained unchanged in the culture broth. It was postulated that active excretion associated with biosynthesis causes a change in cell permeability that prevents further uptake of the antibiotic. This ensures that the producing organism is not only protected from the toxic effects of its own metabolite, but is prevented from degrading it by reactions associated with the intracellular hydrolase (Malik & Vining, 1970, 1972). Induced resistance would be due

to the impermeability of the mycelium to Cm after the onset of production, and the action of Cm hydrolase inside the cell in scavenging any remaining intracellular antibiotic (Vining & Stuttard, 1995). Mosher and co-workers (1990) cloned a 6.5-kb chromosomal fragment from *S. venezuelae* that conferred Cm resistance to *Streptomyces lividans* M252, a Cm-sensitive strain. Two open reading frames (ORF1 and ORF2) were identified in the cloned 2.4-kb fragment of chromosomal DNA. The sequence of protein encoded by ORF1 resembled efflux proteins that confer antibiotic resistance by excluding Cm from the cytoplasm. Sequence analysis of ORF2 identified a kinase, and transformation of a suitable host with the cloned ORF2 gene confirmed the ability to convert Cm to Cm-3'-phosphate. The kinase therefore represented a potential mechanism for Cm resistance in *S. venezuelae*. Cm-resistant transformants carrying the entire 6.5-kb chromosomal fragment were also able to degrade Cm, implying the presence of a cluster of genes encoding the hydrolase and other degrading enzymes (Vining & Stuttard, 1995). Recently it was found that cultures of *S. venezuelae* produce 3'-O-acetylCm in addition to Cm (Groβ *et al.*, 2001). The acetyl derivative was identified by HPLC and NMR spectroscopy, and as part of my study on Cm biosynthesis, was synthesized chemically. The role of 3'-*O*-acetylCm in Cm biosynthesis, and in resistance, will be dealt with later in the thesis.

e. Regulatory mechanisms in Cm biosynthesis

Cm biosynthesis like that of other secondary metabolites is regulated by environmental factors determined by culture conditions. Production of the antibiotic is inversely correlated with the ability of carbon and nitrogen sources to support rapid growth. For *S. venezuelae* to start biosynthesis of Cm, the nitrogen source should be

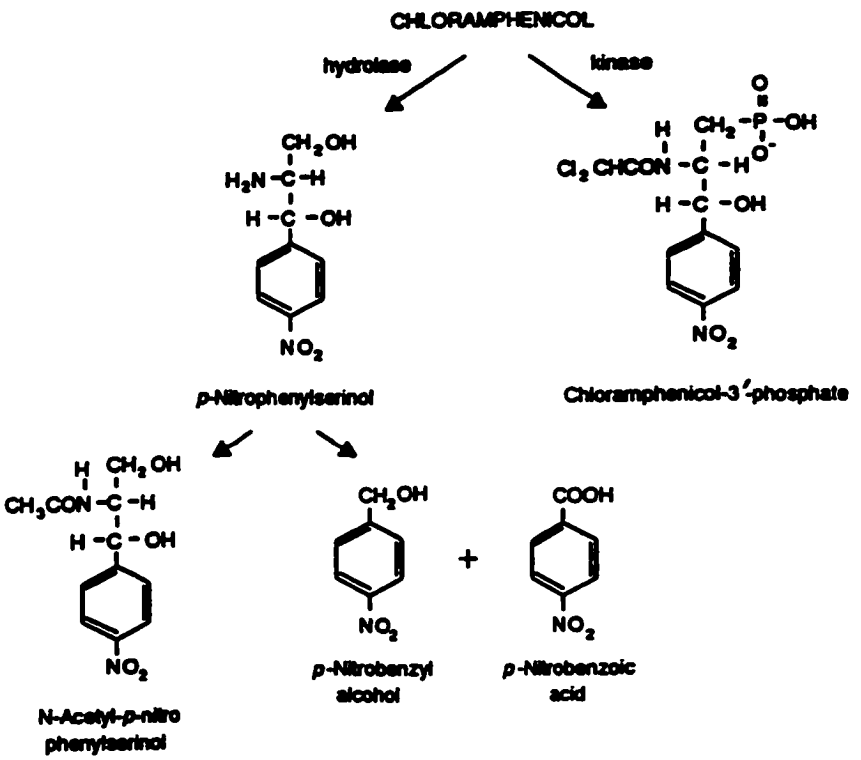


Figure 5. Degradation and modification of chloramphenicol initiated by Cm hydrolase and Cm-3'-kinase (Vining & Stuttard, 1995).

depleted but excess of the carbon source should be present to provide enough energy to extend enzyme activity into the stationary phase (Chatterjee & Vining, 1981; Chatterjee *et al.*, 1982; Shapiro & Vining, 1983). However, if glucose is the carbon source, the excess must be limited, to prevent carbon catabolite repression of the biosynthetic enzymes. Extensive investigation of the factors governing Cm biosynthesis showed that nitrogen depletion causes a change in growth rate that in turn will control the timing of antibiotic production. Unlike many other biosynthetic pathways, that for Cm is not sensitive to the concentration of phosphate in the medium (Vining & Stuttard, 1995).

In many bacteria the shikimate pathway branch point enzymes specific to tryptophan, tyrosine and phenylalanine biosynthesis are subject to end product control. However, in *S. venezuelae* aromatic amino acids, Cm, or shikimic acid pathway intermediates do not repress synthesis of deoxy-*arabino*-heptulosonate-7-phosphate (DAHP) synthase, the entry reaction to the shikimate pathway, nor do they inhibit this enzyme (Lowe & Westlake, 1971, 1972). Moreover, in *S. venezuelae*, Cm and its postulated intermediates have no inhibitory or repressive effects on chorismate mutase, prephenate dehydratase, or anthranilate synthetase activities (Jones & Westlake, 1974). Arylamine synthase activity, which converts chorismic acid and glutamine to *p*-aminophenylalanine and is correlated with Cm production, (Jones & Vining, 1976) was inhibited by 1 mM DL-*p*-aminophenylalanine, but Cm did not show any inhibitory effect. On the other hand, Cm and DL-*p*-aminophenylalanine added at equivalent concentrations to the cultures repressed arylamine synthase activity within the mycelium to a similar degree (Jones & Westlake, 1974). Consistent with the idea that arylamine synthase is the controlling point in biosynthesis of Cm, the activity of the

enzyme rose to a maximum during Cm production and then fell to a very low basal level. Repression of anthranilate synthase, the first enzyme of tryptophan biosynthesis, stimulated arylamine synthase activity in *S. venezuelae*, suggesting a channeling of chorismic acid between the two branch-point enzymes (Jones & Westlake, 1974; Francis *et al.*, 1978).

f. Chromosomal location of Cm biosynthesis genes

Akagawa *et al.* (1979) proposed that the genes for Cm biosynthesis in *S. venezuelae* were plasmid borne. However, elimination of the only plasmid (pUC3) detected in a Cm-producing strain failed to affect antibiotic production (Ahmed & Vining, 1983). Further evidence for location of Cm biosynthesis genes on the chromosome was provided by analyzing mutants blocked in chloramphenicol biosynthesis, and demonstrating linkage between genetic markers for antibiotic production and chromosomal genes for primary pathways (Doull *et al.*, 1986; Vats *et al.*, 1987).

g. Identification of genes associated with Cm biosynthesis

Recently several genes involved in early reactions of Cm biosynthesis have been identified, and their roles in the biosynthesis of the antibiotic have been confirmed by gene disruptions (Figure 6). Among them are the gene for a DAHP synthase (CmlE) required for conversion of phosphoenolpyruvate and erythrose-4-phosphate to DAHP, the gene for 4-amino-4-deoxychorismate synthase (CmlB; Brown *et al.*, 1996), the gene for a chorismate mutase (CmlD) presumed to be responsible for conversion of 4-amino-4-deoxychorismate to 4-amino-4-deoxyprephenate, and the gene for a prephenate

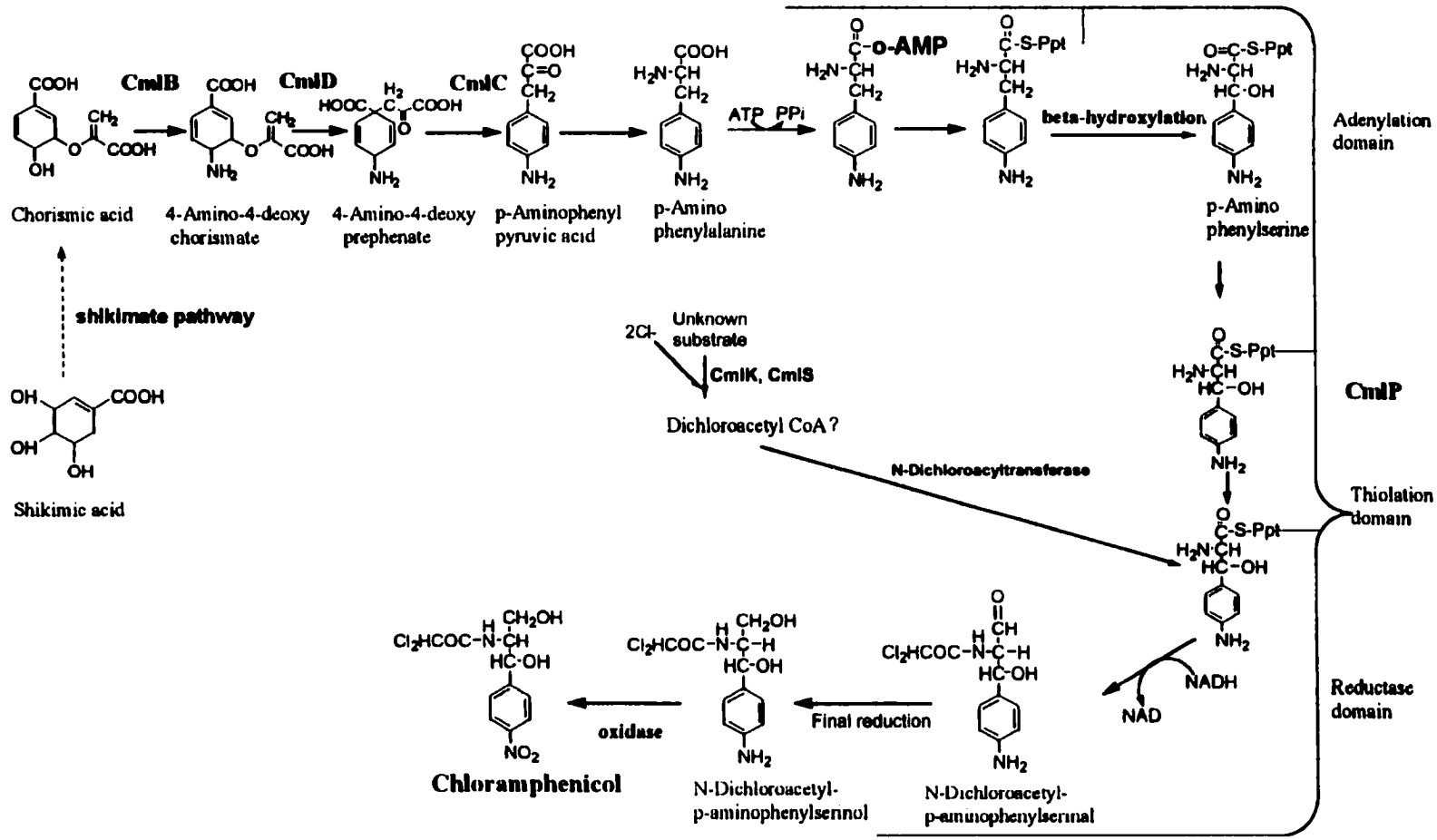


Figure 6. Cm biosynthesis pathway based on the identified genes and their function (modified from He *et al.*, 2001).

S-Ppt indicates the 4'-phosphopantetheinyl group of CmlP.

dehydrogenase (CmlC) believed to convert 4-amino-4-deoxyprephenate to *p*-aminophenylalanine. Other genes with proven roles in the Cm biosynthesis pathway include one encoding a nonribosomal peptide synthetase (CmlP), another encoding a short-chain oxidoreductase (CmlJ), and one (CmlH) with an unknown function. The Cm biosynthesis cluster also contains other genes for which functions have not yet been demonstrated experimentally, but which are predicted from analyses of their deduced amino acid sequences to have roles in the pathway. These include genes for an ABC membrane transporter (CmlI), and for a protein (CmlF) related to known Cm efflux pumps (He *et al.*, 2001), and the genes mentioned earlier that encode enzymes for detoxification of Cm in the producer (Mosher *et al.*, 1990).

IV. Biohalogenation

There is an impressive range of halometabolites in biological systems. Natural halogenated compounds are produced by a variety of terrestrial or marine organisms; more than 3000 have been isolated from sources such as bacteria, fungi, algae, higher plants, marine molluscs, insects and mammals, including humans. Among them are important chlorinated antibiotics such as chlortetracycline, vancomycin, and chloramphenicol. The bacteria producing these antibiotics can usually produce the brominated analogues if chloride is replaced with bromide in the fermentation medium (Neidleman, 1975; van Pée, 1996).

a. Halogenating enzymes

The biohalogenation reaction for production of halometabolites has been the point of interest in many studies aimed at determining the specific substrate, the enzyme responsible, and the mechanism governing the reaction. Despite the large number of

halometabolites isolated, only a few types of halogenating enzymes have been discovered: haloperoxidases (HPO), perhydrolases and halogenases are the three main classes of halogenating enzymes identified to date (van Pée, 1996; Hohaus *et al.*, 1997); in a fourth class is *S*-adenosylmethionine methyltransferase, responsible for biosynthesis of methyl chloride by organisms in the environment. This enzyme catalyzes the methylation of chloride, bromide, and iodide ions without requiring any prosthetic group or cofactor; the methyl group is donated by *S*-adenosyl methionine (Wuosmaa & Hager, 1990). Only the first two types of halogenating enzymes are described in more detail here.

i. Haloperoxidases

Haloperoxidases are a widely distributed family of oxidoreductases (Neidleman & Geigert, 1986; van Pée, 1996) catalyzing the formation of carbon-halogen bonds in the presence of hydrogen peroxide, halide ions, and a suitable organic substrate (Thomas *et al.*, 1970; Libby *et al.*, 1982; Ramakrishnan *et al.*, 1983). Haloperoxidases oxidize halide ions (X^-) with H_2O_2 resulting in the electrophile (X^+) that can react with and halogenate organic substrates.

According to the halide ion they can oxidize, the haloperoxidases are designated chloroperoxidases (Cl^-, Br^-, I^-), bromoperoxidases (Br^-, I^-), or iodoperoxidases (I^-).

Haloperoxidases can be classified into two groups; 1) heme-dependent, 2) nonheme, metal-dependent (Vn^{+5}) (Vilter, 1984; van Pée, 1996).

1. Heme-dependent haloperoxidases

Heme-dependent haloperoxidases contain a heme group, which transfers an electron from the halide ion to the peroxide. The mold *Caldariomyces fumago* produces a heme-dependent chloroperoxidase that catalyzes sequential di-halogenation at C-2 of the 1,3-cyclopentanedione producing the antibiotic caldariomycin (Hager *et al.*, 1966; Morris

& Hager, 1966). However, studies in bacterial species using molecular genetic approaches, such as gene disruption, showed that none of the heme-containing bacterial haloperoxidases is responsible for catalyzing chlorination of substrates in biosynthesis of bacterial antibiotics (van Pée, 1996). A bromoperoxidase-catalase identified in *S. venezuelae* ISP5230 showed similarity to the heme-containing bromoperoxidases (Knoch *et al.*, 1989). Disruption of its gene indicated that the bromoperoxidase-catalase was not required for biosynthesis of Cm (Facey *et al.*, 1996).

2. Metal-dependent non-heme haloperoxidases

These haloperoxidases do not contain heme; instead electron transfer from the halide to the peroxide proceeds via a metal cofactor (e.g., vanadium). They catalyze the halogenation reaction by forming a halogenating agent, which is known to be the relevant hypohalite (HOX) (e.g., hypochlorite in chlorination). This agent results in unspecific halogenation of organic compounds susceptible to electrophilic attack (Neidleman & Geigert, 1986; van Pée, 1996). However, halogenating enzymes in living organisms are not believed to halogenate metabolites by forming free hypohalite as an intermediate because then this could unspecifically halogenate a wide range of substrates (Prutz, 1996). HPO from *Ascophyllum nodosum* is an example of the metal-dependent nonheme haloperoxidases (Vilter, 1984).

ii. Perhydrolases

The enzymes with halogenation activity from *Streptomyces aureofaciens* Tu24 (CPO-T) (van Pée, 1988) and *Pseudomonas pyrrocinia* (CPO-P) (Wiesner *et al.*, 1988) catalyze in vitro chlorination reactions, and do not contain a prosthetic group or any metal ions. They contain a catalytic triad consisting of a serine, a histidine, and an

aspartate similar to that of α/β hydrolases (Hecht *et al.*, 1994), and utilize H_2O_2 to hydrolyze an acyl-enzyme intermediate to its peracid derivative. This peracid oxidizes halide ions and generates hypohalous acid which is the actual halogenating agent (Hoffman *et al.*, 1998)

Investigation of substrate specificities and reaction mechanisms have shown that enzymatic halogenation with bacterial haloperoxidases or perhydrolases is no more specific than chemical halogenation (Ramakrishnan *et al.*, 1983; Bongs & van Pée, 1994). The lack of substrate specificity for these enzymes ruled out involvement of this mechanism in the biosynthesis of halometabolites like 7-chlortetracycline or pyrrolnitrin, and implicated another type of enzyme in the biosynthesis of chlorinated secondary metabolites (van Pée, 1996).

iii. Halogenases

The conclusion that haloperoxidases are not responsible for specific halogenation of biosynthetic intermediates in secondary metabolic pathways, and their inability to oxidize fluoride encouraged a search for alternative halogenating agents. One major obstacle to finding the actual halogenating enzyme was lack of knowledge about the structure of the natural substrates. Also it had to be proved that the "natural" substrate is halogenated *in vivo*. van Pée and colleagues were able to isolate L-7-chlorotryptophan from a mutant of *P. fluorescens* blocked in the second step of pyrrolnitrin biosynthesis. In the expectation that L-tryptophan was the first substrate, they set up experiments to detect the enzyme converting L-tryptophan to L-7-chlorotryptophan. These experiments established that chlorination would take place only if NADH was added to a cell-free extract. The halogenating enzyme from

P. fluorescens was finally purified and characterized, and designated as a halogenase (Keller *et al.*, 2000).

The N-terminal regions of halogenases contain a short segment with high similarity to the NAD-binding domain of NADH-dependent monooxygenases such as a dehydrogenase from *Haemophilus influenzae*, and a thioredoxin reductase from *S. clavuligerus* (van Pée *et al.*, 2000). The NADH requirement for chlorination activity could be explained by the chlorination mechanism proposed for reactions catalyzed by halogenases. In this mechanism FAD is first reduced by NADH and flavin reductase to $FADH_2$, which then reacts with oxygen forming flavin hydroperoxide. The NADH/FAD-dependent halogenases use the flavin hydroperoxide to oxidize the organic substrate to an epoxide. Nucleophilic attack by a chloride ion, and specific removal of water results in the chlorinated product (Figure 7). In an enzyme-catalyzed reaction, the requirements for stereo- and regio-specificity are more likely to be met if the halogenation reaction is initiated by oxidizing the substrate, as proposed for halogenases (Hohaus *et al.*, 1997), than by oxidizing the halide ion as in the haloperoxidase mechanism (Bongs & van Pée, 1994; van Pée, 1996). Several halogenase genes have now been isolated from a variety of organisms. Of these, *pltA*, *pltD*, and *pltM* are involved in pyoluteorin biosynthesis (Nowak-Thompson *et al.*, 1999) while *prnC* is involved in the chlorinating reaction for pyrrolnitrin, and *chl* has a role in the biosynthesis of chlortetracycline (Dairi *et al.*, 1995). The results from these investigations favor the $FADH_2$-dependent halogenases over haloperoxidases as agents for halometabolite biosynthesis.

Specific halogenating enzymes present many possibilities for the production of useful chemicals with commercial application, either as end products or as intermediates. For example, a specific halogenase with high stereoselectivity might be

useful for the kinetic resolution of a racemic substrate by preferential halogenation of one enantiomer.

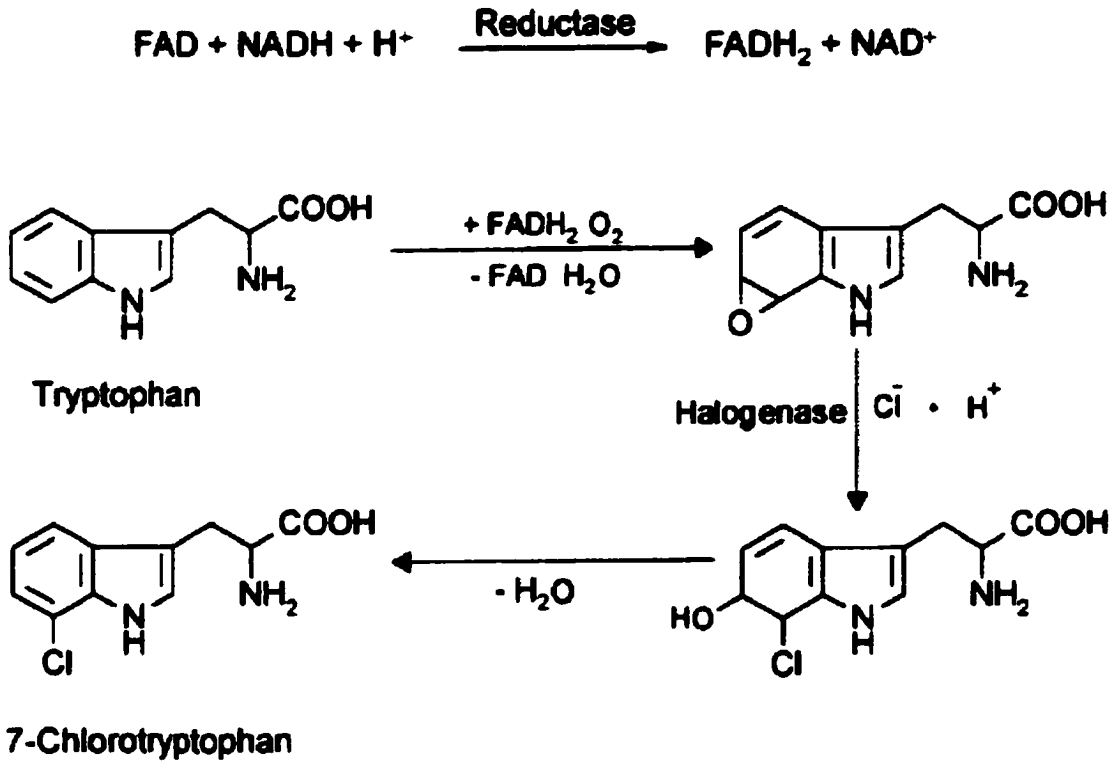


Figure 7. Mechanism of chlorination of tryptophan catalyzed by a halogenase (van Pée *et al.*, 2000).

Materials and Methods

I. Bacterial strains and plasmids

All bacteria and plasmids used in this study are listed in Table 1.

II. Chemicals and biochemicals

Biotechnology grade agarose A was from *BioBasic Inc*, Scarborough, ON, and *tris*-(hydroxymethyl)aminomethane (Tris) from *ICN Biochemicals*, Cleveland, OH. Malt extract, Bacto-Agar, Bacto-Tryptone, and yeast extract were from *Difco Laboratories*, Detroit, MI. Nutrient broth was from *Oxoid Limited*, Basingstoke, UK. Ampicillin (Ap), apramycin (Am) kanamycin (Km), thiostrepton (Ts), and isopropyl-β-D-thiogalactopyranoside (IPTG) were from *Sigma-Aldrich Canada, Ltd.*, Oakville, ON. 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) was from *Diagnostic Chemicals Limited*, Charlottetown, PEI. Sodium dodecyl sulfate (SDS) was a specially purified grade from *B.D.H.*, Toronto, ON. T4 DNA ligase and calf intestinal alkaline phosphatase (CIAP) were purchased from *MBI Fermentas*, Flamborough, ON. Ready To Go DNA Labeling Beads, [α-^{32}P] dCTP, and *Taq* DNA polymerase were purchased from *Pharmacia Biotech*, Baie d'Urfe, QC. Restriction endonucleases used in this study were from, *Pharmacia Biotech*, Baie d'Urfe, QC, or *MBI Fermentas*.

Table 1. Bacterial strains, plasmids, and phages

Strains	Genotype/phenotype	Source/reference
Streptomyces venezuelae		
ISP5230	Wild-type	Doull *et al.*, 1985
cml-2	ISP5230 mutant blocked in chlorination	Doull *et al.*, 1985
VS1101	ISP5230 with ORF11 disrupted by Am^R gene (ISP5230 transformed with pJV527)	This study
VS1102	ISP5230 with ORF12 disrupted by Am^R gene (ISP5230 transformed with pJV508)	This study
VS1103	ISP5230 with ORF13 disrupted by Am^R gene (ISP5230 transformed with pJV518)	This study
VS1104	*cml-2* transformed with pJV507 (containing ORF12 & ORF13)	This study
VS1105	*cml-2* transformed with pJV528 (containing ORF13)	This study
VS1106	*cml-2* transformed with pJV526 (containing ORFs11, 12 and 13)	This study
Escherichia coli		
DH5αF′ IQ	F′/*endA1 hsdR17* (rK- mK+) *supE44 thi-1 recA1 gyr*(Nalr) *relA1* (*lac*(*ZYA-argF*) U169 *deoR* (ϕ80*lac*ZΔM15)	BRL
TOP10F′	F′/*endA1 nupG mcrA* Δ(*mrr-hsdRMS-mrcBC*) ϕ80*lac*ZΔM15 Δ*lac*X74 *deoR recA1 ara*D139 Δ(ara-leu)7697 *galU galK**rpsL* (Str^r)	Invitrogen
ET12567(pUZ8002)	*dam⁻ dcm⁻ hsdM⁻*	Paget, JII[a]

Table 1 continued

Plasmids		
pUC18/19	*ap*^{r}, *LacZ*	Pharmacia
pBluescript II SK+ (pSK+)	phagemid *ap*^{r}, *lacZ*	Stratagene
pJV326	*ts*^{r}, *ap*^{r}, *lacZ*, *oriT*	He *et al.*, 2001
pCR2.1TOPO	*ap*^{r}, *kan*^{r}, *lacZ*	Invitrogen
pUC120A	*Nco*I-*Nco*I cassette containing Am^R gene	Paradkar and Jensen, 1995
pJV501	pCR2.1TOPO containing 290-bp PCR product from *S. venezuelae* (using MPF1 and MPR2)	This study
MP30	Lambda GEM-11 with ~22 kb insert from *S. venezuelae* chromosome (hybridizes with the halogenase PCR product)	This study
MP60	Lambda GEM-11with ~23 kb insert from *S. venezuelae* chromosome (hybridizes with the halogenase PCR product)	This study
pJV502	pSK+ with 6.2-kb *Sac*I-*Sac*I insert from MP30 DNA	This study
pJV503	pSK+ with 6.2-kb *Sac*I-*Sac*I insert from MP60 DNA	This study
pJV504	pUC18 with 1.5-kb *Sal*I-*Sal*I insert from pJV502 (hybridizes with PCR probe)	This study
pJV506	pJV502 with a 2.4-kb *Not*I-*Not*I fragment deleted and residual linear DNA self-ligated; contains 3.8-kb *Not*I-*Sac*I chromosomal fragment	This study
pJV507	pJV326 containing the 3.8-kb *Sac*I-*Xba*I fragment of pJV506 including ORF12 & ORF13	This study
pJV508	pJV507 containing ORF12 disrupted with Am^R gene at *Stu*I site	This study

Table 1 continued

pJV513	pUC18 containing the 1.9-kb *Sac*I-*Sac*I chromosomal fragment from MP30 that lies downstream of the 3.8-kb *Not*I-*Sac*I fragment	This study
pJV516	pJV507 with 1.75-kb *Stu*I-*Not*I deleted	This study
pJV518	pJV516 containing ORF13 disrupted with Am^R gene at *Nco*I site	This study
pJV357	7.5-kb *Bam*HI-*Bam*HI fragment of *S. venezuelae* DNA containing genes upstream of *pabAB*	He *et al.*, 2001
pJV526	pJV326 containing the 6.2-kb *Sac*I-*Sac*I fragment of *S. venezuelae* DNA obtained from pJV502 (containing ORFs 11, 12 & 13)	This study
pJV527	pJV526 containing ORF11 disrupted with Am^R gene at *Not*I site	This study
pJV528	pJV526 with 2.7-kb *Stu*I-*Eco*RV deleted. (containing ORF13)	This study

a. John Innes Institute, Norwich, U.K.

III. Media

Glucose-isoleucine (Cm production) medium contained the following components:

L-Isoleucine	3.75 g
K_2HPO_4	5.25 g
KH_2PO_4	2.25 g
$MgSO_4.7H_2O$	0.10 g
1% / 1% Stock solution of $NaCl/CaCl_2$	4.5 ml
0.2% - $FeSO_4$ stock solution	2.25 ml
Trace elements stock solution	2.25 ml

Water was added to a final volume of 450 ml, then 45-ml portions were dispensed into 250-ml flasks and autoclaved. When the solution was cold, 5 ml of separately autoclaved 32% stock solution of glucose was aseptically added to each flask. The medium was inoculated with 1% (0.5 ml) seed culture of *S. venezuelae*.

GNY medium (Malik & Vining, 1970) was used to grow vegetative mycelium to inoculate Cm-production (glucose-isoleucine) medium; it contained:

Glycerol	20 g
Nutrient broth	8 g
Yeast extract	3 g
K_2HPO_4	5 g
Distilled water	qs 1000 ml

GNY medium was also used to grow *S. venezuelae* for preparation of cell-free extracts in experiments detecting N-acetyltransferase activity. *p*-NPS was added to the culture at a concentration of 0.8 mM.

LB-agar (Sambrook *et al.*, 1989), used to grow cultures of *E. coli* at 37 °C, contained:

Bacto-Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
Bacto-Agar	16 g
Distilled water	qs 1000 ml

If necessary the culture medium was supplemented with 100 µg/ml of ampicillin or 50 µg/ml of apramycin. *E. coli* transformants were selected on LB-agar plates containing ampicillin or ampicillin+apramycin. X-gal and IPTG were used to identify transformants by the "blue-white" screening procedure (Sambrook *et al.*, 1989); colonies with recombinant plasmids produced white colonies after incubation at 37 °C overnight. X-gal agar was prepared by spreading 40 µL from a stock solution (20 mg/ml) of X-gal, and 4 µL from a stock solution (200 mg/l) of IPTG on the surface of each LB-agar plate.

LB-broth was used to grow liquid cultures of *E. coli*. When necessary, LB-broth was supplemented with 100 µg/ml of ampicillin or 50 µg/ml of apramycin to select for growth of cells containing the plasmid of interest. Cultures were shaken overnight at 37 °C.

MYM agar (Stuttard, 1982) was used to grow streptomycete cultures for sporulation, to prepare spore stocks, and when supplemented with apramycin at 50 µg/ml, to grow transconjugants. It contained:

Maltose	4.0 g
Yeast extract	4.0 g

Malt extract	10 g
Agar	15 g
Distilled water	qs 1000 ml

The medium was adjusted to pH 7.0 with NaOH, and with agar omitted was called MYM medium. To determine thiostrepton sensitivity in apramycin-resistant *S. venezuelae* transconjugants, apramycin and thiostrepton were added to MYM agar at final concentrations of 50 μg/ml and 25 μg/ml, respectively.

MYEME medium (Hopwood *et al.*, 1985a) was used to grow all streptomycetes (except *S. armentosus* UC2862) for isolation of genomic DNA. It consisted of:

Yeast extract	3.0 g
Bacto-Peptone	5.0 g
Malt extract	3.0 g
Glucose	10 g
Sucrose	103 g
Distilled water	qs 1000 ml

The medium, adjusted to pH 7.5, was dispensed as 25-ml aliquots in 250-ml Erlenmeyer flasks and autoclaved. Prior to use, 0.05 ml of 2.5 M $MgCl_2$ (5 mM final) and 0.625 ml of 20% w/v glycine (0.5% final) were added to each flask.

MS agar (M. Paget, John Innes Institute, Norwich, personal communication) was used for facilitating conjugation between plasmid-containing strains of *E. coli* and *S. venezuelae* ISP5230. It contained:

Mannitol	20 g

Soy bean flour	20 g
Agar	16 g
Water	qs 1000 ml

The medium was adjusted to pH 7.0, and 250-ml portions were transferred to 1-L Erlenmeyer flasks and autoclaved. When the medium had cooled to about 50 °C, 1 ml of sterile 2.5 M $MgCl_2$ was added to each flask (10 mM final) and the medium was dispensed into Petri plates.

SOB medium was used to grow *E. coli* DH5αF′IQ for preparing competent cells. This medium consisted of:

Bacto-Tryptone	20 g
Yeast extract	5.0 g
NaCl	0.5 g
$MgSO_4$	2.4 g
KCl	0.186 g
Distilled water	qs 1000 ml

SOC medium was used to grow *E. coli* transformants to express the antibiotic resistance gene of the plasmid after transformation. It was prepared by mixing as sterile solutions:

SOB medium	100 ml
1 M glucose	2 ml

Soft agar was used for bioassay of chloramphenicol, and contained 0.5% agar in GNY medium.

The trace elements stock solution used for preparation of glucose-isoleucine medium contained per 100 ml:

$ZnSO_4.7H_2O$	400 mg
$CuSO_4.5H_2O$	18 mg
$MnCl_2 \cdot 4H_2O$	2.4 mg
Ammonium molybdate	1.7 mg
Boric acid	2.6 mg
Distilled water	qs 100 ml

TO agar was used to grow *Streptomyces venezuelae* cultures for sporulation. It contained:

Tomato paste	20 g
Oatmeal	20 g
Agar	15 g
Distilled water	qs 1000 ml

When necessary, apramycin at a final concentration of 50 μg/ml was added to the TO medium for preparation of spores from the knockout strains.

IV. Stock solutions and buffers

Antibiotic stock solutions and their solvents:

Antibiotic	Stock concentration (solvent)	Final concentration (μg/ml)
Ampicillin	100 mg/ml (water)	50-100
Apramycin	50 mg/ml (water)	50
Chloramphenicol	25 mg/ml (ethanol)	25
Kanamycin	100 mg/ml (water)	50
Thiostrepton	50 mg/ml (DMSO)	50
Nalidixic acid	5 mg/ml (water) *	500

* A few drops of 1 M NaOH was added to dissolve nalidixic acid.

Buffer E7 was used for extraction of total proteins in cell-free preparations. It was adjusted to pH 7.0 and contained per 500 ml:

KCl	18.64 g
K_2HPO_4	5.25 g
KH_2PO_4	2.25 g

and

EDTA and 2-mercaptoethanol at concentrations of 0.1 mM and 6 mM, respectively.

Buffer D8 was used to equilibrate the DEAE chromatography column; it was prepared by diluting buffer E7 20-fold with water and adjusting the pH to 8.0.

Chloroform:isoamyl alcohol: a 24:1 mixture of chloroform:isoamyl alcohol was used for purification of DNA samples.

DNase was used in phage DNA purification procedures. A 1-mg/ml stock solution was prepared in 0.02 M sodium acetate, pH 5.2, and was stored at -20 °C.

Ethidium bromide solution: A 10 mg/ml solution was prepared in water, and was wrapped in aluminum foil and stored at room temperature.

0.5 M EDTA was prepared by adding 186 g EDTA.$2H_2O$ to 800 ml water and stirring vigorously. The pH was adjusted to 8.0 with NaOH, and the solution was sterilized by autoclaving.

Hybridizations solutions:

Solution	Components
20X SSC	3 M NaCl (175.3 g/l) and 0.3 M tri-sodium citrate (88.2 g/l), adjusted to pH 7.0 with NaOH. It was sterilized by autoclaving.
Depurination solution	0.25 M HCl
Denaturation solution	1.5 M NaCl and 0.5 M NaOH
Neutralization solution	1.5 M NaCl and 0.5 M Tris-HCl (pH 7.4)
Pre-hybridization solution	6X SSC, 5X Denhardt's reagent, 0.5% (w/v) SDS and 100 µg/ml of denatured, sheared herring sperm DNA
100X Denhardt's reagent	contained 2%(w/v) BSA, 2%(w/v) Ficoll (type 400 Pharmacia), and 2%(w/v) polyvinylpyrrolidone.

IPTG (isopropyl-β-D-thiogalactopyranoside) was dissolved at 200 mg/ml in water and filter sterilized with a 0.22-micron membrane.

λ diluent was used in preparation of PEG-NaCl solution for precipitation of phage particles. It contained:

Tris.HCl (pH 7.5)	10 mM
$MgSO_4$	10 mM

Lysozyme buffer (pH 8.0) was used for lysis of streptomycete mycelia in DNA isolation procedures. It consisted of:

Sucrose	0.3 M
Tris-HCl	25 mM
EDTA	25 mM

Lysozyme was added to a final concentration of 2 mg/ml immediately before use.

Phenol:chloroform was used to denature and extract protein during DNA isolation. It was prepared (Sambrook *et. al.*, 1989) by mixing equal amounts of phenol and chloroform in a brown glass bottle and equilibrating with 0.1 M Tris (pH 7.6). The upper phase (aqueous layer) was decanted, and the lower phase was covered with a 2-cm layer of 0.1 M Tris-HCl (pH 8.0) and capped. The mixture was stored at 4 °C in dark glass bottles.

Pronase: A 20 mg/ml stock solution was prepared in 10 mM Tris-HCl, pH 7.5 containing 10 mM NaCl. The solution was self-digested by incubation at 37 °C for 1 h. Aliquots were stored at –20 °C in tightly-capped 1.5-ml microtubes.

RNaseA: DNase-free RNase (*Pharmacia Biotech*) was dissolved to a final concentration of 10 mg/ml in 10 mM Tris.HCl (pH 7.5) containing 15 mM NaCl. The solution was heated in a boiling water bath for 15 min, and 1-ml aliquots were stored at –20 °C. It was used at a final concentration of 20 µg/ml.

S1 nuclease stop buffer (Sambrook *et al.*, 1989) was used at about 1/10 volume of the reaction mix to stop the S1 nuclease activity after blunt end formation. It contained:

Ammonium acetate	4 M
EDTA (pH 8.0)	50 mM
tRNA	50 µg/ml

SDS solution was prepared at 10% (w/v) in water and the pH was adjusted to 7.2 with a few drops of concentrated HCl.

SM buffer was used for phage elution from agar plugs (Sambrook *et al.*, 1989). It contained:

Tris-HCl (pH 7.5)	50 mM
$MgSO_4.7H_2O$	8 mM
NaCl	100 mM
Gelatin	0.01% (added from a 2% solution)

The mixture was autoclaved and stored in 50-ml aliquots.

Three solutions were used for plasmid DNA isolation by the alkaline lysis method (Sambrook *et al.*, 1989):

Solution I:

Glucose	50 mM
Tris-HCl (pH 8.0)	25 mM
EDTA (pH 8.0)	10 mM

Solution II:

NaOH	0.3 M
SDS	2.0%

Solution III:

5 M Potassium acetate	60.0 ml
Glacial acetic acid	11.5 ml
Distilled water	28.5 ml

TAE buffer (Sambrook *et al.*, 1989) was used for agarose gel electrophoresis of DNA samples. The stock solution was prepared at 50X concentration and adjusted to pH 8.5. It consisted of:

Tris-base	242.0 g
Glacial acetic acid	57.1 ml
0.5 M EDTA (pH 8.0)	100 ml
Water	qs 1000 ml

The stock solution was diluted to1X for agarose gel electrophoresis.

TE buffer, used to dissolve DNA, consisted of:

Tris-HCl	10 mM

EDTA	1 mM

The pH was adjusted to 8.0 with concentrated hydrochloric acid.

Top agarose was used for preparation of phage plaques, and contained:

Nutrient broth	0.8%
Agarose	0.7%

The pH was adjusted to 7.0 before autoclaving.

Wash solution was used to remove *E. coli* cell debris from phage-lift membranes before hybridization procedures. It contained:

20x SSC	25 ml
10% SDS	5 ml
0.5 M EDTA	0.2 ml
Water	qs 100 ml

X-gal was used in X-gal agar plates to detect recombinant *E. coli* transformants. A 20 mg/ml solution of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was prepared in dimethylformamide.

V. Culture conditions

a. Cultures for glycerol stock suspensions

i. *E. coli*

Stock suspensions for long-term storage were prepared by mixing 0.5 ml of sterile 50% glycerol with 0.5 ml of an overnight culture in LB-broth, and stored at –70 °C.

ii. *Streptomyces* spores (Hopwood *et al.*, 1985a)

Spores from a sporulating-culture of *Streptomyces* grown in a Petri plate were collected in 4 ml of sterile distilled water by gently rubbing the surface with a sterile glass spreader. The spore suspension was then filtered through sterile non-absorbent cotton wool to remove mycelial fragments and agar. Spores were pelleted by centrifuging at 8,000 rpm for 10 min, washed twice with water and after centrifuging were resuspended in 25% (v/v) glycerol and stored at –70 °C in 60-µl aliquots.

b. Cultures for genomic DNA extraction from streptomycetes

MYEME medium containing 5 mM $MgCl_2$ and 0.5% glycine was used to grow all *Streptomyces* strains except *S. armentosus* UC2862, for which 25 ml MYM with the above supplements was used. The mycelium produced in a 250-ml flask was harvested after incubation at 30 °C for 48 h on a rotary shaker at 220 rpm.

c. Cultures for plasmid DNA isolation

Strains of *E. coli* harboring the plasmid were grown in 3 ml LB-broth containing 100 µg/ml ampicillin or both ampicillin and apramycin (50 µg/ml) if the *E. coli* had been transformed with a disruptant plasmid. Cells were harvested after overnight incubation at 37 °C on a rotary shaker at 250 rpm.

d. Propagation of *S. venezuelae* phage library

A sample of the Lambda GEM-11 library with an original concentration of 90,000 pfu/µl, was diluted to 10^{-4} with SM buffer, and 100 µl of the diluted phage was mixed with 100 µl of *E. coli* LE392 suspension, vortexed and then incubated at 37 °C for 30 min to allow adsorption of phage particles to the bacteria. Then 3.5 ml molten top agarose (at

47 °C) was added, vortexed gently and poured evenly on the surface of LB agar in a plate that had been pre-warmed to 37 °C. Six plates were prepared and incubated for about 14 h at 37 °C to detect the phage plaques.

The *E. coli* LE392 suspension used as the host for phage propagation was grown at 37 °C in 50 ml LB broth containing 0.2% maltose for 12 h. The cells were centrifuged at 4 °C. The pellet was resuspended in 20 ml of 0.01% $MgSO_4$ solution and kept at 4 °C; it was used within two weeks.

VI. Feeding of labeled substrates and purification of Cm

The NAC esters and several other labeled compounds were separately dissolved in 15 ml water (total 0.75 mmol of each compound), filter sterilized and administered in three equal portions to ten 250 ml flasks (0.5 ml per flask) containing 50 ml culture at 3, 4, and 5 days after inoculation. Cultures were harvested on day 7, and chloramphenicol was extracted twice, with half volumes of ethyl acetate. The combined extracts were back-washed with 20 ml water. Cm was purified by reverse-phase column chromatography (Mega Bond Elut C_{18} column; *Varian*), in which the column was equilibrated with 10% methanol and the sample was applied in this solvent. A gradient from 10% to 30% methanol in water was used to elute Cm and fractions were collected. Those containing Cm were pooled, evaporated under vacuum and was kept in a vial under nitrogen. Isotope incorporations were measured by ^{13}C- NMR analysis of the purified Cm samples.

VII. Measurement of Cm concentration in cultures

Production of Cm by *Streptomyces venezuelae* ISP5230 and mutant strains obtained after inactivation of specific genes was determined as follows: A seed culture of the

strain was prepared by adding 50 µl of a spore suspension to 10 ml of GNY medium and incubating the culture on a rotary shaker (220 rpm) at 30 °C for 24 h. The mycelium was collected by centrifuging 10 ml of the culture in 15 ml sterile plastic tubes at 8,000 rpm for 10 min. The supernatant was discarded and pellet was resuspended in 10 ml sterile saline solution (0.9% NaCl). A 0.5 ml portion of this suspension was used to inoculate 50 ml of glucose-isoleucine medium (1% v/v). The medium used for transconjugants contained apramycin at 50 µg/ml. After incubation at 30 °C for 7 days, a 5 ml sample was clarified and extracted with an equal volume of ethyl acetate. The extract was air dried and the residue was dissolved in 0.5 ml 25% (v/v) methanol; a sample was analyzed by high-performance liquid chromatography (HPLC) with a UV detector at 273 nm, a 10 x 0.46-cm reverse-phase column (CSC-Sil 80A/ODS2; 5 µm) and stepped linear gradients from 100% water to 50% (v/v) methanol (4 min), then 50% to 100% methanol (2 min), followed by 100% methanol for 1 min, and a gradient returning to water (1 min). A solution of Cm (100 µg/ml) prepared in 25% methanol was used as standard in HPLC analysis of the culture extract. Authentic solutions of corynecins (I, II, III, IV) were co-injected with samples obtained from transconjugants for identification of unknown HPLC peaks.

VIII. Acetylation of Cm (Groβ *et al.*, 2001)

A solution of Cm (162 mg) prepared in 1 ml pyridine containing 10 mg 4-(dimethylamino)pyridine (DMAP) was cooled to 0 °C, and acetic anhydride (0.047 ml) was added dropwise with stirring. After 20 min stirring at 0 °C, ice-cold water (3 ml) was added, and the mixture was extracted three times with 20 ml ethyl acetate. The extract was dried over anhydrous $MgSO_4$, and evaporated under vacuum to a brown oil.

IX. Transformation

a. Preparation of competent cells (Sambrook *et al.*, 1989)

E. coli DH5αF' IQ was grown in 2 ml LB-broth by incubating at 37 °C overnight with shaking. From the overnight culture 0.1 ml was added to 25 ml of fresh LB or SOB medium in 250-ml Erlenmeyer flasks, and shaken for 2-3 h at 37 °C. The flask containing the *E. coli* cells was then cooled on ice for 10 min, and the culture was transferred to two pre-cooled centrifuge tubes to pellet the cells by centrifugation (8,000 rpm for 10 min at 4 °C). The supernatant fluid was completely removed, and the pellet, resuspended in 10 ml of cold 0.1 M $CaCl_2$, was kept on ice for 30 min. After centrifugation for 10 min at 4 °C, the cell pellet was collected, resuspended in 2 ml of 0.1 M $CaCl_2$, and cooled on ice for 10 min. After adding an equal volume of 50% glycerol, the competent cells were stored at –70 °C in 150-µL aliquots.

b. Transformation of competent cells

To a microtube containing 150 µL of *E. coli* competent cell suspension kept on ice, 1 µL (0.1-0.2 µg) of plasmid DNA was added and mixed gently by pipeting. The microtube was stored on ice for 30 min, then heated for 90 s in a water bath at 42 °C, after which it was transferred to an ice bath and chilled for 2-5 min. Either LB-broth or SOC broth (0.85) ml was added to the tube and the mixture was shaken at 37 °C for 60 min. The cells were pelleted by brief centrifugation, resuspended in the remaining fluid, and spread on LB-agar containing the appropriate antibiotic and X-gal:IPTG. After incubation at 37 °C for 12-24 h white colonies were screened to detect the plasmid of interest.

c. Conjugal DNA transfer between *E. coli* and *Streptomyces*

The conjugal vector pJV326, which contains the ampicillin and thiostrepton resistance genes, and is a shuttle vector for *E. coli* and streptomycetes, was used to subclone insertionally inactivated ORFs for conjugal disruption. The ORF of interest had been inactivated by inserting an apramycin resistance (Am^R) gene obtained either from pUC120A by digestion with *Nco*I (Paradkar & Jensen, 1995) or from the multi-cloning site of pJV225 (Chang, 1999). A fragment containing the disrupted ORF was ligated into pJV326 and the resulting conjugal disruption plasmid was introduced into *E. coli* ET12567 (pUZ8002). Transformants were selected on LB agar supplemented with Cm (25 μg/ml), Am (50 μg/ml), Km (50 μg/ml) and Ts (25 μg/ml), and the plasmid they contained was conjugally transferred to *S. venezuelae* ISP5230 for allele exchange. Culture conditions described by Flett *et al.* (1997) and Mazodier *et al.* (1989) were used. A single colony of *E. coli* ET12567 (pUZ8002) containing the disruption plasmid was used to inoculate 2 ml of LB-broth containing Cm, apramycin, and Km. The culture was shaken at 37 °C overnight, and a portion (0.1 ml) was used to inoculate 10 ml of fresh LB-broth (containing Cm, Km, and Am). After incubation as above to an OD_{600} of 0.4-0.6, cells were pelleted (8,000 rpm for 10 min) and washed twice with an equal volume of LB-broth. The pellet was resuspended in 1 ml of LB broth; then 0.5 ml was mixed in a microcentrifuge tube with approximately 10^8 spores (120-150 μl of a stock suspension) of *S. venezuelae* ISP5230 that had been heat-shocked in 0.5 ml of 2x YT at 50 °C for 10 min, and allowed to cool to room temperature. The mixture was centrifuged briefly and the pellet, resuspended in 0.3 ml of LB-broth, was spread on MS agar containing 10 mM $MgCl_2$. After incubation at 30 °C for 10-12 h, the agar was overlaid with 1 ml of sterile

distilled water containing 0.5 mg nalidixic acid and 1 mg apramycin. The plate was dried under laminar air flow and incubated at 30 °C until sporulation occurred (3-4 days). Transconjugants of *S. venezuelae* were patched from MS agar to MYM agar containing apramycin and nalidixic acid for plasmid selection and elimination of any remaining *E. coli* cells, respectively.

d. Selection and investigation of transconjugants

To select transformants with double-crossovers, spores appearing after 3-4 days were replica-plated on MYM agar containing either apramycin (50 μg/ml) or apramycin and thiostrepton (25 μg/ml of each). Since plasmid pJV326 is derived from pHJL400 (Larson & Hershberger, 1986), it is lost during propagation without thiostrepton selection. Transconjugants with an Am^RTs^S phenotype were selected and tested by Southern hybridization for double-crossovers to confirm loss of the vector and integration of the disrupted DNA. Stock spore suspensions of disruption mutants were prepared by growing transconjugants on either MYM agar or TO agar containing apramycin. To test transconjugants for the effect of gene disruption on Cm production, apramycin was added to seed culture and Cm-production media. Corynecin II was detected in and purified from transconjugant cultures, by procedures similar to those used for Cm in wild-type cultures (see Section VI) except that 5% methanol was used for both equilibration of the C18 column and application of the sample. The eluant gradient was started from 5% methanol in water and increased very slowly until corynecin II was eluted (10-12% methanol).

X. DNA isolation procedures

a. Isolation of plasmid DNA from *E. coli* cells

The alkaline lysis extraction technique was used (Sambrook *et al.*, 1989). An overnight culture of *E. coli* harboring the plasmid of interest was centrifuged, and the pellet was resuspended by vortexing in 100 µL of cold Solution I. Freshly prepared Solution II (200 µL) was added, mixed in by inversion, and kept on ice for 10 min. Solution III (150 µL) was then mixed in, and after cooling on ice for 10 min, the mixture was centrifuged at 10,000 rpm for 5 min to pellet cell debris. The plasmid-containing supernatant was transferred to another micocentrifuge tube and if necessary was extracted with equal volume of phenol:chloroform or chloroform:isoamyl alcohol. The aqueous phase, separated by centrifugation, was mixed with an equal volume of isopropanol and vortexed briefly. The DNA was pelleted by centrifuging at 10,000 rpm for 5 min, and was washed with 1 ml 70% ethanol. The plasmid DNA was dried under vacuum and redissolved in 75-100 µL of TE buffer. If required RNase A was added at a final concentration of 20 µg/ml, and the solution was stored at –20 °C.

b. Isolation of plasmid DNA from streptomycetes

Streptomycete cultures were grown in MYM medium containing the necessary antibiotic(s) for 36-48 h on a rotary shaker (250 rpm) at 30 °C. The mycelium from 3 ml of a culture was harvested by centrifuging at 8,000 rpm for 10 min, and was washed once with 0.3 M sucrose solution. The washed mycelium was resuspended in 0.2 ml Solution I containing 2 mg/ml of lysozyme The mixture was incubated at 37 °C for 30 min, after which 0.3-0.4 ml of a 2% SDS solution in 0.3 M NaOH, was added and vortexed for 30 s. The mixture was incubated at 70 °C for 15 min, and then cooled to room temperature in

30 min. Finally 0.3 ml of Solution III was mixed in, and after cooling on ice for 15 min, and centrifuging at 10,000 rpm for 5 min, the supernatant was transferred to a new microtube, and extracted once with acid phenol:chloroform. Sodium acetate (3 M, pH 5.2, 1/10 volume) and cold isopropanol (equal volume) were added, and the plasmid DNA precipitated was pelleted by centrifuging at 10,000 rpm for 5 min. The pellet was washed twice with 1 ml cold 70% ethanol, then with 1 ml absolute ethanol, and dried under vacuum. To remove any genomic DNA remaining, the plasmid DNA sample was dissolved in TE buffer, mixed with Solution II and kept on ice for 5 min. Plasmid DNA was extracted with acid phenol:chloroform, and precipitated by adding 1/10 volume sodium acetate (3 M, pH 5.2) and an equal volume of cold isopropanol. DNA was precipitated by centrifugation and the pellet was washed with 70% ethanol. The dried pellet was redissolved in TE buffer, and if larger quantities of plasmid were required, was used to transform into *E. coli*.

c. Isolation of genomic DNA from streptomycetes

The procedure described by Hopwood *et al.* (1985a) was used with minor modifications. Mycelia from a 25-ml culture was collected by centrifugation (10 min at 8,000 rpm). The pellet was washed twice with 10.3% (w/v) sucrose and the cells were pelleted. The supernatant fluid was carefully decanted, and the cells were resuspended in 5 ml of lysozyme buffer supplemented with 2 mg/ml lysozyme and 100 µg/ml RNaseA. After the suspension had been incubated at 37 °C for 30–40 min, 1.2 ml of 0.5 M EDTA and 0.13 ml of pronase (2 mg/ml) were gently mixed in, and incubated at 37 °C for 1.5 h. Then 0.7 ml of 10% (w/v) SDS was mixed in by inverting 5-6 times. The mixture was incubated for 2 h with gentle inversion every 30 min, after which 6 ml of

chloroform:isoamyl alcohol was mixed in thoroughly. The emulsion was centrifuged at 8,000 rpm for 10 min. The upper (aqueous) layer was transferred to a new tube and re-extracted with chloroform:isoamyl alcohol, then transferred to a polypropylene tube containing 0.1 volume (about 0.6 ml) of 3 M sodium acetate (pH 5.2). One volume of cold isopropanol was mixed in gently and swirled until the DNA precipitated. The DNA was then collected by spooling on a glass rod, washed with 70% ethanol and dried under laminar air flow. The spooled DNA was placed in sterile TE buffer (5 ml) and left at 4 °C overnight to dissolve. The solution was extracted twice with 6 ml chloroform:isoamyl alcohol, then transferred to a tube containing 0.1 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of cold absolute ethanol. The contents were gently swirled until the DNA precipitated. The DNA was collected on a glass rod, and was washed in 1 ml of 70% ethanol in a microtube. After 3 min, the ethanol was carefully poured out of the tube, and the DNA was air-dried in a laminar flow cabinet for 1-2 h. It was then dissolved in an appropriate volume (0.5-1 ml) of TE buffer.

XI. DNA manipulations

Samples of DNA were digested with restriction enzymes under the conditions recommended by the enzyme suppliers, using 0.5-1.0 µg DNA. The digestion mixture was incubated at the recommended temperature (usually 37 °C) for at least an hour. If necessary, the mixture was extracted with chloroform:isoamyl alcohol solution and centrifuged at 8,000 rpm for 5 min. The upper phase was transferred to a new microfuge tube, and DNA was precipitated by adding 1/10 volume sodium acetate (3 M, pH 5.2) and 1 volume cold isopropanol. By centrifugation at 8,000 rpm for 5 min, the DNA was pelleted and washed with 70% ethanol. It was dried under vacuum for storage.

For subcloning a DNA fragment, the vector DNA was linearized at the appropriate restriction site, purified with chloroform:isoamyl alcohol, and dephosphorylated with calf-intestine alkaline phosphatase (CIAP) at 37 °C for 40 min. To inactivate the CIAP the incubation mixture was heated at 70 °C for 10 min, then cooled and extracted with chloroform:isoamyl alcohol. The DNA was precipitated by addition of sodium acetate and cold isopropanol. For ligation the insert and vector DNA were mixed in a 1X ligase buffer at 2:1 molar ratio, and T4-DNA ligase (1 unit) was added. The reaction was incubated overnight at 14 °C for both cohesive-end and blunt-end ligations. However, for both types of ligation, incubation at room temperature for 2-3 h gave comparable results. To increase the efficiency of subsequent transformations, the ligation mixture after completion was heated at 65 °C for 10 min to denature the ligase.

To convert DNA overhangs to blunt ends, the DNA sample obtained after digestion with the enzyme creating cohesive ends was purified by extraction with chloroform:isoamyl alcohol, and added to 200 µl of S1 nuclease buffer containing 50 units of S1 nuclease. The reaction mixture was incubated at 37 °C for 30-40 min, then mixed with 20 µl of Stop solution and heated at 70 °C for 10 min. The sample was cooled and the DNA purified with chloroform:isoamyl alcohol solution, was mixed with 1/10 volume of 3 M Na-acetate at pH 5.2 and 1 volume of cold isopropanol. After centrifugation the DNA pellet was washed with 70% ethanol and dried under vacuum, and redissolved in 15-20 µl TE buffer.

a. Electrophoresis of DNA

Agarose gels were prepared in 1X TAE buffer at concentrations between 0.8% to 1.2% (w/v). Electrophoresis was carried out at 90 V for 60-120 min. Gels were stained in

aqueous ethidium bromide (1 mg/ml) solution for 20 min, destained in water for 5 min, and viewed under UV light at 300 nm.

b. Directed nested deletions in DNA inserts

DNA fragments were cloned in the plasmid vector pUC18 or phagemid vector pBluescript II SK(+). The specific recommendations provided in the ExoIII/S1 Deletion Kit of *MBI Fermentas* were used to make nested deletions in the insert. The resulting fragments of DNA were circularized by ligation and used to transform competent cells of *E. coli*. Plasmids that differed in size progressively by 200-250 bp were isolated from transformants for sequencing.

c. Sequence determination and analysis

For sequence homology searches, the BLAST network service at the National Center for Biotechnology Information was accessed through the Internet. Multiple sequences were aligned with the Pileup program of the GCG (Genetics Computer Group) package or a web-based Clustal W program (http://www2.ebi.ac.uk/clustalw/). Nucleotide sequences were analyzed for their %G+C content and to identify open reading frames by the web-based FramePlot 2.3 program (Ishikawa & Hotta, 1999) (http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl.

d. Purification of phage DNA

The general methods of Sambrook *et al.* (1989) were used with some modifications. Cells (100 µl) of *E. coli* LE392 were mixed with 100 µl of purified recombinant phage, vortexed and incubated at 37 °C for 30 min to allow phage adsorption. The mixture was added to 50 ml LB broth in 250 ml flasks containing 0.2 mM $MgSO_4$. The culture was

shaken at 37 °C for 14-16 h, then 25 ml LB broth pre-warmed to 37 °C was added and incubation was continued for 3-4 h. Cell debris was removed by centrifugation (10,000 rpm for 15 min at 4 °C) and the supernatant was incubated with 150 µg RNase, and 100 µg DNase I at 37 °C for 30 min. An equal volume (75 ml) of 20% w/v PEG 8,000 and 2.5 M NaCl in λ diluent was mixed with the phage suspension and the mixture was left on ice for 1 h. After centrifugation (10,000 rpm for 15 min at 4 °C) the supernatant was discarded by inverting the tube on a paper towel to drain all liquids, and the phage pellet was dispersed in 2 ml TE buffer. After 200 µg SDS and 120 µg pronase had been added, the mixture was incubated at 37 °C for 1 h, and extracted three times with equal volumes of chloroform:isoamyl alcohol. The phage DNA was precipitated by adding 1/10 volume of 3 M sodium acetate (pH 5.2) and one volume of cold isopropanol, kept at –20 °C for 2 h and then centrifuged (10,000 rpm for 15 min at room temperature). The DNA pellet was washed twice with 70% and once with 100% ethanol, dried and dissolved in 150-200 µl TE. For further purification the DNA solution was extracted once more with chloroform:isoamyl alcohol.

XII. Southern hybridization

a. Transfer of phage from plaques to a nylon membrane

A nylon membrane was placed on the surface of the top agarose so that it came to direct contact with the plaques. The filter was oriented by stabbing a needle through the membrane and agar at three asymmetric locations, and after 2 min to allow the plaques to adhere, the filters were transferred, DNA side up, to a filter paper saturated with denaturing solution. After 5 min, the filter was placed for 5 min on a filter soaked with neutralizing solution, and then for 5 min on a third filter soaked in 2xSSC. The filter was

dried, DNA side up, on a paper towel for up to one hour, and then sandwiched between sheets of 3MM paper and baked at 80 °C for 2 h to immobilize the DNA on the filter before hybridizing with a ^{32}P-labeled probe. For hybridization the membrane was first immersed in wash solution at 45 °C for 2 h, and bacterial debris was rubbed off the surface with a tissue soaked in wash solution.

b. Transfer of DNA from an agarose gel to a nylon membrane

The agarose gel containing DNA digests fractionated by gel electrophoresis was stained with ethidium bromide, and photographed alongside a fluorescent ruler. Using standard procedures (Sambrook *et al.*, 1989) the gel was first soaked in 0.25 M HCl for 10 min to depurinate the DNA, then rinsed with distilled water, and agitated in denaturing solution for 30 min at room temperature. After washing with distilled water, the gel was shaken in neutralizing solution for 30 min, and finally washed with 10X SSC before the DNA was transferred to a nylon membrane (GeneScreen Plus, *New England Nuclear Inc.*, Boston, MA) by vacuum filtration or capillary diffusion in 10X SSC buffer. The membrane was then air dried for 30-60 min, sandwiched between two sheets of Whatman 3MM paper, and baked in a vacuum oven at 80 °C for 2 h.

c. Preparation of probes

Probes (approximately 100 ng of DNA) were labeled with [α-^{32}P]dCTP by random priming using the procedures recommended in the random priming kit of *Pharmacia Biotech*. The linearized DNA was first denatured by heating 4 min at 95-100 °C and keeping on ice for 5 min. It was then added to the labeling reaction containing 5 µl of a [α-^{32}P]dCTP (10 µCi/µl) solution and incubated at 37 °C for 30 min. The labeled DNA

was denaturated by heating in boiling water for 5 min and cooled on ice before it was added to the hybridization bottle.

d. Hybridization

The nylon membrane on which DNA samples from either agarose gels or plaque lifts had been immobilized by blotting, was washed once with 6X SSC, and rotated in a hybridization bottle with pre-warmed (65 °C) prehybridization solution at 65 °C for 2 h, after which the denatured labeled probe was added. Hybridization proceeded overnight at temperatures between 65 °C and 68 °C. Membranes were washed successively at room temperature twice with 2X and once with 1X SSC solutions containing 0.1% (w/v) SDS, then with 0.1x SSC / 0.1% (w/v) SDS solution at 67 °C for 1 h, and once more at room temperature for 2 min. They were then wrapped in Saran wrap and exposed to a Bio-Rad CS phosphor-imaging screen (*Bio-Rad Laboratories*, Mississauga, Ont.). A Bio-Rad GS525 Molecular Imager was used to scan the screen.

Plaques giving hybridization signals were located and extracted into 400 µl of SM buffer by shaking at 37 °C for 2 h. For a second round of hybridization these phage extracts were titred, and diluted to give about 20 plaques per plate. The plaques were transferred to nylon membranes and hybridized again with the same probe. Single well-separated plaques that hybridized strongly were considered pure and were picked, extracted into SM buffer, and used to isolate the phage DNA.

XIII. PCR techniques and TOPO cloning

Degenerate PCR primers were designed from two conserved regions recognized from the alignment of deduced amino acid sequences of halogenases PrnC (Hammer *et al.*,

1999), PltA, PltM, and PltD (Nowak-Thompson *et al.*, 1999) and Chl (Dairi *et al.*, 1995). To increase the specificity of primer binding to the target sites, codon usage in *Streptomyces* species (Wright & Bibb, 1992) was taken into account while designing primer sequences. The complexity of the forward primer pool was reduced by using, deoxyinosine (Huang & Jeang, 1994) at two positions where there was high amino acid variability. Oligonucleotides were synthesized by *MWG-Biotech* (USA).

Formamide was included in the PCR reaction mixture to increase the likelihood of specific annealing between primers and target. Amplification was optimum with 1% (v/v) formamide and decreased significantly at higher concentrations; at 4% no amplification was detected. Final concentrations of ingredients in 25 µl reaction mixtures were 10 mM Tris-HCl (pH 8.8), 50 mM KCl, 1.25 mM $MgCl_2$, 1% v/v formamide, 0.08% Nonidet P40 (*MBI Fermentas*) and 0.2 mM for each deoxynucleoside triphosphate (dNTP). Each reaction contained 12.5 pmol of each primer, 2.5 ng of genomic DNA and 1.5 U *Taq* DNA polymerase (*MBI Fermentas*). The touch down-PCR method (Don *et al.*, 1991), programmed with a thermal cycler (PTC-150 MiniCycler *MJ Research*), began with a 5 min "hot start" at 96 °C before addition of *Taq* polymerase. The actual amplification programme consisted of thirty six cycles of three steps: denaturation at 96 °C, annealing at 65-55 °C and extension at 72 °C; the initial annealing temperature of 65 °C was reduced 2 °C after every six cycles to reach 55 °C for the final six cycles; each step took 1 min. To facilitate cloning in the pCR2.1TOPO vector, the PCR ended with an extra extension step of 8 min at 72 °C, as recommended in the *Invitrogen* manual, to ensure completion of the *Taq* polymerase-catalyzed addition of single dATPs to the 3'-ends of products. The PCR product amplified from the genomic DNA of S. *venezuelae* ISP5230

was electrophoresed in a 1.2% agarose gel. A gel slice containing the unique ~300-bp PCR product was excised and extracted (UltraClean 15 DNA extraction kit, *Mo Bio Labs*). The DNA fragment was cloned by using the TOPO TA cloning kit (*Invitrogen*). A mixture containing 3.5 µL of the extracted PCR product, 1 µL of salt solution (1.2 M NaCl, 0.06 M $MgCl_2$), 1 µL of TOPO 2.1 Vector and 0.5 µL of distilled water was incubated at room temperature for 10 min; then 3 µL was used to transform competent *E. coli* TOP10F′ cells provided with the kit. The transformed cells were plated on X-gal:IPTG agar containing 100 µg/ml ampicillin. White colonies yielded recombinant plasmid pJV501, containing an amplicon of about 300 bp. The sequence of the DNA cloned was determined by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977) from the M13 reverse primer site of the vector using an automated ABI Prism model 373 DNA sequencer. The PCR product retrieved from pJV501 by *EcoR*I digestion, and labeled with a random primer kit (*Pharmacia Biotech*), was used at high stringency to probe Southern blots of the restriction digests of *S. venezuelae* genomic DNA.

XIV. Phylogenetic trees

Sequences were saved in FASTA format in one file, and the ClustalX programme was used to prepare a dendrogram showing sequence relationships. The TreeView 1.6.1 programme (Page, 1996) was used to view the phylogenetic tree.

XV. Introducing ORFs 11,12 and 13 into mutant *cml-2*

The three genes (ORFs 11, 12 and 13) identified in this study were cloned as various fragments in plasmids pJV507, pJV526 and pJV528. pJV507 was prepared by subcloning the 3.8-kb *Sac*I-*Xba*I fragment of pJV506 in pJV326 and contained ORFs 12 and 13.

pJV526 contained ORFs11, 12 and 13, introduced by subcloning the 6.2-kb *Sac*I-*Sac*I fragment of *S. venezuelae* DNA from pJV502 in pJV326. pJV528 contained only ORF13 as an intact gene, and was prepared by deleting a 2.7-kb *Stu*I-*Eco*RV fragment from pJV526 and self-ligating the remaining linear fragment. These plasmids were transferred conjugally from *E. coli* ET12567 (pUZ8002) into *S. venezuelae* mutant *cml-2* to test for complementation of the defect in chlorination. Transconjugants with a Ts^R phenotype were selected and after two rounds of sporulation on Ts-free MYM medium were tested for Cm production.

Results

Section 1. Feeding experiments

To examine possible roles for CoA derivatives of acetic, acetoacetic, and dichloroacetic acids in biosynthesis of the dichloroacetyl moiety of Cm, ^{13}C-labeled thioesters of the *N*-acetylcysteamine (NAC) analogue of CoA synthesized in the Chemistry Department, Dalhousie University (Lewis, 1998) were added to cultures of *S. venezuelae* during production of Cm. The isotopically labeled compounds were fed at similar molar concentrations and Cm was extracted with ethyl acetate from the cultures 7 days after inoculation. The Cm extract was purified by reverse phase chromatography and analyzed by NMR spectroscopy for ^{13}C enrichment in the carbons of the dichloroacetyl group.

Isotope from sodium [2-^{13}C] acetate was incorporated into both carbons, but ^{13}C enrichment in the dichloromethine carbon was about 8 times higher than in the carbonyl carbon (Table 2). A similar pattern of ^{13}C enrichment was observed after feeding sodium [3-^{13}C]pyruvate or [2-^{13}C]acetyl-NAC. Reinforcing the conclusion that the acetyl moiety had not been incorporated intact, [1,2-$^{13}C_2$]acetyl-NAC was incorporated into the dichloroacetyl moiety only after considerable ^{13}C-^{13}C bond scission; incorporation into the carbonyl carbon was negligible. The results indicated that none of the compounds is a direct precursor of the dichloroacetyl group of Cm, and implied that the chlorinated acyl substituent is not formed via acetyl CoA.

The patterns of ^{13}C enrichment in Cm from L-[3-^{13}C]phenylalanine and ethyl [2,4-$^{13}C_2$]acetoacetate were similar in that both precursors enriched the dichloromethine carbon with ^{13}C, and gave little or no enrichment of the carbonyl carbon. These results

Table 2. Enrichment of ^{13}C from labeled compounds* in carbons of the dichloroacetyl group of chloramphenicol

Labeled compound	Cm isolated (mg)	^{13}C incorporation above natural abundance (%)	
		Cl_2CH-	-C=O
Sodium [2-^{13}C]acetate	14	4.8	0.6
Ethyl [2,4-$^{13}C_2$]acetoacetate	11	0.5	0.2
Sodium [3-^{13}C]pyruvate	12	2.1	0.2
[2-^{13}C]Acetyl NAC	21	2.2	0.2
[1,2-$^{13}C_2$]Acetyl NAC	15	1.5	0.0
L-[3-^{13}C]Phenylalanine	18	0.6	0.0
[2-^{13}C]Acetoacetyl NAC	12	0.8	0.3
Ethyl [1,2-$^{13}C_2$]dichloroacetate	18	0.2	0.3
[1,2-$^{13}C_2$]Dichloroacetyl NAC	10	0.2	0.6

*All compounds contained 99 atom % ^{13}C in the labeled atom, except for [1,2-$^{13}C_2$]acetyl NAC, where the ^{13}C enrichment in each labeled atom was 80%.

are consistent with evidence that in *S. venezuelae* phenylalanine is metabolized to acetoacetate via the homogentisate pathway (N. Ranade and L. C. Vining unpublished), and with the observation that [2-^{13}C]acetoacetyl-NAC enriched mainly the dichloromethine carbon. The latter result implies that the acetoacetyl-NAC was hydrolyzed to [2-^{13}C]acetate. Incorporation of ^{13}C from ethyl [1,2-$^{13}C_2$]dichloroacetate was negligible. Low enrichment of Cm from this compound might be due to its hydrolysis to dichloroacetate, which has been shown in previous studies (Simonsen *et al.*, 1978) not to be incorporated into Cm. Investigations (Lewis, 1998) of the stability of NAC esters under conditions mimicking those existing in the culture medium (in phosphate buffer at pH 7 at 33 °C) showed no hydrolysis after 4 days for acetyl-NAC, and a half life of about 14 h for acetoacetyl NAC making the nonenzymatic hydrolysis an insignificant factor in feeding experiments. Dichloracetyl-NAC, however, showed a low stability (half life 1.5 h) under these conditions. The results strengthen the conclusion that free dichloroacetate is not a direct precursor of the dichloroacetyl group of Cm and also indicate that its ethyl ester is either not taken up by the mycelium or is rapidly hydrolyzed. The very low incorporation of ^{13}C from [1,2-$^{13}C_2$]dichloroacetyl NAC provided further evidence that dichloroacetate is not a pathway intermediate formed by chlorination of acetoacetyl CoA.

Section 2. Isolation of 3'-*O*-acetyl chloramphenicol

During the feeding experiments with labeled compounds, HPLC analysis of cultures showed that Cm was accompanied by a second metabolite with a higher retention time. This compound was purified from culture extracts and characterized. The NMR data from the purified compound matched those of 3'-*O*-acetylCm. Most of ^{13}C isotope

incorporated from [2-^{13}C]acetate was present in the dichloroacetyl group of 3'-*O*-acetylCm rather than in its acetyl moiety. However, significantly more of the label from ethyl [2,4-$^{13}C_2$]acetoacetate was incorporated into the 3'-*O*-acetyl group, implying that this part of the molecule was formed from acetyl CoA derived from the labeled precursor (Groβ *et al.*, 2001). In view of the evidence that *S. venezuelae* lacks chloramphenicol acetyltransferase (Shaw & Hopwood, 1976; Nakano *et al.*, 1977), we investigated the role of 3'-*O*-acetylCm in Cm biosynthesis. 3'-*O*-AcetylCm was chemically synthesized and used to determine whether accumulation of this compound in cultures could be due to inhibition of acetylCm esterase activity often present in streptomycetes (Nakano *et al.*, 1977). If this enzyme is normally active in *S. venezuelae* it might hydrolyze 3'-*O*-acetylCm to release the free antibiotic. The esterase activity was detected in the particulate fraction of sonicated mycelium, but only in cultures harvested during the late growth phase. Mycelium harvested during exponential growth (days 2, 3 and 4 after inoculation of the production medium) did not exhibit significant esterase activity. The 3'-*O*-acetylCm remained intact when it was incubated with boiled mycelium or with the culture fluid passed through a 0.2 μm cellulose acetate filter (Table 3). Similarly, addition of PMSF at 5 mM completely inhibited esterase-catalyzed hydrolyzes of acetyl-Cm. Moreover accumulation of 3'-*O*-acetyl ester of Cm in the broth increased when the esterase inhibitor acetylsalicylic acid (ASA; 12 mM) was added to Cm-producing cultures. ASA remained in the culture for the whole fermentation period, and the initial concentration of Cm remained unchanged. Decrease in concentration of 3'-*O*-acetylCm, measured by HPLC, was used as an indicator of the esterase activity, and control samples

Table 3. Evidence for esterase-catalyzed hydrolysis of acetyl-Cm by the enzyme located within the insoluble fraction of cell membrane.

Sample	% Reduction in 3'-*O*-acetylCm (after 1 h incubation)
Washed mycelium	68
Insoluble fraction of mycelium sonicate	30
Boiled mycelium	NS
Supernatant from mycelium sonicate	NS
Cell-free extract + 5 mM PMSF *	NS
Culture filtrate (passed through 0.2 μm membrane)	NS
Boiled culture-filtrate (passed through 0.2 μm membrane)	NS

NS, Equal or less than 7% due to non-enzymatic hydrolysis and considered non-significant

***PMSF**, Phenylmethylsulfonyl fluoride was used as esterase inhibitor

were used, in which concentration of 3'-*O*-acetylCm was measured immediately after addition.

Section 3. Investigation of acyltransferase activity

Based on the presumption that accumulation of corynecins as minor metabolites accompanying Cm in *S. venezuelae* cultures was due to differences in the substrate preference of an acyl transferase in the Cm biosynthesis pathway, it was expected that the proposed N-dichloroacetyl transferase responsible for dichloroacetylation of the amino acid intermediate *p*-aminophenylserine could be purified by monitoring its N-acetyltransferase activity. This expectation was reinforced by evidence that in the nonproducing mutant *cml-2* low levels of the non-chlorinated congeners were produced (Doull *et al.*, 1985). However, *p*-aminophenylserine was not immediately available for use as the substrate to search for the N-acetyltransferase activity in cell free extracts of *S. venezuelae* and it was necessary to postpone this approach until the amino acid and a reference sample of the N-acetyl derivative could be prepared. In the interim, N-acetyltransferase activity in *S. venezuelae* was investigated with *p*-nitrophenylserinol (*p*-NPS), which was available as a substrate. Purification of this enzyme was expected to provide relevant procedures for isolating *p*-aminophenylserine dichloroacetyltransferase. Moreover, culture conditions in which acetylation of *p*-NPS occurs in *S. venezuelae* have been well established (Gottlieb *et al.*, 1956).

To detect N-acetyltransferase activity in the mycelium of *S. venezuelae* assay conditions using *p*-nitrophenylserinol and acetyl coenzyme A as co-substrates were devised based on Gottlieb *et al.* (1956). The activity was detected in cell-free extracts prepared by sonication of *S. venezuelae* mycelium in buffer E7. Solid ammonium sulfate

was added to the cell-free extract to 40% saturation. After 15 min stirring, the precipitate was pelleted by centrifugation (15,000 rpm, 15 min, at 4 °C) and discarded. The supernatant was brought to 65% saturation with ammonium sulfate and stirred for 15 min in ice bath. The precipitate was collected by centrifugation, redissolved in 15 ml buffer D8 and applied to a DEAE-cellulose column (35 x 2.5 cm) equilibrated with the same buffer. Using a step-wise gradient of 50, 100, 200, 300, 400, 500, 600 mM NaCl solutions, the activity was eluted from the DEAE column when the NaCl concentration in the gradient reached about 400 mM. The active samples combined from the DEAE-cellulose column were concentrated by ultrafiltration (*Amicon* model 12, and UM 20 filter) to 1.2 ml, and applied to a size-exclusion column (TSK-Gel HW-55, 22x1.5 cm; *Supelco*) equilibrated with buffer E8. The acetylating enzyme was eluted with buffer E8 at a volume consistent with a molecular weight of about 43,000 Da. Active fractions were pooled and applied to a hydroxyapatite column (Econo-Pac CHT-II cartridge; *BioRad*) that was equilibrated with 10 mM phosphate buffer, pH 7.2. The target enzyme was not adsorbed on the column, and was collected by washing with 10 mM sodium phosphate at pH 7.2; thus, the enzyme activity was detected in the early tubes. It showed no loss of activity at 4 °C when kept for 2 months, but was sensitive to heat treatment (60 °C for 5 min). The overall purification was about 70 fold, with a yield of about 9%.

Section 4. PCR amplification of a halogenase gene from *S. venezuelae*

The amino acid sequences of five halogenases were aligned and used to design PCR primers from two consensus regions (Figure 8 and Table 4). Using primers MPF1 and MPR2 a unique DNA fragment with a size about 300 bp was amplified from *S. venezuelae* ISP5230 genomic DNA. The PCR product was electrophoresed in an agarose

gel, and the ~300-bp band of DNA was excised. It was extracted from the gel and ligated into the pCR-TOPO vector to give pJV501.

To confirm that the PCR product originated from *S. venezuelae*, it was excised from pJV501 by *Eco*RI digestion, labeled with [α-^{32}P]dCTP and used to probe genomic DNA of *S. venezuelae* ISP5230 digested with different restriction enzymes. Hybridization at high stringency was detected with fragments of various sizes from the genomic DNA digests, confirming that the PCR product originated from *S. venezuelae* (Figure 9).

```
PfPrnC      MTQKSPAN-EHDSNHFDVIILGSGMSGTQMGAILAKQQFRVLIIEESSHPRFTIGESSIP    59
Chl         MTDTTADQTRHGDRPYDVVIIGSGLSGTMLGSILAKHGFRIMLLDGAHHPRFAVGESTIR    60
PltM        -----------MNQYDVIIIGSGIAGALTGAVLAKSGLNVLILDSAQHPRFSVGEAATP    48
PltD        ---MNDVQSGKAPEHYDILLAGNSISVIMLAACLARNKVRVGLLRNRQMPPDLTGEATIP    57
PltA        ----------MSDHDYDVVIIGGGPAGSTMASYLAKAGVKCAVFEKELFEREHVGESLVP    50

PfPrnC      ETSLMNRIIADRYGIPELDHITSFYSTQRYVASST-GIKRNF-GFVFHKPGQEHDPKEFT    117
Chl         QTLVVLRLISDRYGC-EIANLASFQDVLANVSSSH-GQKSNF-GFMFHRDGEEPDPNETS    117
PltM        ESGFLLRLLSKRFDIPEIAYLSHPDKIIQHVGSSACGIKLGF-SFAWHQENAPSSPDHL-    106
PltD        YTSMIFELIADRYGVPEIKNIARTRDIQQKVMPS-SGVKKNL-GFIYHQRSRAVDLGQAL    115
PltA        ATTPVLLEIGVMEKI-EKANFPKKFGAAWTSADSGPEDKMGFQGLDHDFRSAEILFNERK    109

PfPrnC      QCVIPELPWGPESHYYRQDVDAYLLQAAIKYGCKVHQKTTVTEYH-ADKDGVAVTTAQGE    176
Chl         QFRIPSIVGN-AAHFFRQDTDSYMFHAAVRYGCDARQYYRVENIE-FDDGGVTVSGADGS    175
PltM        ---VAPPLKVPEAHLFRQDIDYFALMIALKHGAESRQNIKIESIS-LNDDGVEVALSNAA    162
PltD        QFNVPSEH--GENHLFRPDIDAYLLAAAIGYGAQLVEIDNSPEVL-VEDSGVKVATALGR    172
PltA        QEGVDRDF---TFHVDRGKFDRILLEHAGSLGAKVFQGVEIADVEFLSPGNVIVNAKLGK    166

PfPrnC      R---FTGRYMIDCGGPRAPLATKFKLREEPCRFKTHSRSLYTHMLGVKPFDDIFKVKGQR    233
Chl         T---VRARYLVDASGFRSPLARQLGLREEPSRLKHHARSIFTHMVGVDAIDDHVDMPAEL    232
PltM        P---VKAAFIIDAAAQGSPLSRQLGLRTTEG-LATDTCSFFTHMLNVKSYEDALAPLSRT    218
PltD        W---VTADFMVDGSQGGQVLARQAGLVSQASTQKTRTLEFSTHMLGVVPFDECVQGD---    226
PltA        RSVEIKAKMVVDASGRNVLLGRRLGLREKDPVF---------NQFAIHSWFDNFDRKSAT    217

                                MPF1 →
PfPrnC      W---RWHEGTLHHM-FEGGWLWVIPFNNHPRSTNNLVSVGLQLDPRVYP-KTDISAQQEF    288
Chl         RPPVPWNDGTMHHI-FERGWMWIIPFNNHPGATNPLCSVGIQLDERRYPARPDLTPEEEF    291
PltM        RSPIELFKSTLHHI-FEEGWLWVIPFNNHPQGTNQLCSIGFQFNNAKY--RPTEAPEIEF    275
PltD        -FPGQWHGGTLHHV-FDGGWVGVIPFNNHQHSRNPLVSVLVSLREDLC---PSMDGDQVL    281
PltA        QSPDKVDYIFIHFLPMTWTWVWQIPI------TETITSVGVVTQKQNYT-NSDLTYEEFF    270

                                                          ← MPR2
PfPrnC      DEFLARFPSIGAQFRDAVPVRDWVKTDRLQFSSNACVGDRYCLMLHANGFIDPLFSRGLE    348
Chl         RSHVDRFPAVQRQLKGARSVREWVRTDRMQYSSSRTVGERWCLMSHAAGFIDPLFLRGLS    351
PltM        RKLLKKYPAIGEHFKDAVNAREWIYAPRINYRSVQNVGDRFCLLPQATGFIDPLFSRGLI    335
PltD        AGLIELYPGLGRHLSGARRVREWVLRQPPRQVYRTALERRCLMFDEGAASWDLLFSRRLS    341
PltA        WEAVKTRENLHDALKASEQVRPFKKEADYSYGMKEVCGDSFVLIGDAARFVDPIFSSGVS    330

PfPrnC      NTAVTIHALAARLIKALRDDDFSPERFEYIERLQQKLLDHNDDFVSCCYTAFSDFRLWDA    408
Chl         NTCEIINALSWRLMAALREDDFAVERFAYVEELEQGLLDWNDKLVNNSFISFSHYPLWNS    411
PltM        TTFESILRLAPKVLDAARSNRWQREQFIEVERHCLNAVATNDQLVSCSYEAFSDFHLWNV    395
PltD        NAAELVLALAHRLIKAAHSGDYRSPALNDFVLTQDSIISLSDRIALAAYVSFRDPELWNA    401
PltA        VALNSARIASGDIIEAVKNNDFSKSSFTHYEGMIRNGIKNWYEFIT---LYYRLNILFTA    387

PfPrnC      FHRLWAVGTI---LGQFRLVQAHARFRASRNEGDLDHLDNDPPYLGYLCADMEEYYQLFN    465
Chl         AFRIWASASV---IGGKRILNALTRTKETGDDSHCQALD-DNPYPGLWCP-LDFYKEAFD    467
PltM        WHRVWLSGSN---LGSAFLQKLLHDLEHSGDARQFDAALEAVRFPGCLSLDSPAYESLFR    452
PltD        FARVWLLQSIAATITARKINDAFAKDLDPRVFDEIDQLAED----GFWMPLYRGYKDILN    457
PltA        F-----------------------VQDPRYRLDILQLLQGDVYSGKRLEVLDKMREIIA    423
```

Figure 8. Alignment of halogenases used to design PCR primers MPF1 and MPR2. The conserved regions used are shown in bold letters.
Chl, halogenase from *Streptomyces aureofaciens*; PfPrnC, halogenase PrnC from *Pseudomonas fluorescens*; PltM, PltD, and PltA, halogenases from *P. fluorescens*.

Table 4. Halogenase amino acid sequences and specifications of primers.

Halogenase	Amino acid sequences of consensus regions used for forward primer	for reverse primer
PfPrnC	EGGWLWVIP	IDPLFSRGL
Chl	ERGWMWIIP	IDPLFLRGL
PltM	EEGWLWVIP	IDPLFSRGL
PltD	DGGWVGVIP	NDLLFSRKL
PltA	TNTWVWQIP	VDPIFSSGV

Derived primer sequences (5' to 3')

MPF1: S IRS RSS TGG ITS KGS VWS ATC CCS

MPR2: S ASV YYS SRS VRG AAS AKS RGG TCS

Primer characteristics

MPF1: 25 mer; GC content, 64.6%; Tm, 69.8 °C; degeneracy 12,288 fold; PTS, 196,608

MPR2: 25 mer; GC content 65.3%, Tm, 70.1 °C; degeneracy 147,456 fold; PTS, 147,456

PTS, Potential target site0s; **K**=G+T, **R**=G+A, **S**=G+C, **V**=G+C+A, **W**=A+T, **Y**=C+T, **I**=inosine

- Chl, halogenase from *Streptomyces aureofaciens*; PfPrnC, halogenase PrnC from *Pseudomonas fluorescens*; PltM, PltD, and PltA, halogenases from *P. fluorescens*.

Figure 9. Hybridization of labeled-PCR product with restriction digests of *S. venezuelae* ISP5230 genomic DNA

Lanes: 1. *BamH*I (13 kb); 2. *Kpn*I (>20 kb, 15 kb); 3. *Sal*I (1.5 kb); 4. *Pvu*I (0.6, 0.8 kb);. 5. *Pst*I (8 kb). The >20-kb band in lane 2 corresponded to undigested genomic DNA.

Analysis with the Frameplot program of the nucleotide sequence of the pJV501 insert containing the PCR product amplified from *S. venezuelae* genomic DNA, using primers MPF1 and MPR2, detected a partial ORF with an average codon-third-letter bias of 93.7 mol% G+C (Figure 10), consistent with the range predicted for streptomycete genes (Wright & Bibb, 1992). A BLASTX search with the nucleotide sequence of the PCR product showed the following deduced amino acid sequence identities to GenBank proteins: 37% to a halogenase involved in the biosynthesis of a vancomycin group antibiotic (van Wageningen *et al.*, 1998), 36% to a halogenase from *Amycolatopsis mediterranei* (Pelzer *et al.*, 1999), 36% to a halogenase from *S. lavendulae* (Chiu *et al.*, 2001), 31% to a halogenase from *Xanthomonas oryzae pv. o*ryzae (accession AAG38844), 34% to PltA from *P. fluorescens* (Nowak-Thompson *et al.*, 1999), 32% to PrnC from *Myxococcus fulvus* (Hammer *et al.*, 1999), and 26% to Chl from *S. aureofaciens*. Figure 11 shows an alignment of these sequences. The PCR primers MPF1 and MPR2 amplified genomic DNA fragments from three streptomycetes in addition to *S. venezuelae* (Figure 12) as well as *S. aureofaciens* and *S. coelicolor*, which are not shown in the figure. The size of each PCR product was consistent with the distance (290-bp) between the two consensus sequences from which the primers were designed (Table 4), indicating a potential usefulness of the primers for amplification of halogenase gene fragments from other strains. No amplification occurred using *S. parvulus and S. akiyoshiensis* genomic DNA samples (Figure 12)

A)

```
GACGCTGCGGCGGAAGATCGGGTCGGTGAAGCAGGCGGCGTCTCCGCAGA     50
GGAAGAAGCGGTCGGCGGAGAAGACCTCGGTGTCGTAGGACCAGTCCTGC     100
ACGATCCGGACCTCGTCGACCTGCTCGGCGCCGCCGAGGATGTCCATGGC     150
CTTGGCGCACTTGGCGAGGGTGGACGAGTAGAAGGCGTCGGCGCCCTGCT     200
CGCGCACCTCGGCGGACTTCGATCGGTCGACGACGAGGCCGACGCTGTAC     250
AGGTCGTCCTTGATGGGGATCACCCAAACCCACCCCTTCA               290
```

B)

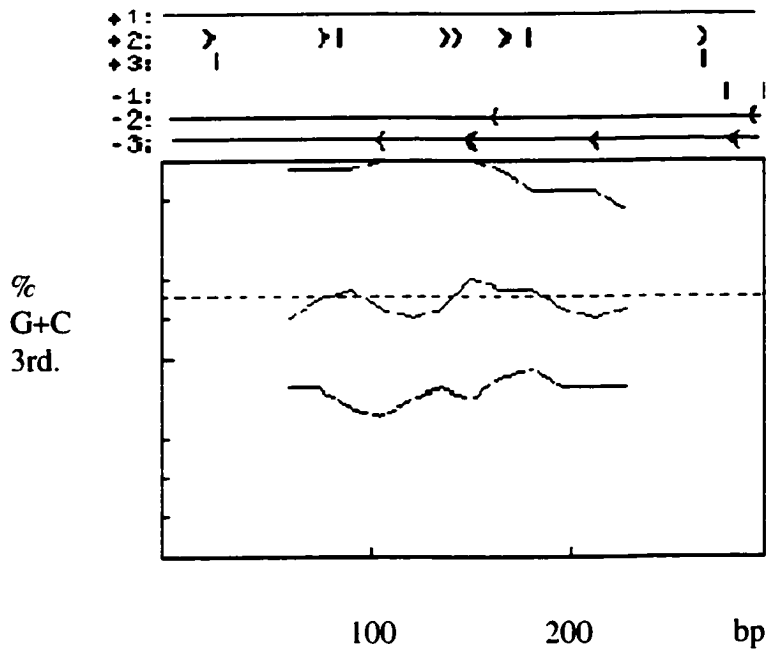


Figure 10. The putative halogenase PCR-product from *S. venezuelae*: **A)** Nucleotide sequence of the fragment cloned in pJV501; **B)** FramePlot analysis of the pJV501 insert sequence. Frame 3 of the complementary (-) strand represents the halogenase-like sequence; Lines in the top panel identify possible open reading frames from relevant start (> and <) and stop (|) codons. The plotted lines show %G+C in the 3rd position of codons for each reading frame; the hatched line represents 65.9% G+C.

Figure 11. Alignment of halogenase sequences with the deduced amino acid sequence of the PCR product cloned in pJV501.

Aohal, halogenase homologue from *Amycolatopsis orientalis* (Van Wageningen *et al.*, 1998); AmHal, halogenase from *A. mediterranei* (Pelzer *et al.*, 1999); SlHal, halogenase from *S. lavendulae* (Chiu *et al.*, 2001); MfPrnC, halogenase PrnC from *Myxococcus fulvus*; Chl, halogenase from *Streptomyces aureofaciens*; PltA, halogenase from *P. fluorescens*; ZoHal, halogenase from *Xanthomonas oryzae* pv. *oryzae* (accession AAG38844).
* indicates identical or conserved residues in all sequences in the alignment. : indicates conserved substitutions, and . indicates semi-conserved substitutions.

```
Aohal           ------------------MSVEDFDVVVAGGGPAGSTVATLVAMQGHRVLLLEKEVFPR
AmHal           ------------------MSVEDFDVVVAGGGPGGSTVATLVAMQGHRVLLLEKEVFPR
SlHal           -----------------MAAVTEEFDVAVVGGGPAGSTFAALVAKQGHRVVVLEKENFPR
MfPrnC          ----------MKPTVNAHHDSNHFDVIILGSGMSGSQMGAILGRQGFRVLIVEESTHPR
Chl             ------------------------------------------------------------
PCR             ------------------------------------------------------------
PltA            ------------------MSDHDYDVVIIGGGPAGSTMASYLAKAGVKCAVFEKELFER
ZoHal           MTSNAVRHTPTAMPAAGPAPGPECPDVLIVGGGPAGCTAAIALAELGWSVTLLEKEQHPR

Aohal           YQIGESLLPATVHG----------------VCRMLGITDELANAGFPVKRGGTFRWGARP
AmHal           YQIGESLLPATVHG----------------VCRMLGISDELANAGFPIKRGGTFRWGARP
SlHal           YQIGESLLPSTIHG----------------VCRLSGAADDLAKAGFPLKRGGTFRWGARP
MfPrnC          FTIGESSIPETSLMNRIIADRYGVPEIEDITSFYSTFKKVSSSTGIKRNFGFVFHK-PGE
Chl             ------------------------------------------------MFHR-DGE
PCR             ------------------------------------------------------------
PltA            EHVGESLVPATTPVLLEIG---VMEKIEKANFPKKFGAAWTSADSGPEDKMGFQGLDHDF
ZoHal           FHIGESLLPMNMPILERLG--------------------VLADVRAIGVLKRGADFPND

Aohal           EPWTFHFGISAKMAGSTSHAYQVERARFDEILLNNAKRKGVVVREGSPVTDVVEDGERVT
AmHal           EPWTFHFGISAKMAGSTSHAYQVERARFDEMLLNNAKRKGVVVREGCAVTDVVEDGERVT
SlHal           EPWTFAFSVSPRMAGPTSVAYQVERSKFDDILLKNARKQGADVREGCSVRGVIEEGERVR
MfPrnC          EHNPTQFTQCVIPELPWGPESHYYRQDVDAYLMHAAIRYGCVVKQKTVIKDYDLSK---T
Chl             EPDPNETSQFRIPSIVGN-AAHFFRQDTDSYMFHAAVRYGCDARQYYRVENIEFDD---G
PCR             ------------------------------------------------------------
PltA            RSAEILFNERKQEGVDRDFTFHVDRGKFDRILLEHAGSLGAKVFQGVEIADVEFLSPG-N
ZoHal           SGGYNTFRFSHALDAKADFAFQVPRAQFDQVLFQRARAAGVDAREQVSVEQVAFDGEQ-P

Aohal           GLRYTDADGNEREVSARFVIDASGNKSRLYSKVGGSRNYS-------EFFRSLALFGYFE
AmHal           GARYTDPDGTEREVSARFVIDASGNKSRLYTKVGGSRNYS-------EFFRSLALFGYFE
SlHal           GLTYADADGNEREIRARYVVDASGNKSRLYNKVGGTREYS-------DFFRSLALFGYFE
MfPrnC          GVAVTTTQG--EHFTARYMIDCGGPRAPLALKFGLREEPCRYKTHSRTLYTHMVGVKPFD
Chl             GVTVSGADG--STVRARYLVDASGFRSPLARQLGLREEPSRLKHHARSIFTHMVGVDAID
PCR             ------------------------------------------------------------
PltA            VIVNAKLGKRSVEIKAKMVVDASGRNVLLGRRLGLREKDP--------VFNQFAIHSWFD
ZoHal           LLQARTVDGGVQQFRPRYLLDASGRDTFLGTRLKLKRANAK--------HQSAALFSHFR

Aohal           GGKRLPAPVSG------NILSVAFDSGWFWYIPLSDTLTSVGAVVRREDAEKIQGDREK-
AmHal           GGKRLPEPVSG------NILSVAFDSGWFWYIPLSDTLTSVGAVVRREDAEKIQGDREK-
SlHal           NGKRMPEPNRF------NILCVAFESGWFWYIPLSDTLTSVGAVVRREMAEKIQGDPEK-
MfPrnC          DIFKPKGQRWR---WHEGTLHHMFHGGWLWVIPFNNHSRATNGLVSVGLQLDPRIHPKT-
Chl             DHVDMPAELRPPVPWNDGTMHHIFERGWMWIIPFNNHPGATNPLCSVGIQLDERRYPARP
PCR             ------------------------KGWVWVIPIKDDLYSVGLVVDRSKSAEVREQGADA
PltA            NFDRKSATQSPDKVDYIFIHFLPMTNTWVWQIPITETITSVGVVTQKQNYTNS-------
ZoHal           GVTRRPGEDAG------NISIYRHAHGWMWLIPLPEDIMSVGAVCYPEYMKTRKGD----
                                        *.* **: :    :.. :

Aohal           ALNALIAECPLISEYLANATRVTTGKYGELRVRKD-YSYQQETYWRPGMILIGDAACFVD
AmHal           ALNTLIAECPLISEYLADATRVTTGRYGELRVRKD-YSYQQETYWRPGMILVGDAACFVD
SlHal           ALRALIDECPMIADYLSDATRVTEGQYGEIRVRKD-YSYHHTTFTRPGMMLVGDAACFVD
MfPrnC          EIPAQQEFDEFLARFPTIAAQFKDARPVRDWVKSDRLQYSSKSTVGDRYCLMLHAAGFID
Chl             DLTPEEEFRSHVDRFPAVQRQLKGARSVREWVRTDRMQYSSSRTVGERWCLMSHAAGFID
PCR             FYSSTLAKCAKAMDILGGAEQVDEVRIVQD------WSYDTEVFSADRFFLCGDAACFTD
PltA            DLTYEEFFWEAVKTRENLHDALKASEQVRPFKKEADYSYGMKEVCGDSFVLIGDAARFVD
ZoHal           ---SEAFLMRTLALNPELNARMLDAERVAPVHATGNYAYECTRMAGPRWLMLGDAYTFVD
                                  .  .              *           :  .* * *
```

Fig 11.

Fig. 11 continued

```
Aohal       PVFSSGVHLATYSALLAARSINSVLAGDLDEKTALNEFEMRYRREYGVFYEFLVSFYQMN
AmHal       PVFSSGVHLATYSALLAARSINSVLAGDLDEKTALNEFELRYRREYGVFYEFLVSFYQMN
SlHal       PVFSSGVHLATYSSLLAARSLNSVLEGKVDEDAAMKEFEARYRREYGVFYEFLVSFYEMH
MfPrnC      PLFSRGLENTSVTIHALAARLIKALRDDDFSPERFEYIDRLQQKLLEHNDDFVSCCYTAF
Chl         PLFLRGLSNTCEIINALSWRLMAALREDDFAVERFAYVEELEQGLLDWNDKLVNNSFISF
PCR         PIFRR-----------------------------------------------------
PltA        PIFSSGVSVALNSARIASGDIIEAVKNNDFSKSSFTHYEGMIRNGIKNWYEFITLYYRLN
ZoHal       PMFSSGVFLAMHGAERGAAMVDAALRAPQSEAKLQRALQRELTRGVDEFIFSLSHRTSLV
            *:*

Aohal       -----VNEESYFWQAKKVTQNQSTDIESFVELIGGVSSGETALT--------AADRIAAR
AmHal       -----VNEESYFWQAKKVTQNQSTDVESFVELIGGVSSGETALT--------AADRIAAR
SlHal       -----RDENSYFWQAKKVTANSEPELQSFVELIGGVSSGEQGLMNT----DALSERLADN
MfPrnC      SDFELWDAFHRLWAVGTMLGQFRLVQAHARFREDRNEAHLDHLDDNPPHLGYLCADMDAY
Chl         SHYPLWNSAFRIWASASVIGGKRILNALTRTKETGDDSHCQALDDNPYPG--LWCPLDFY
PCR         -----------------------------------------------------------
PltA        ------------ILFTAFVQDPRYRLDILQLLQGDVYSGK-------------------R
ZoHal       P-----------------------------------------------------------

Aohal       SAEFAAAVDQMASGDGDNMVPMFKSTVVKQAMQEAGQVQMKALLG---EDAEPELPLFPG
AmHal       SAEFAAAVDEMAGGDGDNMVPMFKSTVVQQAMQEAGQVQMKALLG---EDAEPELPLFPG
SlHal       SEEFATAVEKFAANEDGSSVPLFSSSVVRNAMQEAGQVQMRALLG---EDAEPETPMFPG
MfPrnC      CDLFDAAKAEVESVSEKRASPKEAAARIHALIEAQEFARPLFSFGYCITGANRNLNNSKY
Chl         KEAFDELTELCEAVDAGHTTAEEAARLLEQRVRESDWMLPALGFN------DPDTHHINP
PCR         -----------------------------------------------------------
PltA        LEVLDKMREIIAAVESDPEHLWHKYLGDMQVPTAKPAF----------------------
ZoHal       --RAALACTARLAQRWVWSWPLIASCWLSPALHIATYA----------------------

Aohal       GLVTSPDGMKWLPHHPA----------------------------
AmHal       GLVTSPERMKWLPHHPA----------------------------
SlHal       GLVSSPDGLYWLPATTA----------------------------
MfPrnC      SLVPALRLLHWTQKGAPPEVKKYFDYNPMFSLLKSYVGNRLALALK
Chl         TADKMIRIAEWATGHHRPEIRELLAASAEEVRAAMRVKP-------
PCR         -----------------------------------------------
PltA        -----------------------------------------------
ZoHal       -----------------------------------------------
```

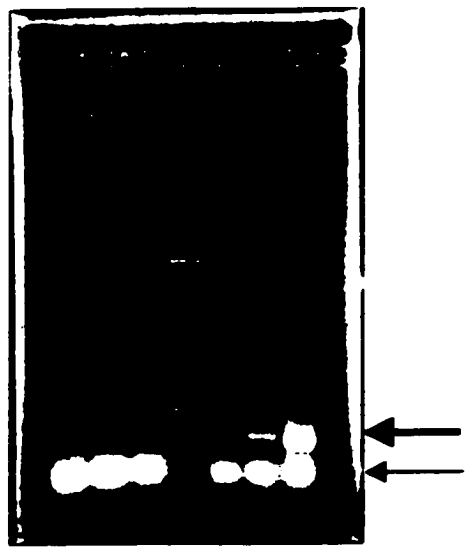
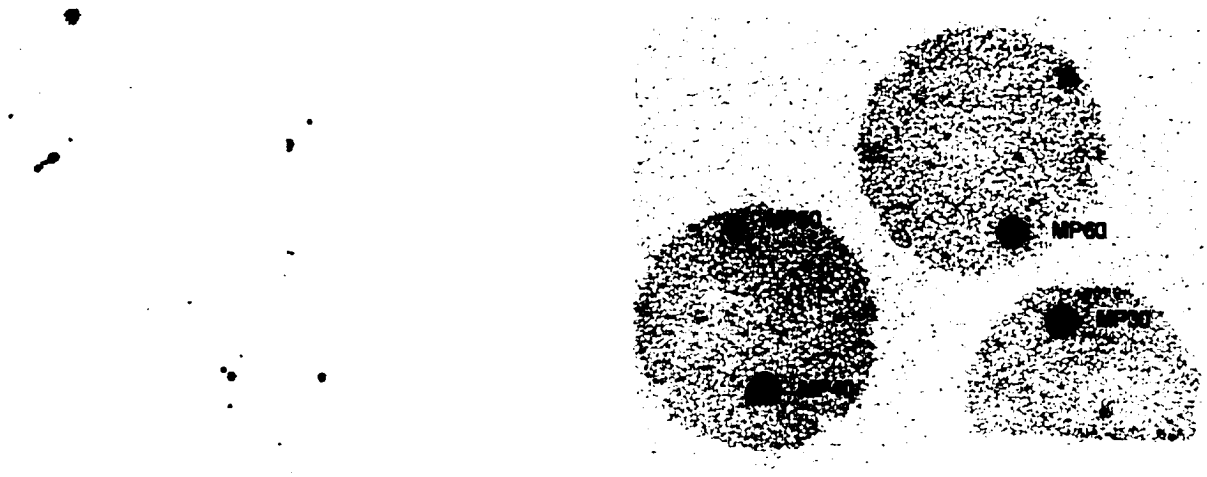


Figure 12. Agarose gel electrophoresis of products from genomic DNA of six streptomycete strains amplified by PCR with primers MPF1 and MPR2. Left-to-right: *Streptomyces akiyoshiensis* (-), *Streptomyces griseoviridis* (+), *Streptomyces glaucescens* (+), DNA Ladder Mix (*MBI Fermentas*), *Streptomyces parvulus* (-), *Streptomyces armentosus* UC2862 (+), and *S. venezuelae* ISP5230 (+).
(-), no amplification; (+), PCR product of about 300 bp (shown by the thicker arrow); the thin arrow identifies excess primers

I. Screening a genomic-library with the PCR probe

The *S. venezuelae* PCR fragment was used as a hybridization probe to detect the full-length gene in a Lambda GEM-11 library of *S. venezuelae* ISP 5230 genomic DNA. Assuming that the genome size of *S. venezuelae* is similar to that of other streptomycete genomes (about 8-10 Mb; Gladek & Zakrzewska, 1984; Hopwood *et al.*, 1985a; Redenbach *et al.*, 1996), and that the genome was randomly cut into fragments of 9-23-kb during construction of the Lambda GEM-11 library (Facey *et al.*, 1996), the number of phage plaques to be screened for a 99.9% probability of finding a specific gene (Hopwood *et al.*, 1985) was about 6000.

Probing the *S. venezuelae* genomic library with the [α-^{32}P]-labeled PCR product yielded several hybridizing plaques (Figure 13-A). Picking and patching the recombinant phages on a lawn of *E. coli* LE392 for a second round of hybridization, yielded four phages (Figure 13-B). Recombinant phages MP30, MP40, MP50, and MP60 at 10^{-4}-10^{-6} dilution were plated and probed again with the labeled-PCR product to obtain pure plaques. DNA from four purified plaques was isolated and analyzed by gel electrophoresis after digestion with *Sac*I.



A) B)

Figure 13. Hybridization signals from recombinant phages probed with the labeled PCR product. **A)** preliminary screen; **B)** well-isolated plaques labeled in a second round of hybridization using the PCR product (pJV501) as the probe.

Since recombinant Lambda GEM-11 DNA has *Sac*I restriction sites at each end of the insert, digestion with *Sac*I released the left and right lambda arms as 20-kb and 9-kb fragments, respectively. The inserts yielded several *Sac*I-*Sac*I fragments, the number depending on the presence of internal *Sac*I sites. Based on the pattern of restriction fragments (Figure 14-A) recombinant phages MP30, MP40 and MP50 DNA contained identical fragments of the *S. venezuelae* chromosome. These included a 6.2-kb *Sac*I-*Sac*I fragment that hybridized with the labeled PCR product and was present in MP60 DNA as well (Figure 14-B).

To identify smaller fragments containing the halogenase gene, the 6.2-kb *Sac*I-*Sac*I fragments from phages MP30 and MP60 DNA were purified and subcloned in pBluescript II SK+ to give pJV502 and pJV503, respectively (Figure 15). Restriction digests of pJV502 and pJV503 (Figure 16, A-I & B-I), probed with the PCR product (Figure 16 A-II & B-II) showed that pJV502 and pJV503 contained the same 6.2-kb *Sac*I-*Sac*I segment of the *S. venezuelae* chromosome subcloned in different orientations. The pJV502 insert was subcloned to find a smaller fragment containing the halogenase sequence. Plasmid pJV502 was digested with *Not*I and a 2.4-kb fragment was excised. The remaining 6.8-kb fragment (Figure 16 A-I lane 2) including the vector and approximately 3.8-kb *Not*I-*Sac*I of *S. venezuelae* chromosome hybridized with the labeled PCR product (Figure 16 A-II lane 2). The linear DNA containing the 3.8-kb *Not*I-*Sac*I segment was purified from an agarose gel, and recircularized to give pJV506 (Figure 17).

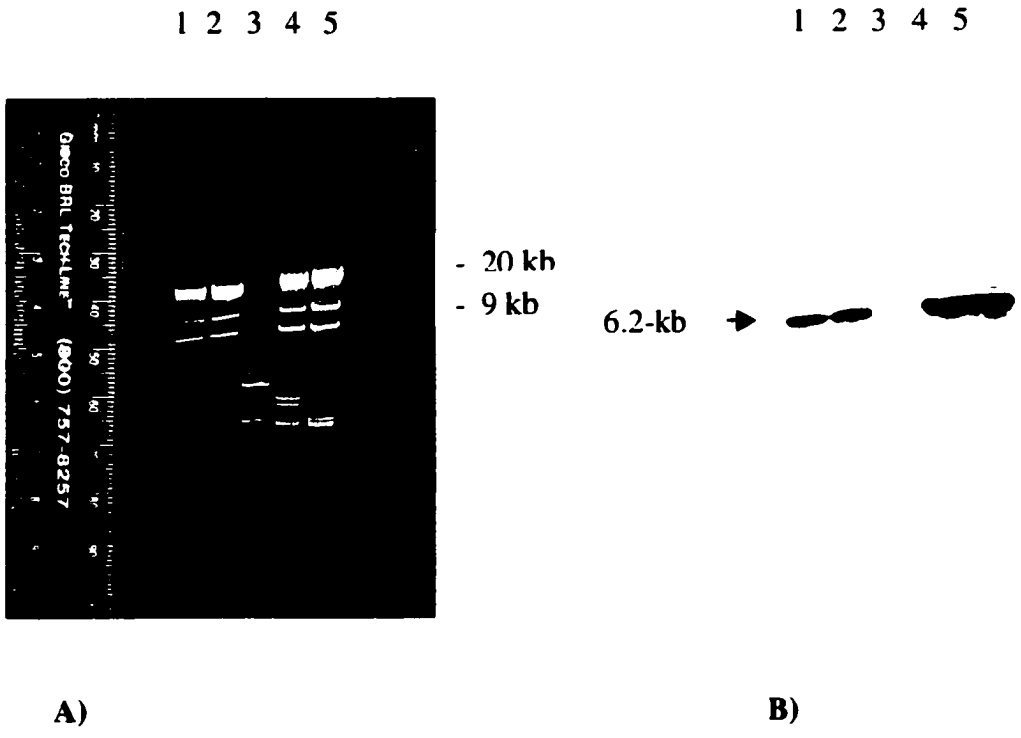


Figure 14. *Sac*I fragments obtained by digesting DNA from recombinant phages: **A)** separated by agarose gel electrophoresis: lane 1, MP30; lane 2, MP40; lane 3, DNA size marker; lane 4, MP50; lane 5, MP60. Bands at 9- and 20-kb are right and left arms of λ GEM-11, respectively. The arrows indicate 6.2-kb.

B) Hybridization signals obtained by probing the *Sac*I digest with labeled PCR product.

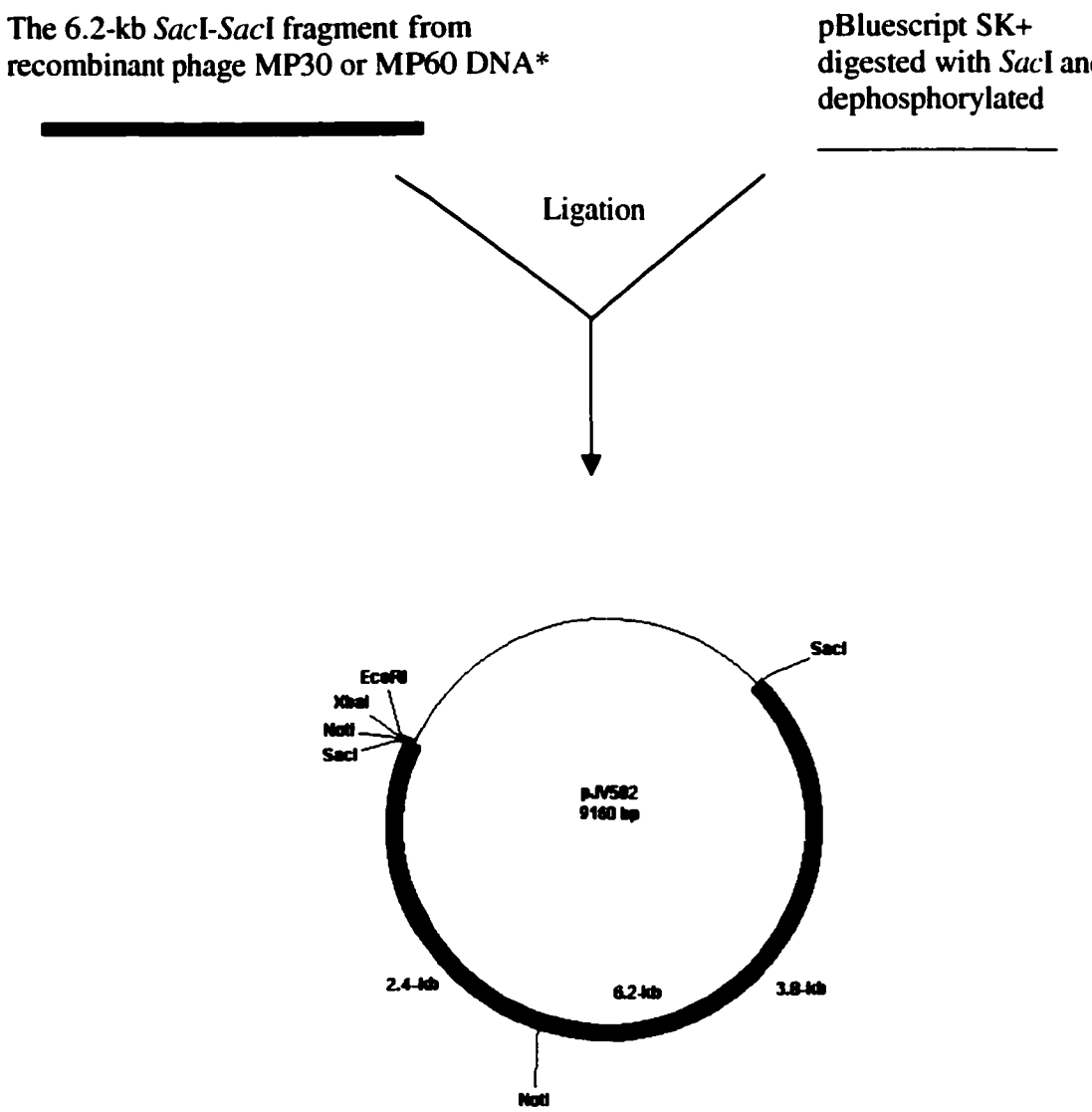


Figure 15. Construction of pJV502 by ligating the 6.2-kb *Sac*I-*Sac*I obtained from recombinant phage MP30 DNA with pBluescript SK+.
pJV503 was constructed by a similar procedure using the 6.2-kb *Sac*I-*Sac*I fragment obtained from phage MP60 DNA.

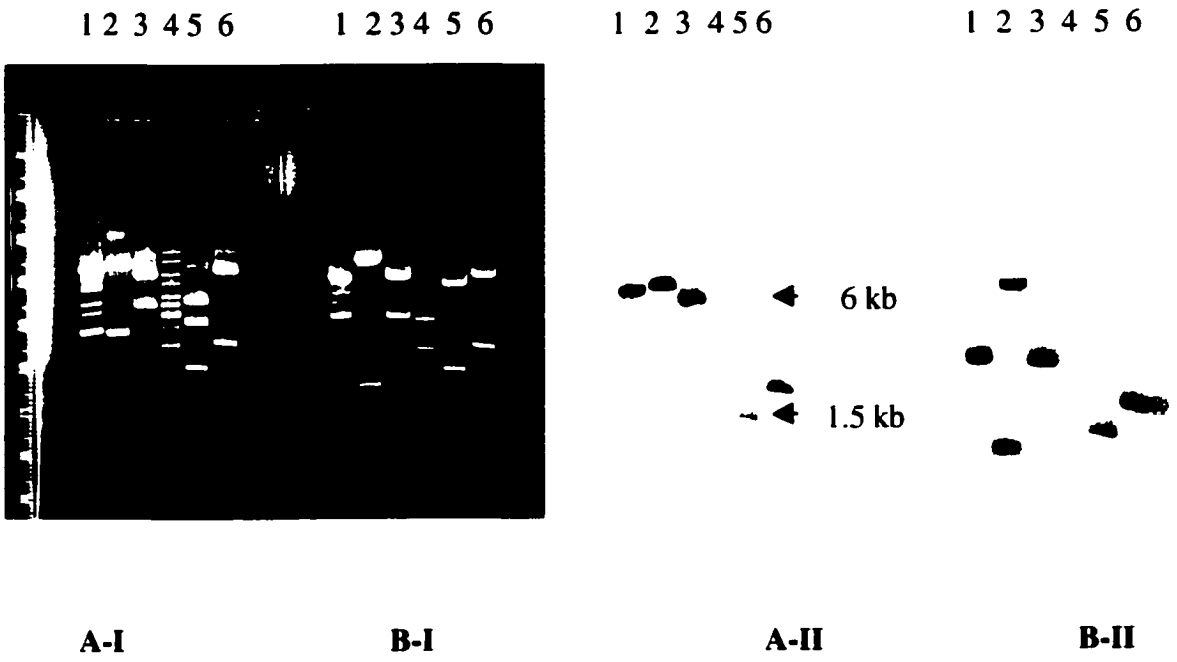


Figure 16. **A-I)** Agarose gel electrophoresis of pJV502 restriction fragments obtained with: lane1, *Bam*HI; lane 2, *Not*I; lane3, *Pst*I; lane 4, DNA marker; lane5, *Sal*I; lane 6, *Bgl*II

B-I) Agarose gel electrophoresis of pJV503 restriction fragments obtained with: lane1, *Bam*HI; lane 2, *Nco*I; lane 3, *Pst*I; lane 4, DNA marker; lane 5, *Sal*I; lane 6, *Bgl*II.

A-II and **B-II)** Hybridization signals from restriction fragments in A-I and B-I probed with the PCR product. The 6.8-kb fragment that remained after digestion of pJV502 with *Not*I (A-I, Lane2; shown by arrow) and the 1.5-kb *Sal*I fragment of pJV502 (A-I, lane 5; shown by arrow), which showed hybridization (A-II, lanes 2 and 5 respectively), were purified from agarose gels.

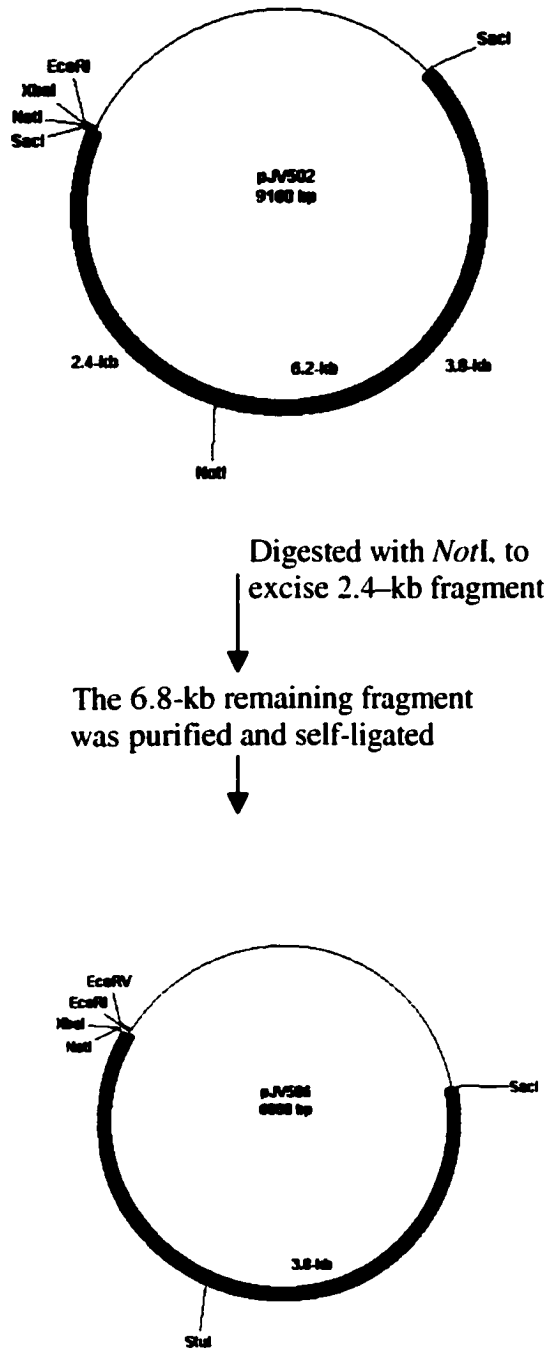


Figure 17. Construction of pJV506 by digesting pJV502 with *Not*I and religating the remaining 6.8-kb fragment.

a. Sequence of pJV504 and pJV506

A 1.5-kb *Sal*I-*Sal*I fragment that hybridized with the PCR product and was present in both pJV502 and pJV503 was subcloned in pUC18 from pJV502 to give pJV504 (Figure 18). Preliminary sequencing of both ends of pJV504 from M13 reverse and forward primers showed high similarities with known halogenases. Digestion of pJV506 with *Sal*I yielded the hybridizing 1.5-kb fragment, indicating that halogenase gene sequences were present in pJV506 (Figure 19). Therefore, subclones and nested deletion clones of the 3.8-kb *Not*I-*Sac*I insert of pJV506 were prepared and the DNA segment was completely sequenced.

b. Analysis of pJV506 sequence

Analyses of the nt sequence of the 3.8 kb *Not*I-*Sac*I insert in pJV506 with the web-based Frameplot program (http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl; Ishikawa & Hotta, 1999) detected three ORFs (Figure 20), all of them were similarly oriented. One (ORF11) was truncated at the 5'-end, but the other two (ORF12, and ORF13) were complete. Partial ORF11 (692 bp) was located at the 5'-end of the 3.8-kb *Not*I-*Sac*I fragment. It encoded the C-terminal region of a protein with high similarity to adenylating enzymes. Moreover, 163 nucleotides of partial ORF11 (from its unique *Bam*HI site to the 5'-end of the 3.8-kb *Not*I-*Sac*I fragment) showed total identity with the 3'-end of the incomplete ORF11 (*CmlK*) reported in *S. venezuelae* (see Figures 21, 22) by He *et al.* (2001). Incomplete ORF11of He *et al.* is a 3'-truncated gene located in a 7.5-kb *Bam*HI-*Bam*HI fragment of the *S. venezuelae* chromosome containing genes involved in Cm biosynthesis.

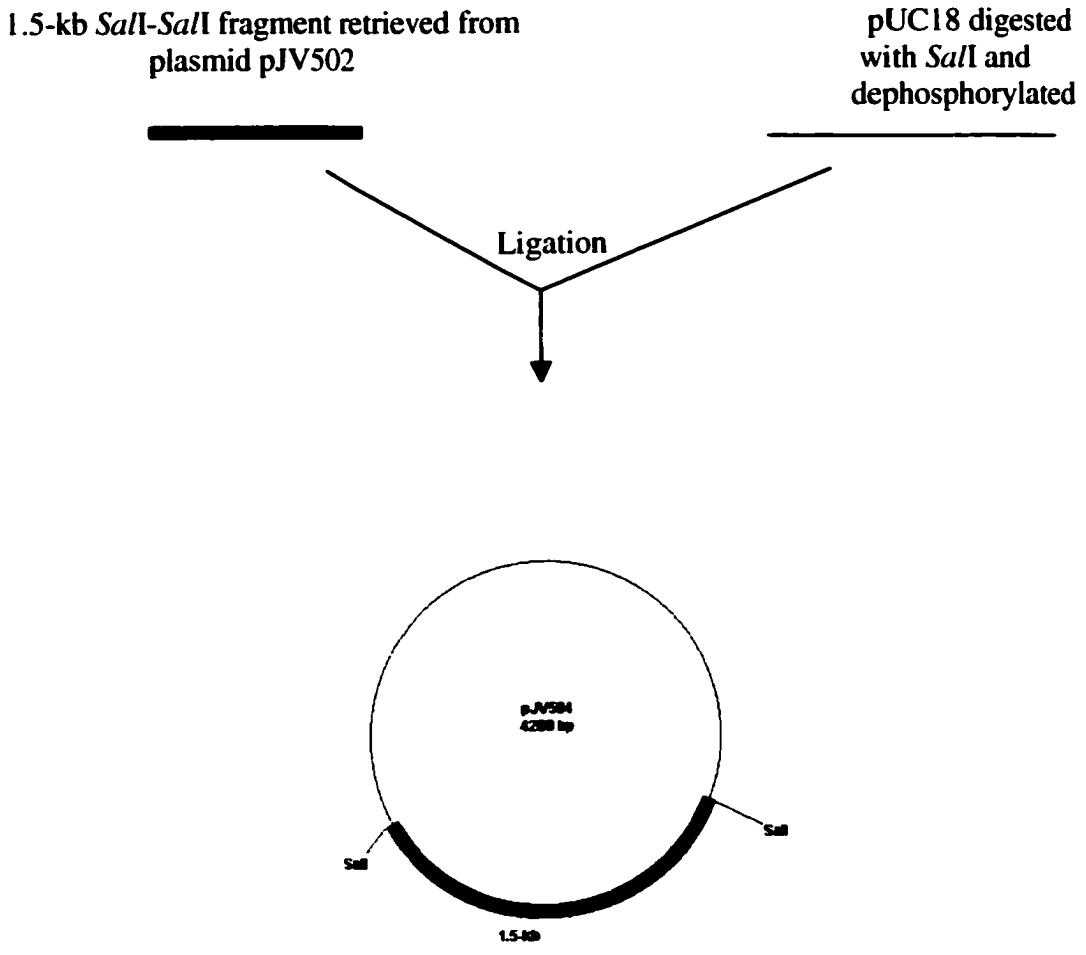


Figure 18. Constructing pJV504 by ligating the 1.5-kb *Sal*I-*Sal*I fragment from pJV502 with pUC18.

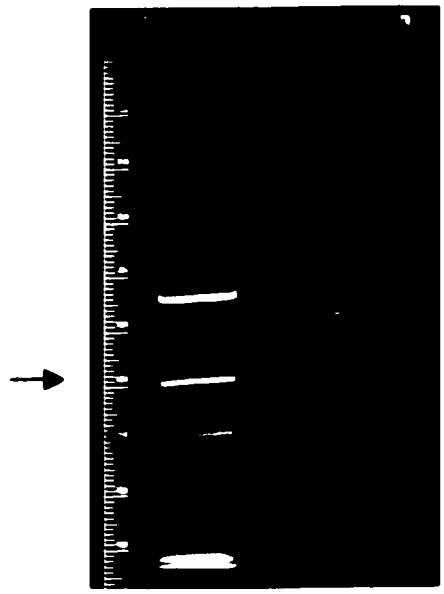


Figure 19. Restriction fragments from *Sal*I digestion of pJV506. The 1.5-kb *Sal*I-*Sal*I fragment indicated by the arrow, was also present in pJV502.

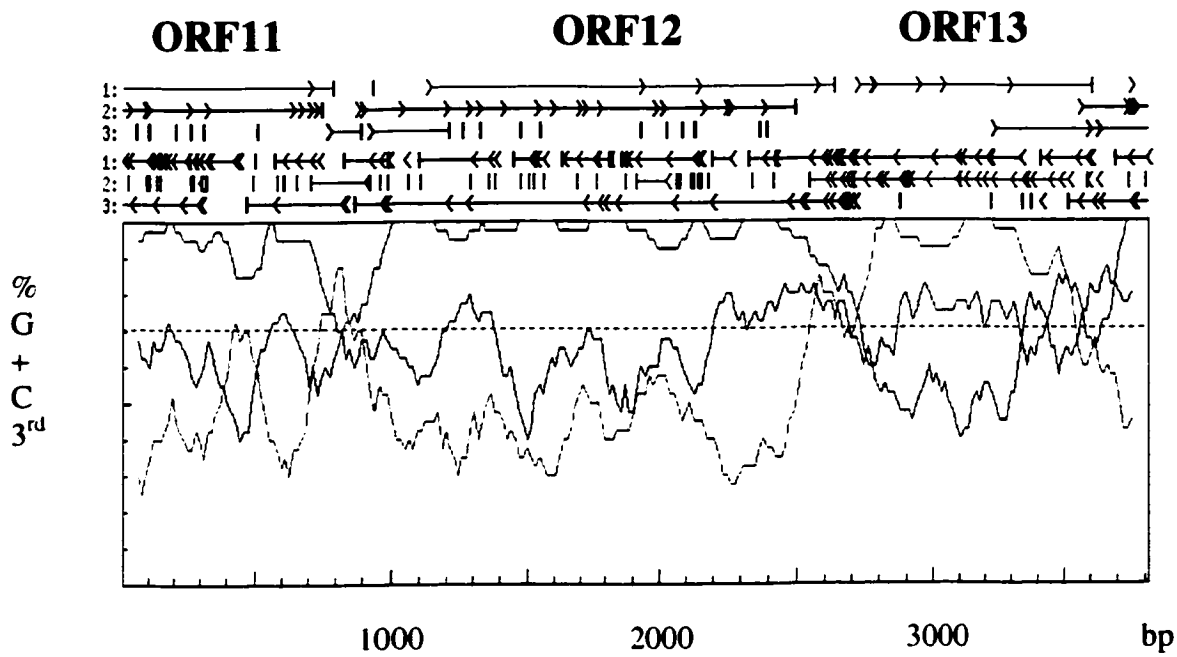


Figure 20. Frameplot analysis of the 3.8-kb *Not*I-*Sac*I fragment (insert of pJV506) containing ORF11 (encoding an AMP-ligase) and ORF12 (encoding a halogenase).
Top panel: Lines identify possible ORFs from relevant start and stop codons. > and < are potential start codons, and | indicates stop codons.
Bottom Panel: Variations in %G+C at the 3rd position of codons in the selected reading frame are plotted; dashed line represents 71% G+C.

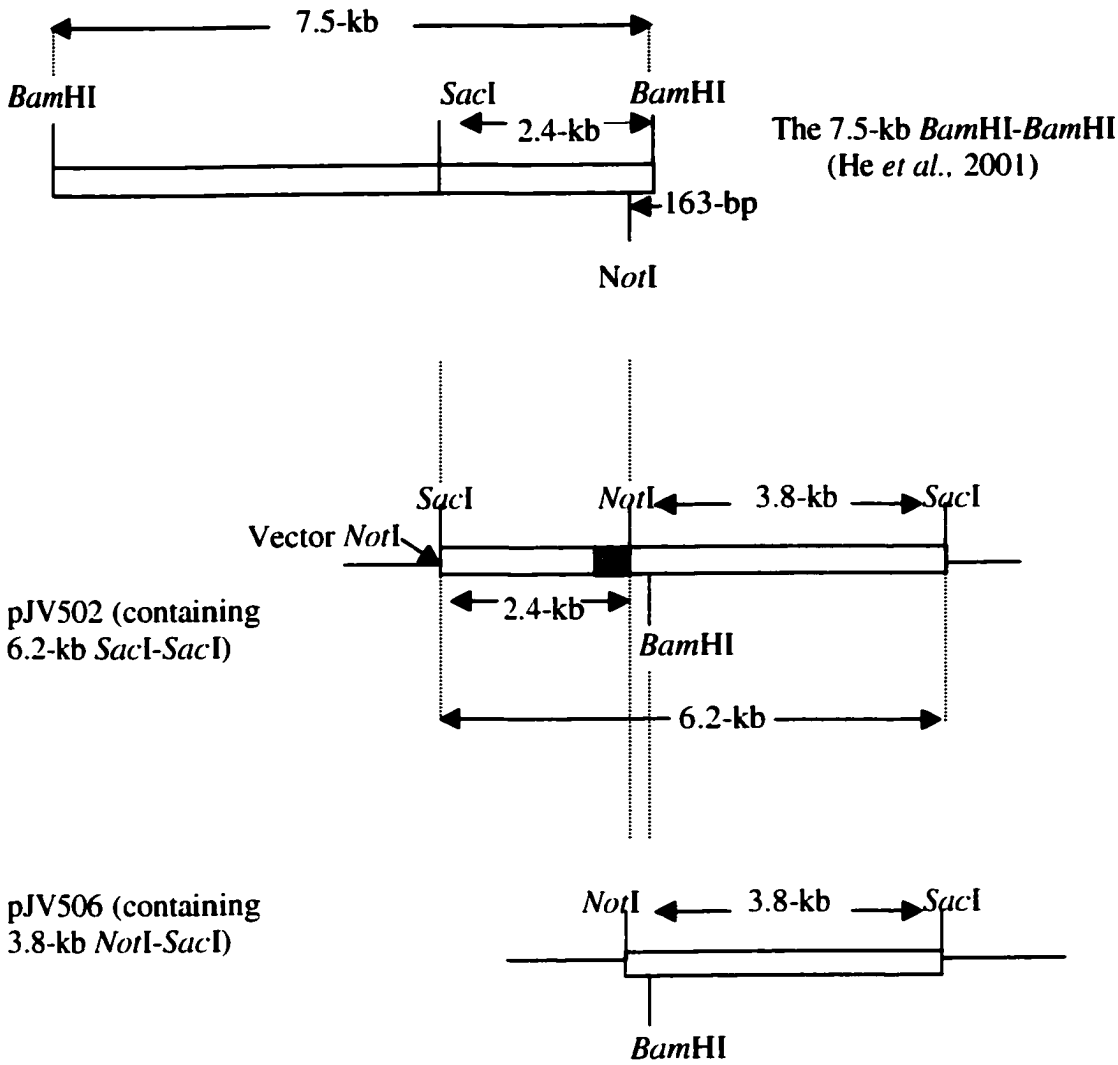


Figure 21. Relationships between pJV502, pJV506 and the overlapping 2.4-kb *Sac*I-*Bam*HI segment of DNA in the 7.5-kb *Bam*HI-*Bam*HI fragment reported by He *et al.* (2001), from which 430 bp, shown by the hatched area, was used to complete the sequence of ORF11 (see Fig. 22).

II. Identification of ORFs

To obtain the complete sequence of the region that included the initiation codon of ORF11, the known sequence (430 bp) determined as part of the 7.5-kb *Bam*HI-*Bam*HI fragment (He *et al.*, 2001) was added to the sequence at the 5'-end of the 3.8-kb *Not*I-*Sac*I fragment, and gave a total sequence of 4177 bp (Figures 20-23).

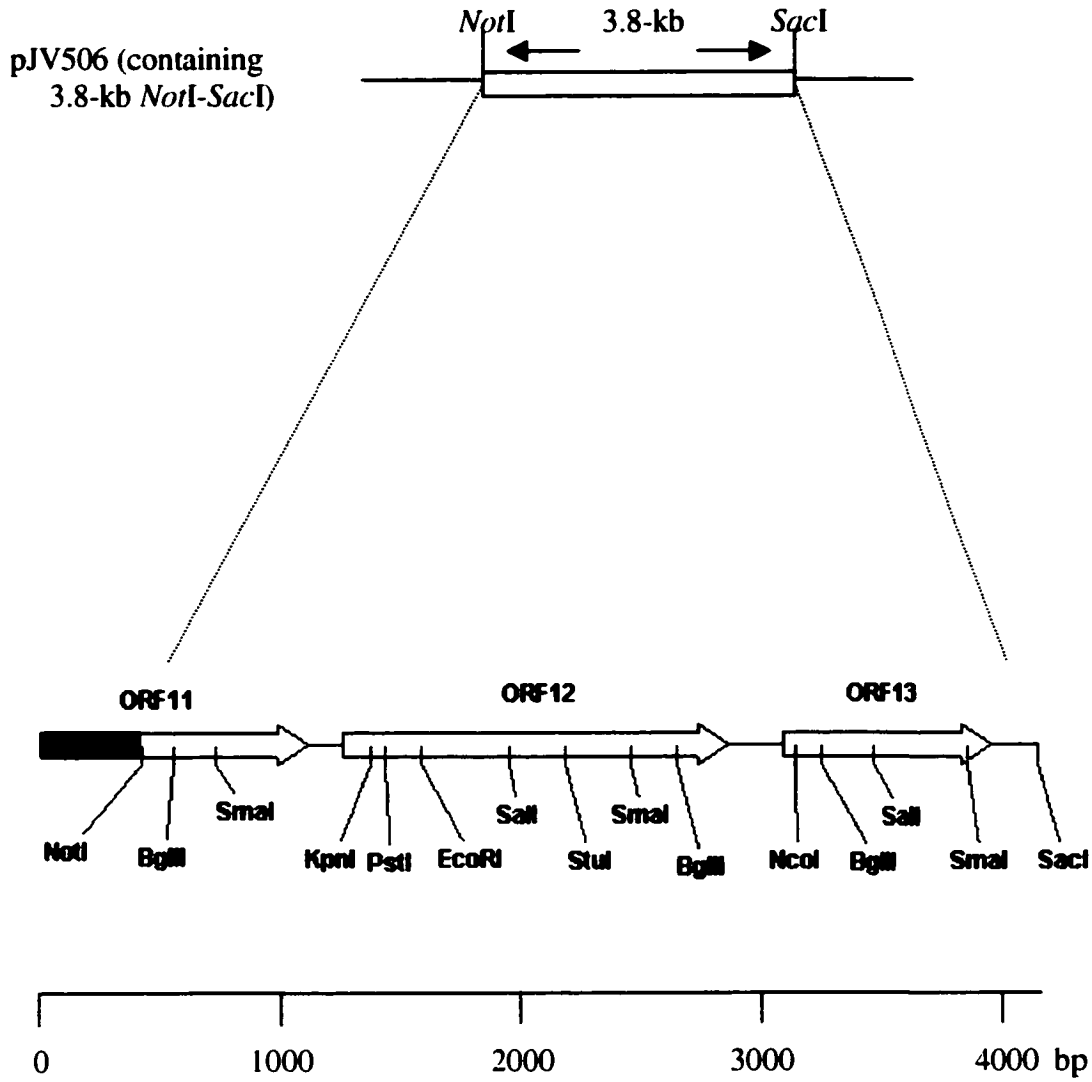


Figure 22. Restriction map of the 3.8-kb *Not*I-*Sac*I DNA fragment cloned in pJV506. Nucleotide sequence for the 5'-end of ORF11 (the hatched area; 430 bp) was obtained from the 7.5-kb *Bam*HI-*Bam*HI fragment (see Fig. 21) reported by He *et al.* (2001).

Figure 23. The 4177-bp nucleotide sequence including the 3.8-kb *Not*I-*Sac*I fragment subcloned in pJV506, and 430-bp added to complete ORF11.
Arrows mark the initiation codons and indicate the direction of ORFs. The plausible RBS for ORF12 and ORF13 are marked with asterisks. The stop codons (TGA) are shown with italic-bold letters. The presumed transcriptional termination sequence of ORF11 is underlined with a thick line. Selected restriction enzymes are marked above the sequences.

ORF11

CTGGCGCTCCACGGC**GTG**AAGCCGCGGGAGACCGTGGTCGTCCACGGCGA 50

CTACTCGCTGGACTCGGTGGCGGCCCTTTTCGCCCTCTTCCTCGACCGGA 100

ACGTCGTGGTGCCGGTGGTCACGCTCACCGACACCACCCTCGACACCCTC 150

GTCGAGCACTGCCGGCGGGCCATCTCATCAGGACGACCGGGGCGGGTACG 200

GAGGTCGAGGACCTCGCCGCACGGGCCGAACCCGCCGCACGGGCCGAACC 250

CGCCGGACCGGCCCAGCTCGCCGGCGCCGAGGAACCCACCGACGGGGCCG 300

AGGCCGTGCACGGCGCGGCTCGGCGACCCTGACGCGGCCGGACTCGTGCT 350

GCTCAGCAGCGGCAGCACCGGCGCGCCCAAGGTCATCCTGCACAGTCTCG 400

*Not*I

ACGTGCTCGTCGGCGAGAAGCTGAAGAAGCGGCCGCGCCGCCGGGAGAAC 450

GCCCTCAACATCCTCATGGTGCTGATGTTCGATCACATCGGGGGCATCAA 500

CTCCCTGCTCAGCACCCTCCTGGTCGGGGGCACCGCGGTCCTCCCCCGCC 550

*BamH*I

AGCGCACCCCGGACGAGATCTGCGCGCTGATCGAGCGCCACCGGATCCTC 600

GTCCTCCCCACCAGCCCGACCTATCTCAACCTGATCATGGTCGGCGACTA 650

CCACCGCACACACGACCTCAGCAGCCTGCGCCTGATCACGTACGGCACCG 700

AGCCGATGTCCGAGGAACTCCTCCTGCGGGTCCGCAAGACCTTCCCGGGG 750

GTCCGGCTGCTCCAGACCTTCGGGACGAGCGAGACCGGCATCGCCACCAC 800

CACCAGCGAATCGTCGACGGCGAACTCCAGCTCAGGAGCCGGACCCAGTT 850

0CCTCGGCTATCTGCTACGACGACGACTCGCTGACCGAGGACCGCTGGTTC 900

CGGACCGGGGACCTGGTCGAGGAGGCGGCCCACGGCTACCTCCGGATCAA 950

GGGCCGGGCCAAGGAGGTCATCAACGTCGGCGGCGAGAAGCTCCTCCCGC 1000

TGGAACTCGAATCGGTCCTCATGGGCAGCCCCCTCGTCGAGGACTGCGTC 1050

GTGTACGGCAGGCCGAACGCGATCACCGGCCAGTCCGTGTGCGTCGACAC 1100

GATCGCGATGTCGGACCGC**TGA**AGAAGCTGCGGTCCCGGCCGGGCGCCGA 1150

CGAGGAGTGAACGGCCTCCTCCCGCCCGTACCCCGCACCCCTTTCACAGC 1200

GCTTCCTCCTCGCCGCCCCGGCGCGGAGTCCCGTCAGCCTGTCCCCGTCC 1250

****** ORF12

C**AAGGAG**TAC**GTGATG**ACACGATCGAAGGTGGCGATCATCGGCGGAGGGC 1300

CGGCCGGTAGTGTCGCGGGCCTCACCCTGCACAAGCTGGGCCACGACGTC 1350

Fig. 23

*Kpn*I

ACGATCTACGAGCGGTCCGCCTTCCCCCGGTACCGCGTCGGCGAGTCCCT 1400

*Pst*I

GCTGCCCGGCACGATGTCGATCCTCAACCGCCTCGGGCTGCAGGAGAAGA 1450

TCGACGCGCAGAACTACGTCAAGAAGCCCTCGGCGACCTTCCTCTGGGGC 1500

CAGGACCAGGCCCCGTGGACGTTCTCCTTCGCGGCCCCCAAGGTCGCGCC 1550

*Eco*RI

CTGGGTCTTCGACCACGCCGTCCAGGTGAAGCGCGAGGAATTCGACAAGC 1600

TCCTCCTCGACGAGGCCCGGAGCCGCGGGATAACCGTCCACGAGGAGACC 1650

CCGGTCACCGACGTGGACCTGTCCGACCCCGACCGCGTCGTCCTGACCGT 1700

GCGCCGGGGCGGCGAGAGCGTCACCGTCGAGAGCGACTTCGTCATCGACG 1750

CGGGCGGCTCGGGCGGTCCGATCTCCCGCAAGCTCGGCGTGCGCCAGTAC 1800

GACGAGTTCTACCGGAACTTCGCGGTCTGGTCGTACTTCAAGCTGAAGGA 1850

CCCCTTCGAGGGGGACCTCAAGGGCACCACGTACTCGATCACCTTCGAGG 1900

ACGGCTGGGTCTGGATGATCCCCATCAAGGACGACCTGTACAGCGTCGGC 1950

CTCGTCGTCGACGAGGTCCGGATCGTGCAGGACTGGTCCTACGACACCGA 2000

GGTCTTCTCCGCCGACCGCTTCTTCCTCTGCGGAGACGCCGCCTGCTTCA 2050

CCGACCCGCTGTTCTCCCAGGGCGTGCACCTGGCCTCGCAGTCGGCGGTG 2100

TCGGCCGCCGCGGCCATCGACCGCATCACCCGGCACGGGGACGAGAAGGA 2150

*Stu*I

CGCGGTGCACGCCTGGTACAACCGCACCTACCGCGAGGCCTACGAGCAGT 2200

ACCACCAGTTCCTCGCCTCCTTCTACACCTTCGCCTCCTTCACGGAGCCC 2250

GACTCCGAGTTCTGGCGGAAGCGGCGCATCACGGAGTCCGACGACGACCG 2300

GCTGACCCGAAAGAAGTGGTTCGAGAGCCTGGCGGGCAACGGCCCGGAGG 2350

ACCCCTCCGGGACCGTGGCGTCCTTCCGCGACCGGGCATCCACGATGATC 2400

GCGATCGGCCGCCACCAGCGTCCGGAACTCAGCGACGACTTCAGCGAGGC 2450

CGAGCTGAACCCGGCCCGGGTCCGCTGGATCAGCGACCTCACCAAGCGGC 2500

TGAACAGCATCACCCGCTTCAAGTGGACCGGCGGCAAGGCCGTGCTCAAG 2550

CAGCACTACCGCGTCGAGCCGATCGGCTTCCGGCTGGAACAGCGCGAGGT 2600

CCTCGCCAACGGGGAAGGGCTCGACATGGCCCAGTACCCGATGGACGACG 2650

*Bgl*II

AGGCCCGGCAGATCTTCCAGGACCTCGCCGAGGAGGAGTTCGGCTACAAG 2700

ACGCTGGTCAAGCGCCTCGGCGCGGTCGGCCGGCAGGAGCTGAGCACCCA 2750

GATCGTCGTCCGGCTGATGGAGGCGGGCCTCCTCACCGGCTACGACGCGC 2800

AGGGCGAGAAGGTCTTCGTCCAGGGCCGGCTCCACTTCGGCGGCGTCGGG 2850

GTCGAGTACGAGGTC**TGA**CCGCCGCCGCACGACCACCCGCCGGGCAGCCC 2900

CGTCCCGTTCACGGGAGCGGACGGGACGGGCCCCGGCGGGCTCGGGGGCC 2950

GGTGCACCCACCGGCCACCACCCACCGGCCACCGCCCACCGGCCACCGGC 3000

CCCCGTAACACCCCCCACACCCCACACCCCTCCCACCCCCCACGCCCCTT 3050

****

ORF13

CCACACCCCACCCCTCATCCACCCTCGCGCCCCGAAGGAGGCGGTCCG**AT** 3100

GCGGCACACGAGATTGCGGGACCTGGAGGTCTCCCGGATCGGCCTGGGGA 3150

*Nco*I

CCATGGGCATGTCCTTCGGCTACACGGGCTCCGGCTCGGACGACGCGGGC 3200

TCCGTCCGCGCCCTGCACCGCGCCCTCGACCTCGGCGTCACCTTCCTCGA 3250

CACCGCCGAGATCTACGGGCCCTACACCAACGAACGGCTCGTCGGCCGAG 3300

CGCTCGGCGCCCGCCGCGACCAGGTGGTGCTCGCCACGAAGTTCGGCCTC 3350

GTCTCGCACGCGGGCGGCGGCCCCGGACAGCTCGACAGCAGCCCCGCGAA 3400

CGTCCGTACCGCCGTGGAGGGTTCGCTGCGGCGGCTCGGGACCGACCACA 3450

TCGACCTCTGCTACCAGCACCGCGTCGACACCGTCGGCGCCCTCGCCGAG 3500

CTGGTCGCCGAGGGCAAGATCCGCCACATCGGCCTCTCCGAGGCGGGCCC 3550

CGCCACGATCCGCCGCGCGCACGCCGTCCACCCGGTCACCGCGGTCCAGT 3600

CCGAGTACTCCCTGTGGACCCGCGACCCCGAGGAGGCGGTCCTGCCGGTC 3650

CTGCGCGAACTGGGCATCGGCTTCGTGGCGTACTCCCCGCTCGGCCACGG 3700

CTTCCTCACCGGCACGGTCCGCGCGGCCGGCCGCTTCACGCCTCGGACCT 3750

CCGGGCGGACAACCCCCGCTTCACCGCGGAGAACCTGCCCCGGACACCTG 3800

CGCATCGCCGACGAGATCGCCGCCGTCGCCGCCGAGGCGGACGCCACCCC 3850

GGCGCAGGTCGCCCTCGCCTGGCTGCTCACCCGGGCTCGCCCCGCCGAGA 3900

GCGAACATCCTTGGCCACGGGGCGAGTTGCGGCAAGGCTGTGGGGGTCAC 3950

ATCCCGCATCCTGGAGGCGTTCCA**TGA**CCGCCGAGCACGCGAACACCCTG 4000

TACGAGGCCGTGGGCGGCGCCGACGCGCTGCGCAGGCTTCCGAGACCTTC 4050

TACCAGGGGGTCCTCGCCGACCCCCTCCTCGCCCCCGTCTTCGCCGACTT 4100

CACCGCCGCCCACGTGGAGCACGTCGCCGTGTGGCTGGCGGAGGTGTTCT 4150

CGGGGCCCGCCGACTTCACCGCGGAGCT 4177

Fig. 23

a. ORF11 (encoding CmlK, an AMP-ligase)

i. Sequence analysis

The complete sequence of ORF11 included an initiation codon and showed a high similarity to proteins in GenBank with AMP-ligase activity that function as activating enzymes. The complete sequence of ORF11 contained 1107 bp with the strong third-letter G+C bias (93.8%) in codons expected for translated streptomycete DNA. The stop codon of ORF11 was located at nt 690-692, inside the 3.8-kb *Not*I-*Sac*I fragment of pJV506 (Figure 20). Immediately downstream of the ORF11 stop codon, the GCG Terminator programme identified repeat sequences (TGAAGAAGCTGCGGTCCCGGCCGGGCGCCGACGAGGAGTGAA) presumed to be involved in transcriptional termination. The deduced amino acid sequence of ORF11 (CmlK; 368 aa; Figure 24) showed high similarity (25% identity over 304 amino acids) to salicylyl-AMP ligase (PchD) of *Pseudomonas aeruginosa* (Quadri *et al.*, 1999), to 2,3-dihydroxybenzoate-AMP ligase (DhbE) of *Bacillus subtilis* (Rowland *et al.*, 1996; 42% identity over 73 aa), and to a homologue of acetate-CoA ligase (EntA) in *Bacillus subtilis* (42% identity over 73 amino acids) involved in activating chorismate for conversion to 2,3-dihydroxybenzoate (Adams & Schumann, 1993). Figure 25 shows an alignment of CmlK with similar proteins identified in GenBank by using CmlK as the query sequence with BLASTP.

```
VKPRETVVVHGDYSLDSVAALFALFLDRNVVVPVVTLTDTTLDTLVEHCR      50

RAISSGRPGRVRRSRTSPHGPNPPHGPNPPDRPSSPAPRNPPTGPRPCTA     100

RLGDPDAAGLVLLSSGSTGAPKVILHSLDVLVGEKLKKRPRRRENALNIL     150

MVLMFDHIGGINSLLSTLLVGGTAVLPRQRTPDEICALIERHRILVLPTS     200

PTYLNLIMVGDYHRTHDLSSLRLITYGTEPMSEELLLRVRKTFPGVRLLQ     250

TFGTSETGIATTTSESSTANSSSGAGPSSSAICYDDDSLTEDRWFRTGDL     300

VEEAAHGYLRIKGRAKEVINVGGEKLLPLELESVLMGSPLVEDCVVYGRP     350

NAITGQSVCVDTIAMSDR                                      368
```

Figure 24. The deduced amino acid sequence (CmlK) of ORF11.

Figure 25. Alignment of the aa sequence of CmlK with sequences of similar proteins with AMP-ligase activity.
DhbE, 2,3-dihydroxybenzoate-AMP ligase of *Bacillus subtilis* (Rowland *et al.*, 1996); EntA, homologue of acetate-CoA ligase in *Bacillus subtilis* (Adams & Schumann, 1993); PchD, salicylyl-AMP ligase of *Pseudomonas aeruginosa* (Quadri *et al.*, 1999); CmlK, AMP-ligase of Cm biosynthesis. Abbreviations are as shown in Figure 11.

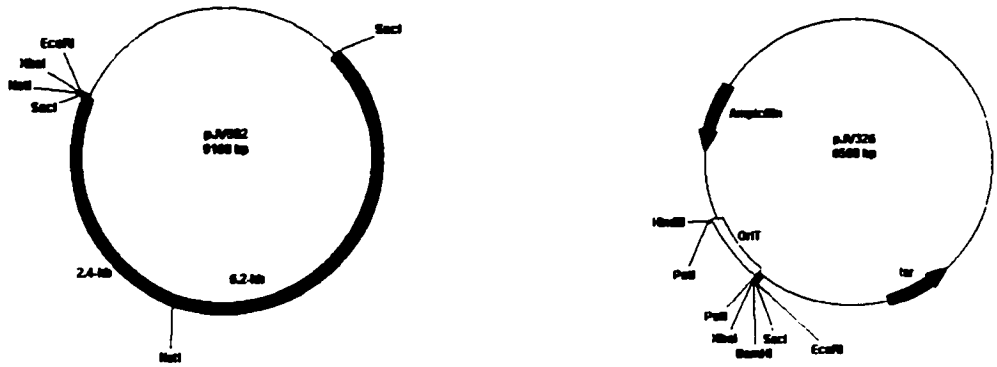
```
DhbE        ---------MLKGFTPWPDELAETYRKNGCWAGETFGDLLRDRAAKYGDRIAITCGNTHW 51
EntA        ------------------------------------------------------------
PchD        MTSSPVTPSAVDDAPDWPAAFVRRYLDAGHWQDQNFAEALAASAARHPRRIALCDDDQRL 60
CmlK        ------------------------------------------------------------

DhbE        SYRELDTRADRLAAGFQKLGIQQMDRVVVQLPNIKEFFEVIFALFRLG--ALPVFALPSH 109
EntA        ------------------------------------------------------------
PchD        SYADLLQRCRRLAAGLRQAGLAHGDTVVLHLPNGIAFVETCFALFQLG--VRPVLALPAH 118
CmlK        --------------------VKPRETVVVHGDYSLDSVAALFALFLDRNVVVPVVTLTDT 40

DhbE        RSSEITYFCEFAEAAAYIIPDAYSGFDYRSLARQVQSKLPTLKNIIVAGEAEEFLPLEDL 169
EntA        ------------------------------------------------------------
PchD        RQHEISGFCRFAEAKAYIGAERIDGFDPRPMARELLASGACRMALIHGEAEAPLQALAPL 178
CmlK        TLDTLVEHCRRAISSGRPGRVRRSRTSPHGPNPPHGPNPP--------DRPSSPAPRNPP 92

DhbE        H-AEPVKLPEVKSSDVAFLQLSGGSTGLSKLIPRTHDDYIYSLKRSVEVCWLDHSTVYLA 228
EntA        ------------------------------------------------------------
PchD        YQADALEDCAARAEDIACFQLSGGTTGTPKLIPRRHREYLYNVRASAEVCGFDEHTVYLT 238
CmlK        TGPRPCTARLGDPDAAGLVLLSSGSTGAPKVILHSLDVLVG---EKLKKRPRRRENALNI 149

DhbE        ALPMAHNYPLSSPGVLGVLYAGGRVVLSPSPSPDDAFPLIEREKVTITALVPPLAMVWMD 288
EntA        ------------------------------------------------------------
PchD        GLPMAHNFTLCCPGVIGTLLAGGRVVVSQRADPEHCFALIARERVTHTALVPPLAMLWLD 298
CmlK        LMVLMFDHIGGINSLLSTLLVGGTAVLPRQRTPDEICALIERHRILVLPTSPTYLNLIMV 209

DhbE        AASSRRDDLSSLQVLQVGGAKFSAEAARRVKAVFGCTLQQVFGMAEGLVNYTRLDDPEEI 348
EntA        ------------------------------------------------R LDDPEEI 8
PchD        AQESRRADLSSLRLLQVGGSRLGSSAAQRVEPVLGCQLQQVLGMAEGLICYTRLDDPPER 358
CmlK        GDYHRTHDLSSLRLITYG----------------------------------T EPMSEE 234
                                                                 :   *

DhbE        IVNTQGKPMSPYDEMRVWDDHDRDVKPGETGHLLTRGPYTIRGYYKAEEHNAASFTEDGF 408
EntA        IVNTQGKPMSPIDEMRVWDDHDRDVKPGETGHLLTRGPYTIRGYYKAEEHNAASFTEDGF 68
PchD        VLHTQGRPLSPDDEVRVVDAEGREVGPGEVGELTVRGPYTIRGYYRLPEHNAKAFSADGF 418
CmlK        LLLRVRKTFPGVRLLQTFGTSETGIATTTSESSTANSSSGAGPSSSAICYDDDSLTEDRW 294
            ::    :.:.    ::. .    : .       ....          ::  ::: * :

DhbE        YRTGDIVRLTRDGYIVVEGRAKDQINRGGEKVAAEEVENHLLAHPAVHDAAMVSMPDQFL 468
EntA        YRTGDIVRLTRDGYIVVEGRAKDQINRGGEKVAAEEVENHLLAHPAVHDAAMVSMPDQFL 128
PchD        YRTGDRVSRDKDGYLVVEGRDKDQINRGGEKIAAEEVENLLIAHPQVHDATVVAMPDSLL 478
CmlK        FRTGDLVEEAAHGYLRIKGRAKEVINVGGEKLLPLELESVLMGSPLVEDCVVYGRPNAIT 354
            :**** *    .**: ::** *: ** ****: . *:*. *:. * *.*..: . *: :

DhbE        GERSCVFIIPRDEAPKAAELKAFLRERGLAAYKIPDRVEFVESFPQTGVGKVSKKALREA 528
EntA        GERSCVFIIPRDEAPKAAELKAFLRERGLAAYKIPDRVEFVESFPQTGVGKVSKKALREA 188
PchD        GERTCAFVIPRQPAPSALKLKQYLHACGLAAFKVPDRIELVPAFPQTGIGKISKKDLRER 538
CmlK        GQSVCVDTIAMSDR---------------------------------------------- 368
            *:  *.  *. .

DhbE        ISEKLLAGFKK 539
EntA        ISEKL------ 193
PchD        LRRELEARA-- 547
CmlK        -----------
```

Fig. 25

ii. Functional analysis of CmlK

To investigate the function of CmlK in Cm biosynthesis, the wild-type gene was replaced by an insertionally-inactivated copy through allele exchange between homologous regions of DNA. A vector suitable for disrupting CmlK was constructed by subcloning a 6.2-kb *Sac*I-*Sac*I fragment of the *S. venezuelae* chromosome obtained from pJV502 in the multiple cloning site of pJV326. The plasmid obtained (pJV526; Figure 26) was linearized at its unique *Not*I site inside ORF11 and ligated with the 1.5-kb Am^R gene excised with *Not*I from pJV225 (Chang, 1999), giving pJV527 (Figure 27). The pJV527 construct was conjugally transferred from its *E. coli* host to *S. venezuelae* ISP5230, generating the transconjugant VS1101.

Southern hybridization experiments were carried out to confirm that a double-crossover event in the mutant strain VS1101 had replaced the native ORF11 with its disrupted allele. For this, genomic DNA extracted from the Am^RTs^S strain VS1101 was completely digested with *Bgl*II. The restriction fragments were probed with the 400-bp *Sal*I-*Not*I fragment from pJV506 containing part of ORF11. The probe hybridized with a band of about 4-kb, corresponding in size to the insertionally inactivated ORF11. In contrast, genomic DNA from *S. venezuelae* ISP5230 gave the expected signal at about 2.6 kb. When the 1.5-kb Am^R gene was used as a probe, strain VS1101 gave the signal at about 4 kb confirming that its DNA contained the Am^R gene, while the wild-type strain gave no signal (Figure 28). The results indicated that the disrupted copy of ORF11 had replaced the wild-type gene by homologous recombination through a double-crossover event that had excised the vector.



pJV502 digested with *Sac*I and the 6.2-kb *Sac*I-*Sac*I fragment purified

digestion of pJV326 with *Sac*I and dephosphorylation

Ligation

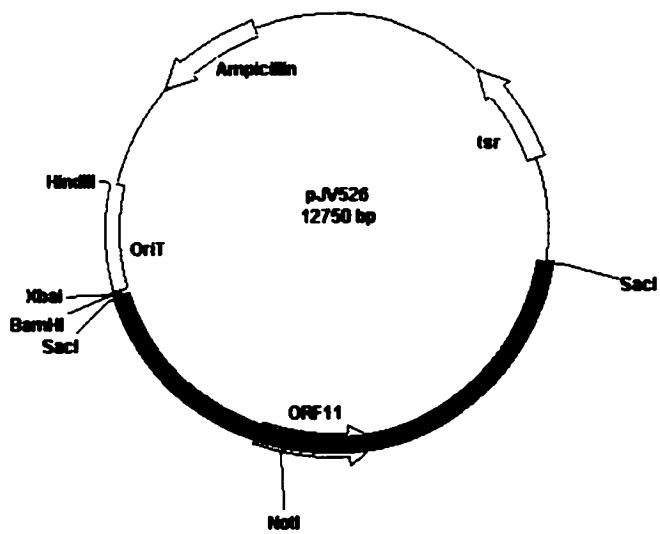


Figure 26. Construction of pJV526 by ligating the 6.2-kb *Sac*I-*Sac*I fragment from pJV502 into pJV326.

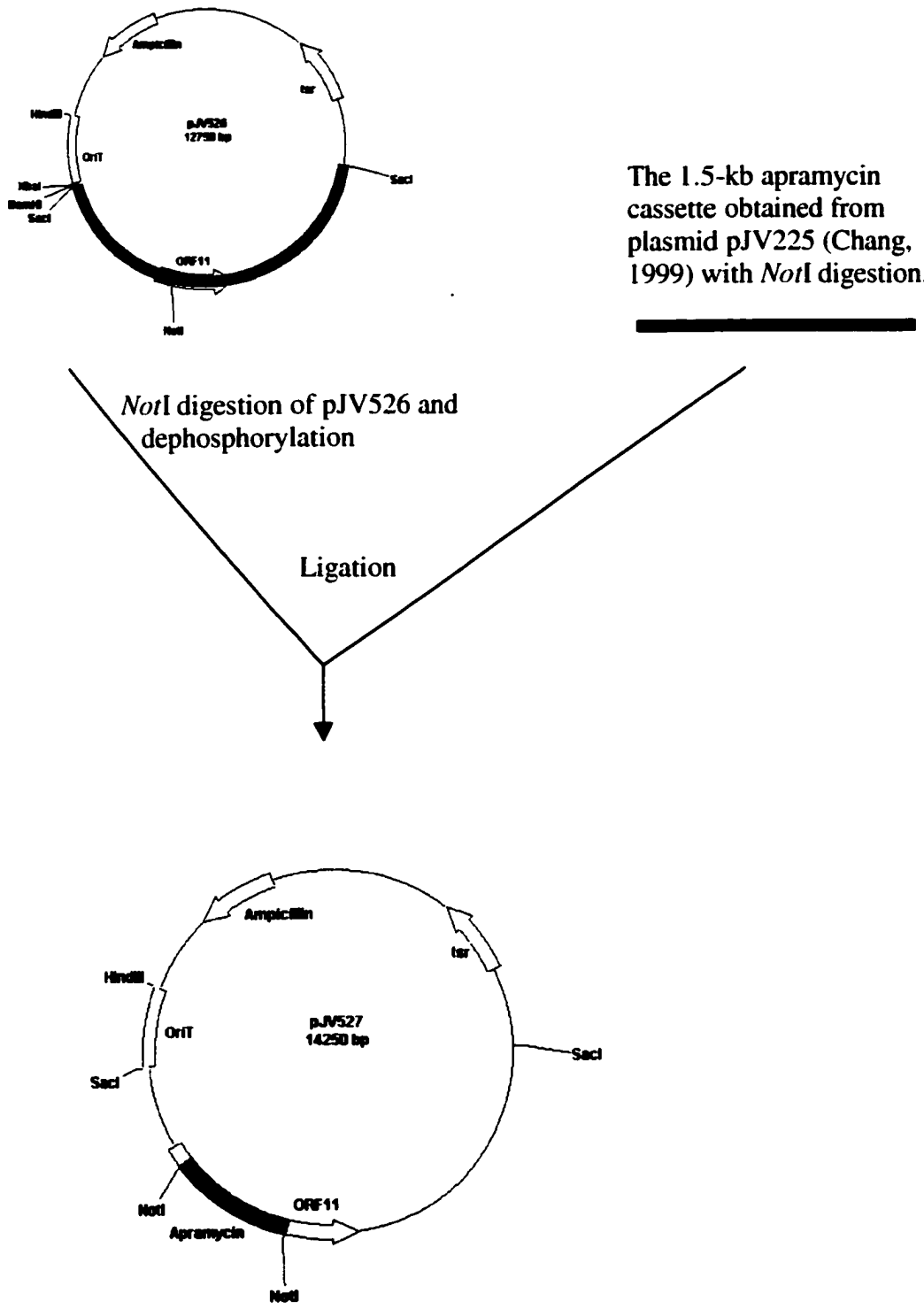


Figure 27. Construction of disruption plasmid pJV527 by ligating a *Not*I-apramycin cassette into the unique *Not*I site inside ORF11.

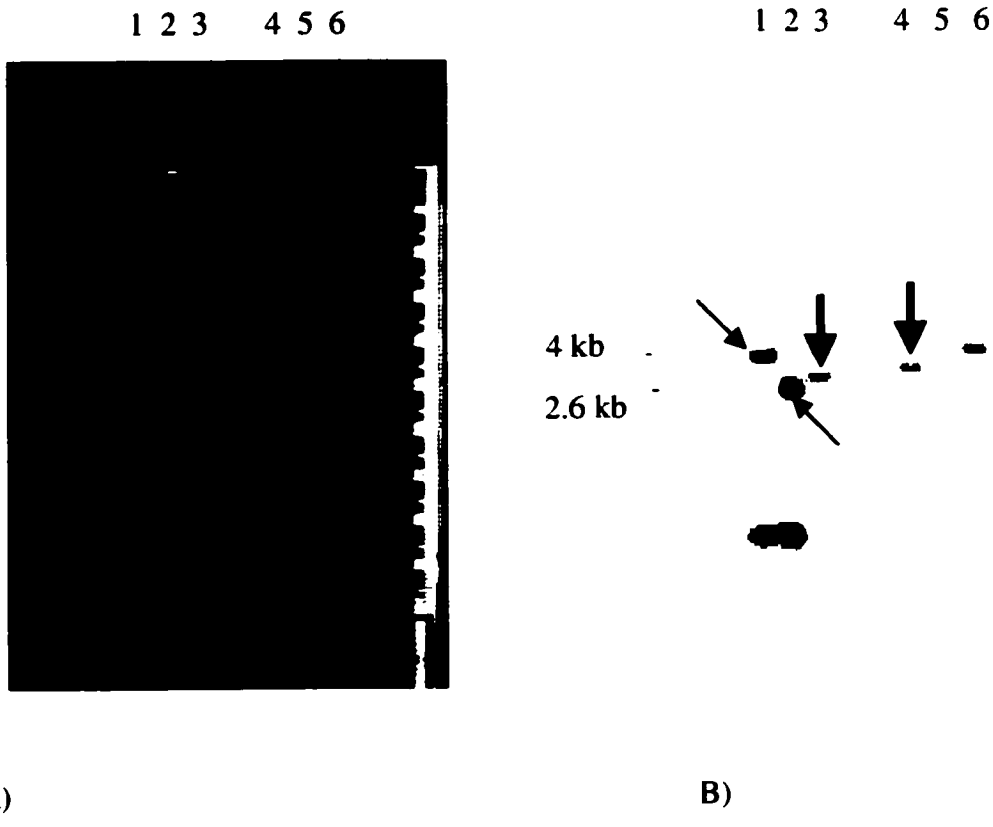


Figure 28. Evidence from Southern hybridization that the genome of strain VS1101 contains the Am^R gene inserted by double crossover.
A) Agarose gel electrophoresis of genomic DNA samples digested with *Bgl*II.
B) Hybridization of the Southern blots of (A) probed with the ^{32}P-labeled *Sal*I-*Not*I fragment of pJV506 containing part of ORF11. The size of the hybridizing *Bgl*II fragment from VS1101 (lane 1) was 1.5-kb larger than that from the wild-type (lane 2), as indicated by the thin arrows. No hybridization occurred in lane 5 when the blot was probed with the Am^R gene whereas lane 6 gave a signal at about 4-kb.
Hybridization signals in lanes 3 and 4 (indicated by thick arrows) are from a 3.28-kb marker fragment

Lanes: 1 and 6, mutant VS1101; 2 and 5, *S. venezuelae* ISP5230; 3 and 4, DNA size marker.

b. ORF12 (encoding CmlS, a halogenase)

i. Sequence analysis

A second open reading frame (ORF12) was identified in the 3.8-kb *Not*I-*Sac*I fragment downstream of the ORF11 stop codon. Whether this ORF starts with the GTG at nt 832-834 located 101 bp downstream of ORF11, or with the ATG at nt 835-837 located 103 bp downstream of ORF11, (Figure 20) is not certain. However, a putative Shine-Dalgarno sequence (AAGGAG; ΔG = -7.5 kcal mol^{-1}) was present 6 bp upstream of the ATG initiation codon, suggesting this as the more likely start of the ORF12 coding sequence. The first in-frame stop codon was at nt 2437-2439, giving a 1608- or 1605-bp gene encoding a protein (CmlS) of 535 (or 534) aa. The frequency of codons containing G or C in the third position was 97%, typical of *Streptomyces* coding sequences.

In restriction digests of *S. venezuelae* ISP5230 genomic DNA, the sizes of fragments that hybridized with the labeled PCR product (see Figure 9) were consistent with the locations of restriction sites in the 3.8-kb *Not*I-*Sac*I fragment (see Figure 23). The presence in *Bam*HI and *Kpn*I digests of *S. venezuelae* genomic DNA of 13- and 15-kb hybridizing fragments, respectively (Figure 9; lanes 1, 2), indicated that the cloned fragment contained a single site for each enzyme. Likewise, the presence of a 1.5-kb hybridizing fragment in the *Sal*I digest of genomic DNA (see Figure 9, lane 3) was compatible with two *Sal*I restriction sites 1.5-kb from each other, one inside ORF12 and the other in ORF13 on the chromosomal restriction map (see Figure 22).

The deduced amino acid sequence of CmlS (Figure 29) exhibited 38% identity (over 231 aa) to the halogenase from *Amycolatopsis orientalis* (GenBank accession number

```
          *  *  *  *     *
VMTRSKVAIIGGGPAGSVAGLTLHKLGHDVTIYERSAFPRYRVGESLLPG  50

TMSILNRLGLQEKIDAQNYVKKPSATFLWGQDQAPWTFSFAAPKVAPWVF 100

DHAVQVKREEFDKLLLDEARSRGITVHEETPVTDVDLSDPDRVVLTVRRG 150

GESVTVESDFVIDAGGSGGPISRKLGVRQYDEFYRNFAVWSYFKLKDPFE 200

GDLKGTTYSITFEDGWVWMIPIKDDLYSVGLVVDEVRIVQDWSYDTEVFS 250

ADRFFLCGDAACFTDPLFSQGVHLASQSAVSAAAAIDRITRHGDEKDAVH 300

AWYNRTYREAYEQYHQFLASFYTFASFTEPDSEFWRKRRITESDDDRLTR 350

KKWFESLAGNGPEDPSGTVASFRDRASTMIAIGRHQRPELSDDFSEAELN 400

PARVRWISDLTKRLNSITRFKWTGGKAVLKQHYRVEPIGFRLEQREVLAN 450

GEGLDMAQYPMDDEARQIFQDLAEEEFGYKTLVKRLGAVGRQELSTQIVV 500

RLMEAGLLTGYDAQGEKVFVQGRLHFGGVGVEYEV                 535
```

Figure 29. The deduced amino acid sequence (CmlS) of ORF12. The putative NAD(P)H binding site that matches the consensus GxGx$_2$(G/A)x$_3$(G/A)x$_6$G is marked with asterisks. Aspartate-259 (D) and other conserved residues of the putative FAD binding site that resembles the consensus sequence GxxxxxGDAxHxxxPxxxxGxxxxxxD are underlined.

AAB49291), 32% identity (over 249 aa) to PltA from *P. fluorescens* (Nowak-Thompson *et al.*, 1999), and 32% identity (over 228 aa) to a putative reductase/halogenase from *Xanthomonas oryzae* pv. *oryzae* (accession number AY010120). CmlS also exhibited 29% identity to PrnC from *Myxococcus fulvus* and *P. fluorescens* (Kirner *et al.*, 1998), and 27% identity to PltM from *P. fluorescens* (Nowak-Thompson *et al.*, 1999). Figure 30 compares the sequence of CmlS with halogenase sequences retrieved from GenBank.

CmlS contains the characteristic motif sequence GxGx$_2$(G/A)x$_3$(G/A)x$_6$G (Figure 29), which forms the β–α–β structure of the NAD(P)H-binding site (Scrutton *et al.*, 1990). The NAD(P)H binding sites of halogenases are compared in Figure 31.

A putative FAD-binding site can be recognized in the central region of CmlS. It resembles the consensus (GxxxxxGDAxHxxxPxxxxGxxxxxxD) found in FAD-dependent monooxygenases such as the *p*-hydroxybenzoate hydroxylases from *Acinetobacter calcoaceticus* (AcPobA) and *P. fluorescens* (PfPobA) (Dimarco *et al.*, 1993). Figure 32 compares the putative FAD binding site of CmlS with similar sites in other halogenases and two hydroxylases.

The N-terminal region of CmlS showed sequence similarity to a number of oxidoreductases, such as a probable electron transfer chain enzyme (GenBank T34627) from *S. coelicolor*, another such enzyme from *Sulfolobus solfataricus* (GenBank AAK42886), and the FAD-binding monooxygenase from *Mycobacterium tuberculosis* (GenBank AAK46067). This suggests that CmlS may include a similar function, which I postulate to be required for reduction of FAD by the halogenase itself.

Figure 30. Alignment of the CmlS sequence with the sequences of halogenases reported from other microorganisms.
AmHal, halogenase homologue from *Amycolatopsis mediterranei*; Xhal, putative reductase/halogenase from *Xanthomonas oryzae* pv. oryzae. (accession AAG38844); MfPrnC, halogenase PrnC from *Myxococcus fulvus*; PfPrnC, halogenase PrnC from *P. fluorescens* Pf-5; PltM, PltD and PltA, halogenases from *P. fluorescens*; CmlS, Cm halogenase from *S. venezuelae* ISP5230. Abbreviations are as shown in Figure 11.

```
AmHal   ------------------------------MSVEDFDVVVAGGGPGGSTVATLVAMQGH 29
Xhal    MHDARPLQVRSRMTSNAVRHTPTAMPAAGPAPGPECPDVLIVGGGPAGCTAAIALAELGW 60
MfPrnC  ----------------------MKPTVNAHHDSNHFDVIILGSGMSGSQMGAILGRQGF 37
PfPrnC  ----------------------MTQKSPANEHDSNHFDVIILGSGMSGTQMGAILAKQQF 38
pltM    --------------------------------MNQYDVIIIGSGIAGALTGAVLAKSGL 27
PltD    ------------------------MNDVQSGKAPEHYDILLAGNSISVIMLAACLARNKV 36
CmlS    --------------------------------VMTRSKVAIIGGGPAGSVAGLTLHKLGH 27
PltA    -----------------------------MSDHDYDVVIIGGGPAGSTMASYLAKAGV 29
                                                .: : *.. .    .  :

AmHal   RVLLLEKEVFPRYQIGESLLPATVHGVCRMLGISDELANAGFPIKRGGTFRWGARPEPWT 89
Xhal    SVTLLEKEQHPRFHIGESLLPMNMPILERLGVLADVRAIG--VLKRGADFPNDSGG---Y 115
MfPrnC  RVLIVEESTHPRFTIGESSIPETSLMNRIIADRYGVPEIEDITSFYSTFKKVSSSTG-IK 96
PfPrnC  RVLIIEESSHPRFTIGESSIPETSLMNRIIADRYGIPELDHITSFYSTQRYVASSTG-IK 97
pltM    NVLILDSAQHPRFSVGEAATPESGFLLRLLSKRFDIPEIAYLSHPDKIIQHVGSSACGIK 87
PltD    RVGLLRNRQMPPDLTGEATIPYTSMIFELIADRYGVPEIKNIARTRDIQQKVMPSSG-VK 95
CmlS    DVTIYERSAFPRYRVGESLLPGTMSILNRLGLQEKIDAQNYVKKPSATFLWGQDQAP--- 84
PltA    KCAVFEKELFEREHVGESLVPATTPVLLEIGVMEKIEKANFPKKFGAAWTSADSGPEDKM 99
           :           **:  * .      :

AmHal   FHFGISAKMAGS----------------TSHAYQVERARFDEMLLNNAKRKGVVVREGCA 133
Xhal    NTFRFSHALDAK----------------ADFAFQVPRAQFDQVLFQRARAAGVDAREQVS 159
MfPrnC  RNFGFVFHKPGEEHNPTQFTQCVIPELPWGPESHYYRQDVDAYLMHAAIRYGCVVKQKTV 156
PfPrnC  RNFGFVFHKPGQEHDPKEFTQCVIPELPWGPESHYYRQDVDAYLLQAAIKYGCKVHQKTT 157
pltM    LGFSFAWHQENAPSSP----DHLVAPPLKVPEAHLFRQDIDYFALMIALKHGAESRQNIK 143
PltD    KNLGFIYHQRSRAVDLG--QALQFNVPSEHGENHLFRPDIDAYLLAAAIGYGAQLVEIDN 153
CmlS    WTFSFAAPKVAP--------------WVFDHAVQVKREEFDKLLLDEARSRGITVHEETP 130
PltA    GFQGLDHDFRSAEILFN---ERKQEGVDRDFTFHVDRGKFDRILLEHAGSLGAKVFQGVE 146
           :                                 :  *  .*   :  *   *    :

AmHal   VTDVVEDGERVTGARYTDPDGTEREVSARFVIDASGNKSRLYTKVGGSRNYSEFFRS-LA 192
Xhal    VEQVAFDGEQPLLQART-VDGGVQQFRPRYLLDASGRDTFLGTRLKLKR-ANAKHQS-AA 216
MfPrnC  IKDYDLSKTGVAVTTTQ-----GEHFTARYMIDCGGPRAPLALKFGLREEPCRYKTHSRT 211
PfPrnC  VTEYHADKDGVAVTTAQ-----GERFTGRYMIDCGGPRAPLATKFKLREEPCRFKTHSRS 212
pltM    IESISLNDDGVEVALSN-----AAPVKAAFIIDAAAQGSPLSRQLGLRTTEG-LATDTCS 197
PltD    SPEVLVEDSGVKVATAL-----GRWVTADFMVDGSQGGQVLARQAGLVSQASTQKTRTLE 208
CmlS    VTDVDLSDPDRVVLTVR-RGGESVTVESDFVIDAGGSGGPISRKLGVRQYDEFYRNF--A 187
PltA    IADVEFLSPGNVIVNAK-LGKRSVEIKAKMVVDASGRNVLLGRRLGLREKDPVFNQFAIH 205
           .                       .    ::* .     :  :

AmHal   LFGYFEGGKRLPEPVSG-----------NILSVAFDSGWFWYIPLS------DTLTSVG 234
Xhal    LFSHFRGVTRRPGEDAG-----------NISIYRHAHGWMWLIPLP------EDIMSVG 258
MfPrnC  LYTHMVGVKPFDDIFKPKGQRWR---WHEGTLHHMFHGGWLWVIPFNNHSRATNGLVSVG 268
PfPrnC  LYTHMLGVKPFDDIFKVKGQRWR---WHEGTLHHMFEGGWLWVIPFNNHPRSTNNLVSVG 269
pltM    FFTHMLNVKSYEDALAPLSRTRSPIELFKSTLHHIFEEGWLWVIPFNNHPQGTNQLCSIG 257
PltD    FSTHMLGVVPFDECVQGDFPG----QWHGGTLHHVFDGGWVGVIPFNNHQHSRNPLVSVL 264
CmlS    VWSYFKLKDPFEGDLKG-----------TTYSITFEDGWVWMIPIKDDLYSVGLVVDEV 235
PltA    SWFDNFDRKSATQSPDK--------VDYIFIHFLPMTNTWVWQIPIT------ETITSVG 251
                                                *.  **:         : .
```

Fig. 30

Fig. 30 continued

```
AmHal       AVVRREDAEKIQGDREKALNT-----LIAECPLISEYLADATRVTTGRYGELRVRKDYSY 289
Xhal        AVCYPEYMKTRKGDSEAFLMR-----TLALNPELNARMLDAERVAP-----VHATGNYAY 308
MfPrnC      LQLDPRIHPKTEIPAQQEFDE-----FLARFPTIAAQFKDARPVRDWVK-----SDRLQY 318
PfPrnC      LQLDPRVYPKTDISAQQEFDE-----FLARFPSIGAQFRDAVPVRDWVK-----TDRLQF 319
pltM        FQFNNAKYRPTEAP-EIEFRK-----LLKKYPAIGEHFKDAVNAREWIY-----APRINY 306
PltD        VSLREDLCPSMDGD--QVLAG-----LIELYPGLGRHLSGARRVREWVLR---QPPRQVY 314
CmlS        RIVQDWSYDTEVFSADRFFLCGDAACFTDPLFSQGVHLASQSAVSAAAAIDRITRHGDEK 295
PltA        VVTQKQNYTNSDLTYEEFFWE-----AVKTRENLHDALKASEQVRPFKK-----EADYSY 301
                                :                   :     .

AmHal       QQETYWRPG--MILVGDAACFVDPVFSSGVHLATYSALLAARSINSVLAGDLDEKTALNE 347
Xhal        ECTRMAGPR--WLMLGDAYTFVDPMFSSGVFLAMHGAERGAAMVDAALRAPQSEAKLQRA 366
MfPrnC      SSKSTVGDR--YCLMLHAAGFIDPLFSRGLENTSVTIHALAARLIKALRDDDFSPERFEY 376
PfPrnC      SSNACVGDR--YCLMLHANGFIDPLFSRGLENTAVTIHALAARLIKALRDDDFSPERFEY 377
pltM        RSVQNVGDR--FCLLPQATGFIDPLFSRGLITTFESILRLAPKVLDAARSNRWQREQFIE 364
PltD        R--TALERR--CLMFDEGAASNDLLFSRKLSNAAELVLALAHRLIKAAHSGDYRSPALND 370
CmlS        DAVHAWYNRTYREAYEQYHQFLASFYTFASFTEPDSEFWRKRRITESDDDRLTRKKWFES 355
PltA        GMKEVCGDS--FVLIGDAARFVDPIFSSGVSVALNSARIASGDIIEAVKNNDFSKSSFTH 359
                          .       .::                 :

AmHal       FELRYRR-EYGVFYEFLVSFYQMNVNEESYFWQAKKVTQNQSTDVESFVELIGGVSSGET 406
Xhal        LQRELTR-GVDEF---------------------------------------------- 378
MfPrnC      IDRLQQK-LLEHNDDFVSCCYTAFSDFELWDAFHRLWAVGTMLGQFRLVQAHARFREDRN 435
PfPrnC      IERLQQK-LLDHNDDFVSCCYTAFSDFRLWDAFHRLWAVGTILGQFRLVQAHARFRASRN 436
pltM        VERHCLN-AVATNDQLVSCSYEAFSDFHLWNVWHRVWLSGSNLGSAFLQKLLHDLEHSGD 423
PltD        FVLTQDS-IISLSDRIALAAYVSFRDPELWNAFARVWLLQSIAATITARKINDAFAKDLD 429
CmlS        LAGNGPEDPSGTVASFRDRASTMIAIGRHQRPELSDDFSEAELNPARVRWISDLTKRLNS 415
PltA        YEGMIRN-GIKNWYEFITLYYRLN-------ILFTAFVQDPRYRLDILQLLQGDVYSGKR 411

AmHal       ALTAADRIAARSAEFAAAVDEMAGGDGDNMVPMFKSTVVQQAMQEAGQVQMKALLGEDAE 466
Xhal        ------------------------------------------------------------
MfPrnC      EAHLDHLDDNPPHLGYLCADMDAYCDLFDAAKAEVESVSEKRASPKEAAARIHALIEAQE 495
PfPrnC      EGDLDHLDNDPPYLGYLCADMEEYYQLFNDAKAEVEAVSAGRKPADEAAARIHALIDERD 496
pltM        ARQFDAALEAVRFPGCLSLDSPAYESLFRQSCQVMQQAREQARPVAETANALHELIKEHE 483
PltD        PRVFDEID-QLAEDGFWMPLYRGYKDILNTTLGLCDDVKSAKVSAAHAASSIFAELANAS 488
CmlS        ITRFKWTGGKAVLKQHYRVEPIGFRLEQREVLANGEGLDMAQYPMDDEARQIFQDLAEEE 475
PltA        LEVLDKMREIIAAVESDPEHLWHKYLGDMQVPTAKPAF---------------------- 449
```

Fig. 30

Fig. 30 continued

```
AmHal          PELPLFPGG---------------------------------------------------- 475
Xhal           ------------------------------------------------------------
MfPrnC         ---FARP--LFSFGYCITGANRNLNNSKYSLVPALRLLHWTQKGAPPEVKKYFDYNPMFS 550
PfPrnC         ---FAKP--MFGFGYCITGDKPQLNNSKYSLLPAMRLMYWTQTRAPAEVKKYFDYNPMFA 551
pltM           ---AELL--PLGY------------------------------------------------ 491
PltD           ---FVPPIFDFANPHARVYQLTTLRKLKALWWGLMQVPSEVG-----------------R 528
CmlS           FGYKTLVKRLGAVGRQELSTQIVVRLMEAGLLTGYDAQGEK-----------------VF 518
PltA           ------------------------------------------------------------

AmHal          LVTSPERMKWLPHHPA 491
Xhal           ----------------
MfPrnC         LLKSYVGNRLALALK- 565
PfPrnC         LLKAYITTRIGLALKK 567
pltM           ---SRISNRFILKV-- 502
PltD           LIFYRSFRKPSLRKES 544
CmlS           VQGRLHFGGVGVEYEV 534
PltA           ----------------
```

```
MfPrnC  (20)  GSGMSGSQMGAILGRQGF
PltM    (10)  GSGIAGALTGAVLAKSGL
PltA    (12)  GGGPAGSTMASYLAKAGV
CmlS    (11)  GGGPAGSVAGLTLHKLGH
AoHal   (12)  GGGPAGSTVATLVAMQGH
                *.* :*:  .  :   *

Consensus:    GxGx2(G/A)x3(G/A)x6G
```

Figure 31. Comparison of the putative NAD(P)H binding site of CmlS with similar sequences in other halogenases.
MfPrnC, halogenase PrnC from *Myxococcus fulvus*; PltM and PltA, halogenase from *P. fluorescens* Pf-5; CmlS, halogenase from *S. venezuelae* ISP5230; AoHal, halogenase from *A. orientalis*. The number in parantheses following the name of halogenase indicates the position of the first amino acid in the sequence.

```
Chl      (226)  VGERWCLMSHAAGFIDPLFLRGLSNTCEIINALS
MfPrnC   (324)  VGDRYCLMLHAAGFIDPLFSRGLENTSVTIHALA
PltM     (312)  VGDRFCLLPQATGFIDPLFSRGLITTFESILRLA
AoHal    (295)  WRPGMILVGDAACFVDPVFSSGVHLATYSALLAA
PltA     (307)  CGDSFVLIGDAARFVDPIFSSGVSVALNSARIAS
CmlS     (250)  SADRFFLCGDAACFTDPLFSQGVHLASQSAVSAA
AcPobA   (279)  RFGKLFLAGDAAHIVPPTGAKGLNLAASDIAYLS
PfPobA   (277)  QHGRLFLAGDAAHIVPPTGAKGLNLAASDVSTLY
                      *  .*: :  *     *

Consensus:          GxxxxxGDAxHxxxPxxxxGxxxxxxD
```

Figure 32. Comparison of putative FAD-binding sites in CmlS and other halogenases with the flavin-binding sites identified in PobA (hydroxylase) proteins (Dimarco *et al.*, 1993). Consensus residues are shown by bold letters.
Chl, halogenase from *Streptomyces aureofaciens*; MfPrnC, halogenase PrnC from *Myxococcus fulvus*; PltM and PltA halogenase from *P. fluorescens* Pf-5; AoHal, halogenase homologue from *Amycolatopsis orientalis*; PltA, halogenase from *P. fluorescens*; CmlS, halogenase from *S. venezuelae* ISP5230; AcPobA, hydroxylase from *Acinetobacter calcoaceticus*; PfPobA, hydroxylase from *P. fluorescens*. The number in parantheses following the name of the protein indicates the position of the first amino acid in the sequence.

ii. Phylogenetic comparison of halogenases and monooxygenases

BLASTP revealed close similarity between the aa sequences of halogenases and certain monooxygenases. To compare their relatedness, sequences were retrieved from GenBank and used in ClustalX to obtain a phylogenetic dendrogram (Figure 33).

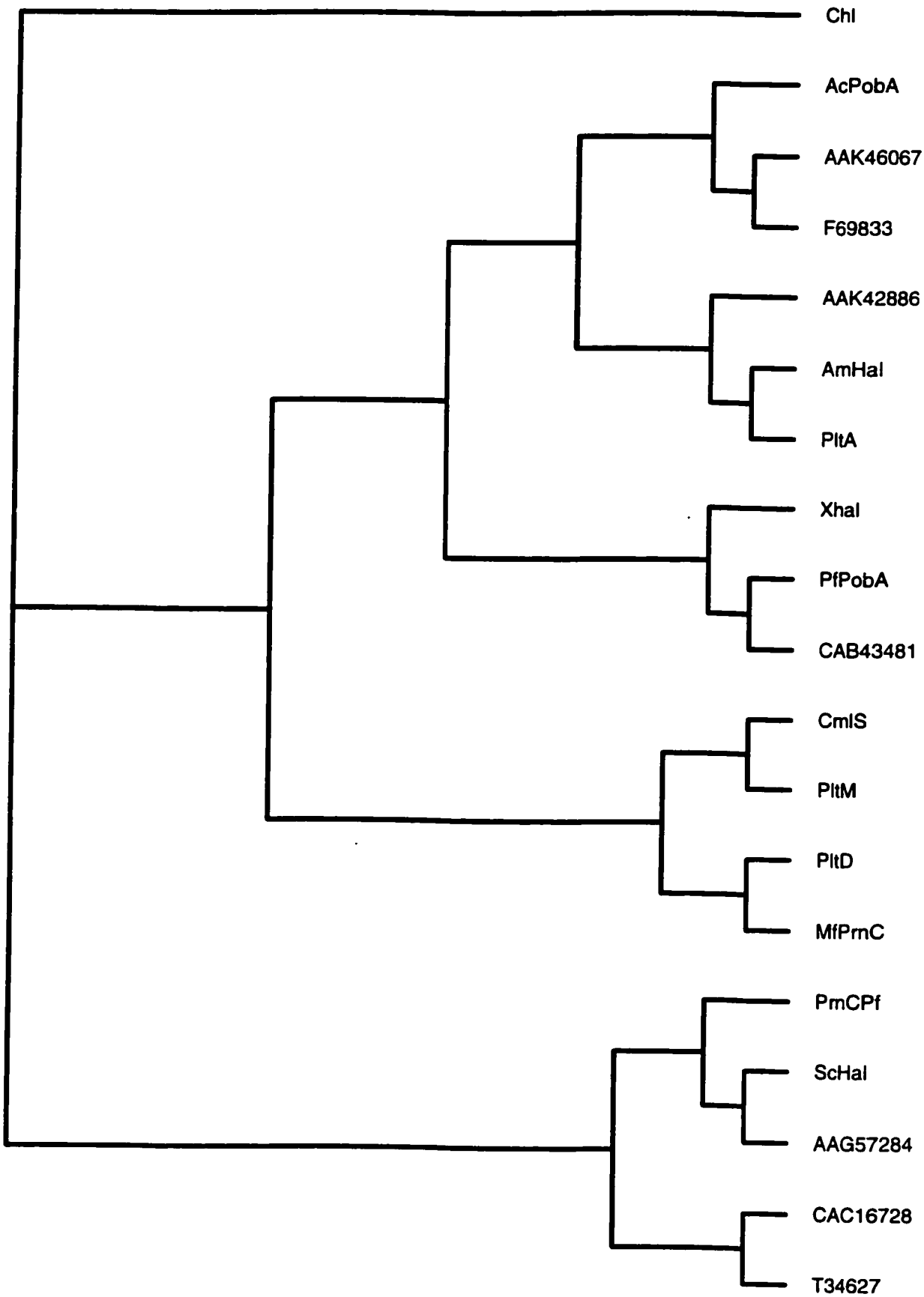
iii. Functional analysis of CmlS

To disrupt ORF12, the 3.8-kb *Sac*I-*Xba*I fragment containing the gene was retrieved from pJV506 and inserted into the multiple cloning site of pJV326 to give pJV507 (Figure 34). ORF12 was disrupted by inserting into its unique *Stu*I site a 1.5-kb *Eco*RV fragment containing the Am^R gene obtained from pJV225 (Chang. 1999), giving pJV508 (Figure 35). The disruption plasmid pJV508 was conjugally transferred from *E. coli* to *S. venezuelae* ISP5230. Transconjugants were screened to obtain mutant strain VS1102 with an Am^RTs^S phenotype.

Disruption of ORF12 by insertion of the Am^R gene and selection of a double-crossover mutant (i.e., Am^RTs^S transconjugant) was confirmed by Southern hybridization analyses of Southern blots of genomic DNA from the disrupted and wild-type strains. The DNA extracted from VS1102 was digested with *Sma*I and the restriction fragments were probed with a labeled 1.5-kb *Sal*I-*Sal*I fragment from pJV504 containing part of ORF12 (Figure 36). A signal at ~3.2-kb, corresponded to the insertionally inactivated ORF in VS1102. Genomic DNA from *S. venezuelae* ISP5230 digested with *Sma*I gave a ~1.7-kb signal as expected for the wild-type genome. Also consistent with replacement of the wild type gene with the inactivated copy, hybridization with the ^{32}P-labeled Am^R gene as a probe, gave a signal at ~3.2-kb from strain VS1102, but no signal from *S. venezuelae* genomic DNA.

Figure 33. A dendrogram depicting the relatedness of halogenases to certain monooxygenases.

AAK46067, FAD-binding monooxygenase from *Mycobacterium tuberculosis* CDC1551 (accession AAK46067); F69833, monooxygenase homologue yhjG (tetracycline 6-hydroxylase superfamily) from *Bacillus subtilis* (accession F69833); CAB43481, *p*-hydroxybenzoate hydroxylase from *Pseudomonas* sp. (accession CAB43481); AAK42886, the electron transfer oxidoreductase from *Sulfolobus solfataricus* (accession AAK42886); AAG57284, putative oxidoreductase from *Escherichia coli* O157:H7 EDL933 (accession AAG57284); CAC16728, putative monooxygenase from *Streptomyces coelicolor* (accession CAC16728); T34627, probable electron transfer oxidoreductase from *Streptomyces coelicolor* (accession T34627); Chl, halogenase from *S. aureofaciens*; Xhal, putative reductase/halogenase from *Xanthomonas oryzae* pv. oryzae (accession AAG38844); MfPrnC, halogenase PrnC from *Myxococcus fulvus*; PltM, PltD, and PltA, halogenase from *P. fluorescens*; PrnCPf, halogenase PrnC from *P. fluorescens*; CmlS, halogenase from *S. venezuelae*; AcPobA , *p*-hydroxybenzoate hydroxylases from *Acinetobacter calcoaceticus*; PfPobA, *p*-hydroxybenzoate hydroxylases from *P. fluorescens*; AmHal, halogenase from *Amycolatopsis mediterranei*; ScHal, halogenase from *S. coelicolor* (accession CAB99155).


Chl
AcPobA
AAK46067
F69833
AAK42886
AmHal
PltA
Xhal
PfPobA
CAB43481
CmlS
PltM
PltD
MfPrnC
PrnCPf
ScHal
AAG57284
CAC16728
T34627

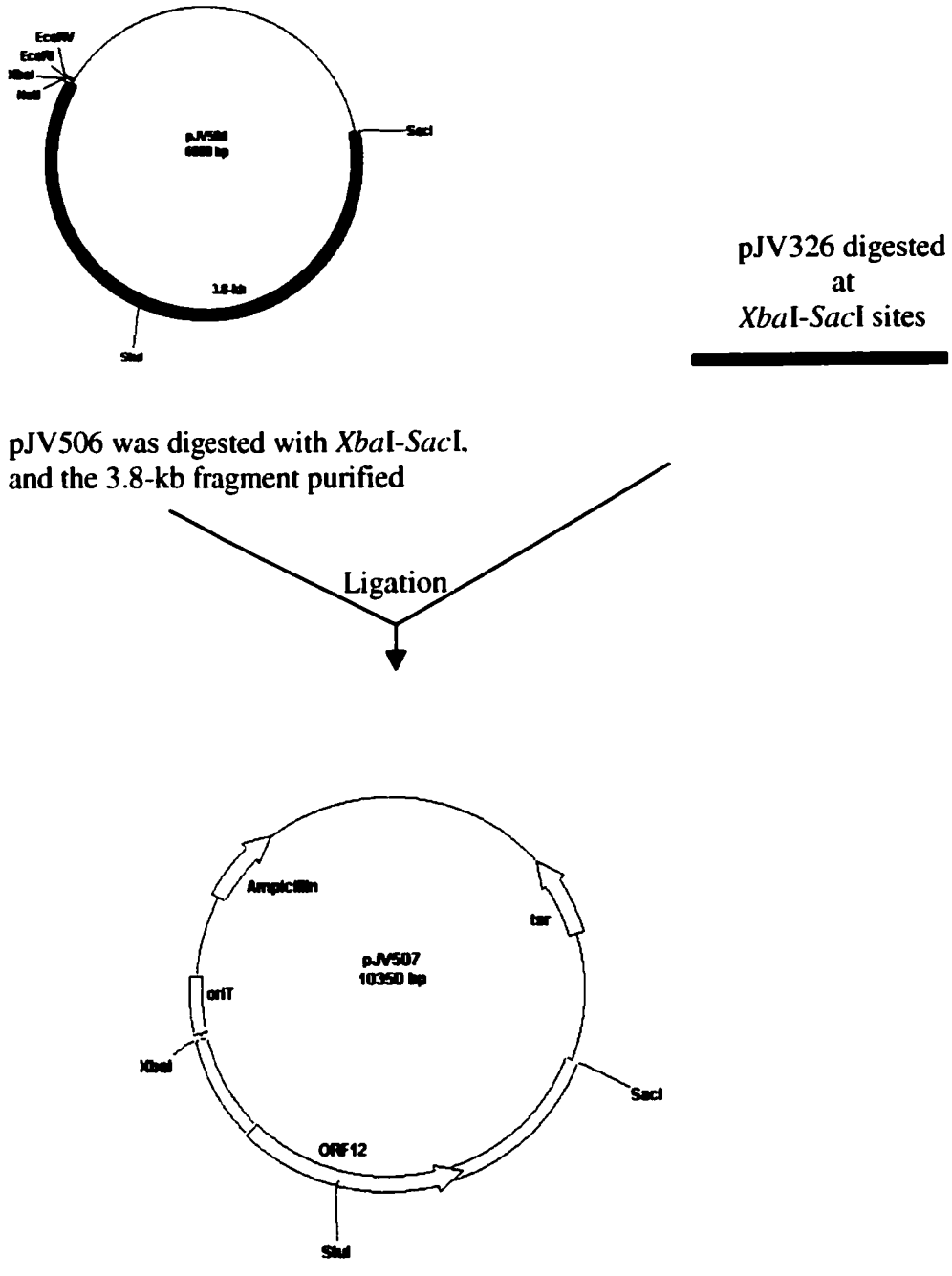


Figure 34. Construction of pJV507 by ligating the 3.8-kb *Sac*I-*Xba*I fragment from pJV506 into the multiple cloning site of pJV326.

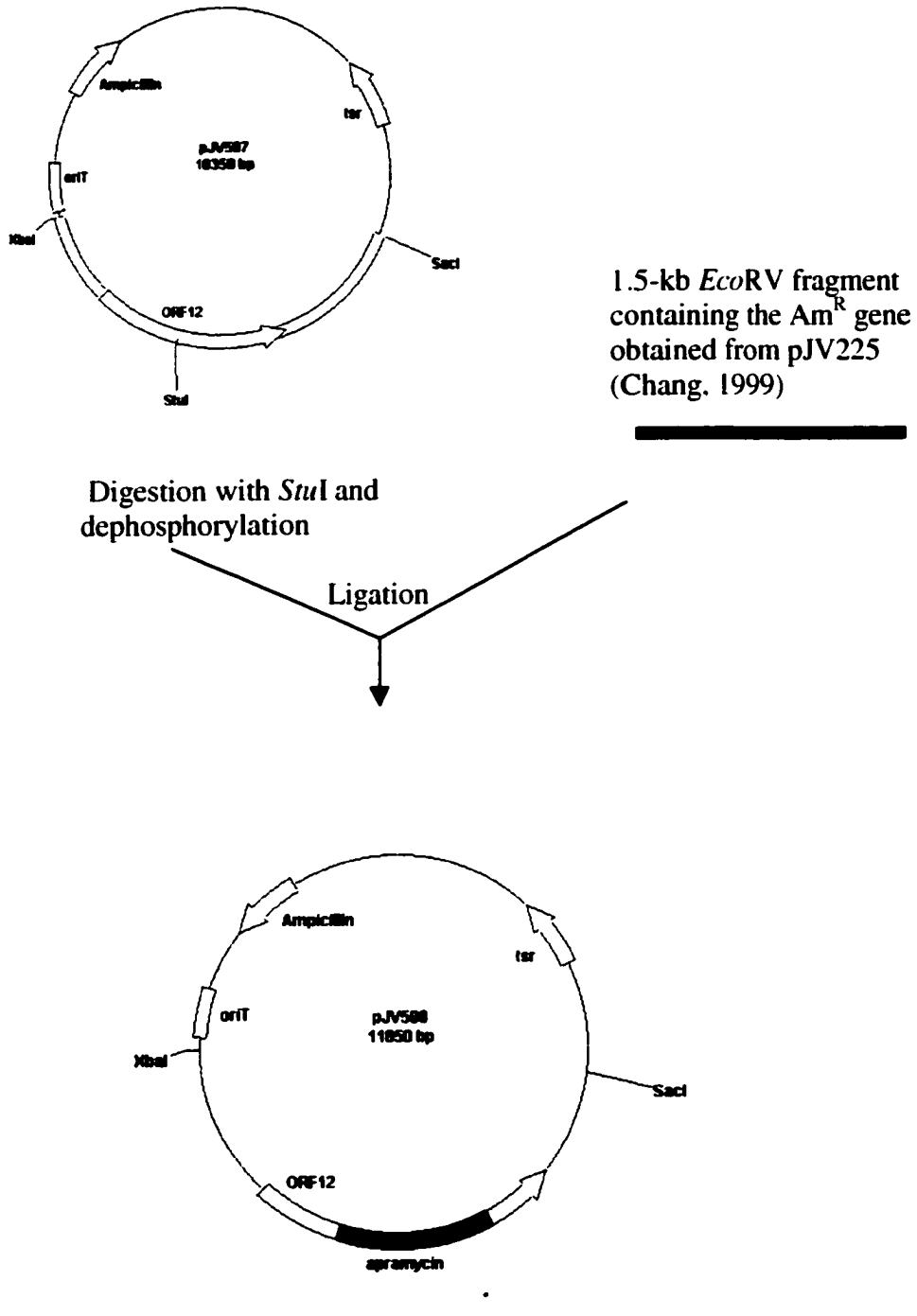


Figure 35. Construction of disruption plasmid pJV508 by ligating the *Eco*RV-apramycin cassette into *Stu*I site of ORF12.

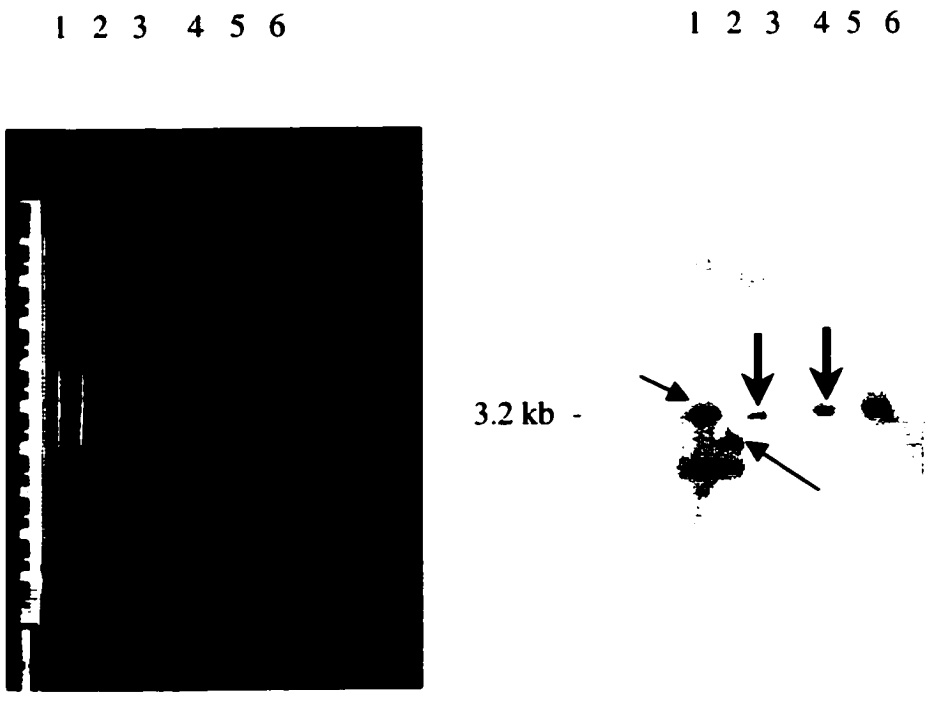


Figure 36. Evidence from Southern hybridization for a double cross-over in mutant strain VS1102.
A) Gel electrophoresis of genomic DNA samples digested with *Sma*I.
B) Hybridization of Southern blots of the genomic digests with an ORF12 probe showing the 1.5 kb increase in size of the *Sma*I fragment from VS1102 compared with that from the wild-type (difference between lanes 1 and 2 shown with thin arrows). No hybridization occurred in lane 5 with an Am^R gene probe, whereas lane 6 gave a signal at about 3.2-kb. Thick arrows mark 3.28-kb.

Lanes: 1 and 6, mutant VS1102; 2 and 5, *S. venezuelae* ISP5230; 3 and 4, DNA size marker.

iv. Analysis of disrupted mutants VS1101 and VS1102

Disruption of either the CmlK or the CmlS gene blocked Cm production under conditions where wild-type cultures produced about 45 μg/ml of the antibiotic. Each of the blocked mutants VS1101 and VS1102 accumulated a series of compounds showing HPLC peaks with retention times (Rt) of 5.49, 6.08, 6.68 and 7.08 min. The compound with Rt 6.08 gave the dominant peak. Since the Rt values were identical to those of corynecins I, II, III, and IV respectively, reference samples of each of these compounds were co-injected individually with samples of the mutant extract to confirm the identity of the peaks by HPLC analysis. In each case the reference compound was co-eluted with one of the peaks, causing its area to increase. The major compound accumulated in cultures of VS1101 and VS1102 was identified as corynecin II, a Cm-analogue containing a propionyl in place of a dichloroacetyl substituent (Vining & Stuttard, 1995). The minor peaks were identified as the acetyl (corynecin I), isobutyryl (corynecin III) and 2-methylbutyryl (corynecin IV) congeners. Table 5 shows the retention times (Rt) of corynecins and Cm by HPLC. Figure 37 shows HPLC chromatograms of the extract from mutant VS1102; similar results were obtained when an extract from strain VS1101 was used. Corynecin II constituted the major component of the extract; corynecins I, III, and IV were the minor components.

The compound identified by HPLC as corynecin II was purified by reverse-phase chromatography from an extract of strain VS1102 and analyzed by ^{1}H-NMR spectroscopy; comparison of the NMR data with a standard sample confirmed its identity as corynecin II (Appendix I). These results indicated that CmlK and CmlS function in the chlorination reaction of Cm biosynthesis.

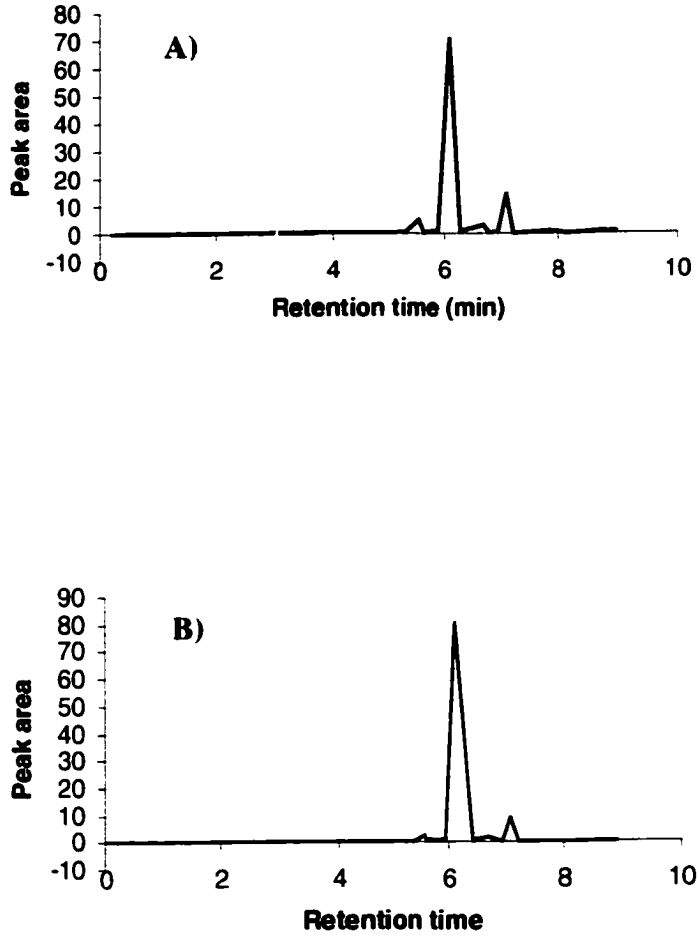


Figure 37. Chromatograms from HPLC analyses of **A)** the extract from mutant strain VS1102. The Rt of the major peak (6.08 min) corresponded to that of corynecin II.
B) the extract from VS1102 mixed and co-injected with a reference sample of corynecin II; the area of only the peak at Rt 6.08 min is increased, confirming its identity.

Table 5. Retention times of Cm and corynecins using RP-HPLC [a].

Compound	Rt (min)
Corynecin I	5.46
Corynecin II	6.05
Cm	6.7
Corynecin III	6.68
Corynecin IV	7.07
Products accumulated	5.49, 6.08[b]
by mutants VS1101 and VS1102	6.68, 7.08

a- Details of the HPLC analysis are given in the Methods & Materials section.
b- Co-injection of the product accumulated by mutant strain VS1101 or VS1102 with authentic samples of each corynecin yielded enhanced peaks at the retention time predicted, indicating the identity of each peak from the mutant strains.

c. ORF13 (encoding CmlT, an aldo/keto reductase)

i. Sequence analysis

The third ORF (ORF13) identified in the 3.8-kb *Not*I-*Sac*I fragment (Figure 20) started with an ATG codon located at nt 2670-2672, which is 230 bp downstream of the stop codon of ORF12. The ORF is preceded by the likely Shine-Dalgarno sequence GGAGG (ΔG = -7.8 kcal mol^{-1}). The first in-frame TGA was at nt 3546-3548 giving an ORF containing 879 bp encoding a protein of 292 aa (CmlT: Figure 38). The codon-third-letter mol% G+C of ORF13 was 90%, as expected for a *Streptomyces* gene. A BLASTP search revealed close identity between the deduced amino acid sequence of ORF13 and aldo/keto reductases from *Deinococcus radiodurans* (58% over 228 aa; White *et al.*, 1999), *Mesorhizobium loti* (60% identity over 211 aa; Kaneko *et al.*, 2001), and *Streptomyces clavuligerus* (56% over 210 aa; Mosher *et al.*, 1999). Figure 39 compares the aa sequence of CmlT with sequences retrieved from GenBank for three of the aldo/keto reductases with high similarity to CmlT.

ii. Functional analysis of CmlT

To disrupt ORF13, plasmid pJV507 containing the 3.8-kb chromosomal fragment that included ORF13 was digested with *Stu*I and *Not*I to remove a 1.75-kb fragment including ORF11 and part of ORF12. The residual linear plasmid fragment containing ORF13 was blunt-ended with S1 nuclease and separated from other products by gel electrophoresis. It was purified from the agarose gel and religated to give pJV516.

```
MRHTRLRDLEVSRIGLGTMGMSFGYTGSGSDDAGSVRALHRALDLGVTFL      50

DTAEIYGPYTNERLVGRALGARRDQVVLATKFGLVSHAGGGPGQLDSSPA      100

NVRTAVEGSLRRLGTDHIDLCYQHRVDTVGALAELVAEGKIRHIGLSEAG      150

PATIRRAHAVHPVTAVQSEYSLWTRDPEEAVLPVLRELGIGFVAYSPLGH      200

GFLTGTVRAAGRFTPRTSGRTTPASPRRTCPGHLRIADEIAAVAAEADAT      250

PAQVALAWLLTRARPAESEHPWPRGELRQGCGGHIPHPGGVP              292
```

Figure 38. The deduced amino acid sequence (CmlT) of ORF13.

```
CmlT        --------------------MRHTRLRDLEVSRIGLGTMGMSFGYTGSGSDDAGSVRAL 39
MlAldo      ----------------MTLATRKLGSQGLTVSAIGLGCMGMSQSYG--PADEAESIATL 41
DrAldo      ---------------MTTLPTRQLR--DLTVSALGLGCMGMSEFYG--DADEAESLRTL 40
ScAldo      MSRSPPESPAGSVSAAVPRPPVRALR--DLPVSAQGLGCLPTTDFYG--RPDRARATATI 56
                                 *     .* **  *** :  :  *    .* * :  ::

CmlT        HRALDLGVTFLDTAEIYGPYTNERLVGRALGARRDQVVLATKFGLVSH-AGGGPGQLDSS 98
MlAldo      HRAIELGCTFLDTAEVYGPHTNEALLGRALKGKRDQVTIATKFGFRIE-NGKQLSGVDSR 100
DrAldo      DRALDLGVTFFDTADIYGPYHNEELLGRWLRGKRDRVVLATKFGIVRDPADPSRRGFSGR 100
ScAldo      RAAVDAGVTLLDTADVQGLGAGEELLGRAVAGRRDEVLIATKFGMVRS-SDGASQGLCGE 115
              *:: * *::***:: *   .* *:** : .:**.* :*****:      .     . .

CmlT        PANVRTAVEGSLRRLGTDHIDLCYQHRVD-------TVGALAELVAEGKIRHIGLSEAGP 151
MlAldo      PEHIREVVEASLGRLATDHIDLLYQHRVDPDVPMEDVAGAVGKLVAEGKVRFFGLSEAGS 160
DrAldo      PEYVRQSIEGSLNRLGTDHVDLYYLHRVDPDTPIEDTVGAMADLVQEGKVRALGLSECSA 160
ScAldo      PSYVRAACERSLRRLGTDRIDLYYQHWTDPAVPIEETVGAVAELVREGKVRRLGLSEPSA 175
            *  :*   * ** **.**::** * * .*       ..**:..** ***:* :**** ..

CmlT        ATIRRAHAVHPVTAVQSEYSLWTRDPEE-AVLPVLRELGIGFVAYSPLGHGFLTGTVRAA 210
MlAldc      ANIRRAHAVFPVSALQSEYSLWERNLEP-EIIPLLKELGIGLVPFSPLGRGFLTGDVKRA 219
DrAldo      DTLRRAHAVHPISAVQSEYSLWTRDPEENGVLAACRELGVGFVPYSPLGRGFLAGQLKSP 220
ScAldc      ATLRRADAVHPVTAVQSEWSLWSRGIED-EVVPVCRELGIGIVAYAPLGRGFLTGTIRTT 234
              .:***.**.*::*:***:*** *. *   ::.  :***:*:*.::***:***:* :: .

CmlT        GRFTPRTSGRTTPASPRRTCPG--TCASPTRSPPSPRGGR-------------------- 248
MlAldc      EDYPEGDYRRNDPRYQGENYDANVEAAGTVRAIAAARGVKPGQIALAWLLGKGIGFGMDI 279
DrAldo      DDLAPDDFRRHNPRFQDENFQKNLDLVREIERLAAEKGCTPSQLALAWVLAQG----ENL 276
ScAldc      DDLGDEDFRRGQPRFSAPALARNRSLLHRLRPVADGLGLTLAQLALAWLHHRG----EDV 290
                      *  *                  .  .   *

CmlT        --HPGAGR-----PRLAAHP-------------------GGRRYS--------- 267
MlAldc      VPIPGTKRRKYLEENVAAASLKLDAAEMAALDEALAPEKISGPRYGERGMAMVDR 334
DrAldo      VPIPGTKRVKYLEENVGALDVRLTPDELRQIDATFP--VAAGDRYP--DMRHVNR 327
ScAldo      VPIPGTANPAHLADNLAAASIRLDDRSLAEVTAAIS-HPVSGERYTPALLAMIGN 344
              **: .      .:.*                    .* **
```

Figure 39. Alignment of the CmlT sequence with other aldo/keto reductases retrieved from GenBank.

Aldo/keto reductases are: CmlT, from *S. venezuelae*; MlAldo, from *Mesorhizobium loti* (Kaneko *et al.*, 2001); DrAldo, from *Deinococcus radiodurans* (White *et al.*, 1999); ScAldo, from *Streptomyces clavuligerus* (Mosher *et al.*, 1999). Abbreviations are as shown in Figure 11.

ORF13 was disrupted by inserting the apramycin resistance gene from pUC120A (Paradkar & Jensen, 1995) into its unique *Nco*I site, giving pJV518 (Figure 40). The disrupted plasmid was conjugally transferred from its *E. coli* host cell to *S. venezuelae* ISP5230 to give transconjugant VS1103 containing the insertionally inactivated copy of ORF13 in its chromosome.

To show that the double-crossover mutant had been obtained, genomic DNA from VS1103 was extracted and digested with *Sma*I. The Southern blot prepared from *Sma*I restriction fragments was probed with a labeled 0.75-kb *Sma*I-*Sal*I fragment containing part of ORF13. Detection of a ~2.9-kb signal in VS1103 (Figure 41) confirmed the double-crossover. Probing *Sma*I-digested genomic DNA from *S. venezuelae* ISP5230 gave a signal at ~1.4 kb, consistent with the restriction map (Figure 23). When DNA samples from disrupted mutant and wild-type strains were probed with the ^{32}P-labeled Am^R gene, only DNA from the mutant strain VS1103 hybridized, and gave a signal at ~2.9-kb. HPLC analyses of the extract from mutant VS1103 showed that Cm was produced at the same level as in the wild-type. This result indicated that the product of ORF13 is not essential for Cm biosynthesis. Although it probably has a role in primary metabolism, no phenotypic change was detected after its disruption.

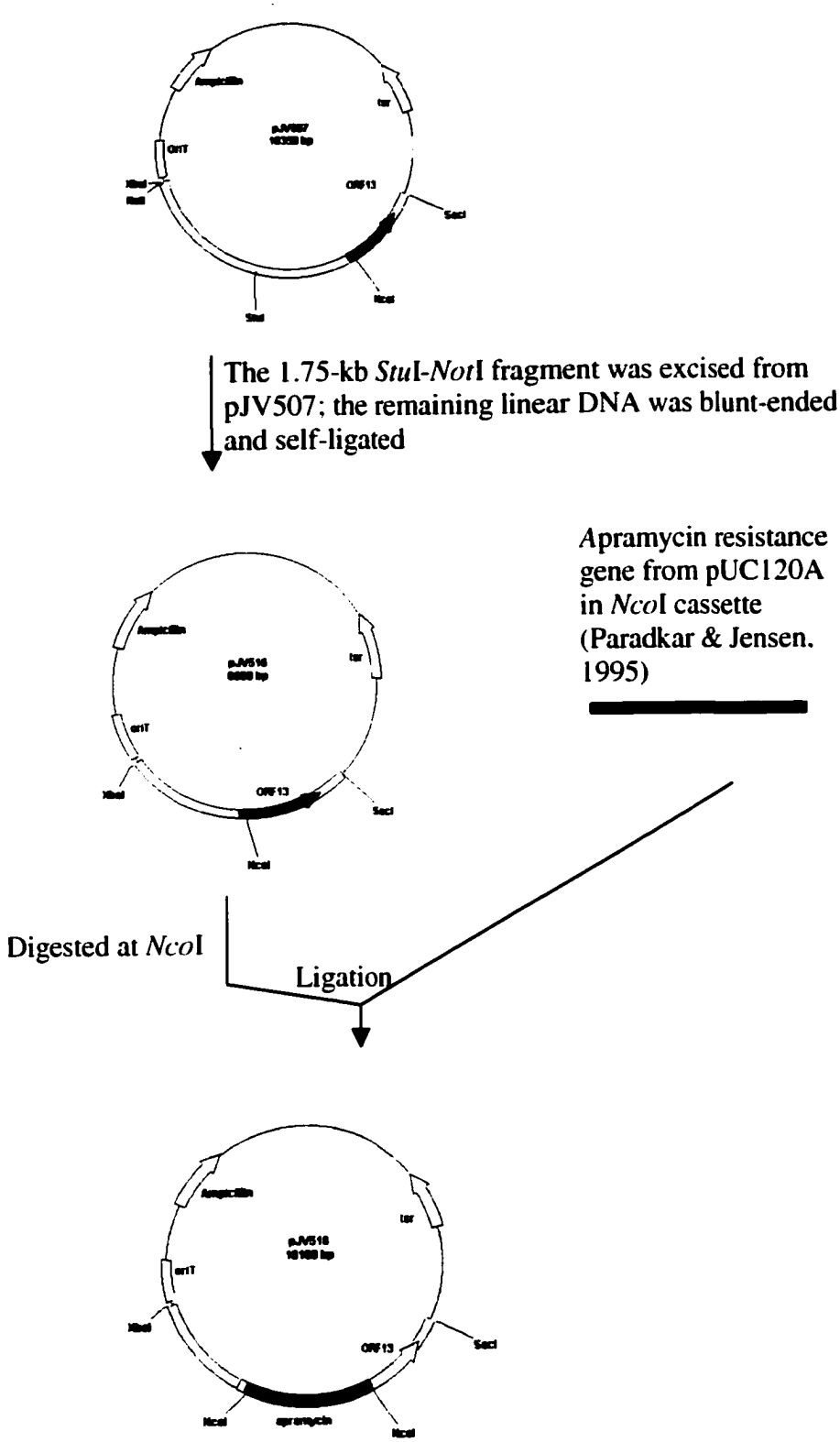


Figure 40. Construction of disruption vector pJV518 containing ORF13 disrupted by an *Nco*I-apramycin resistance gene cassette.

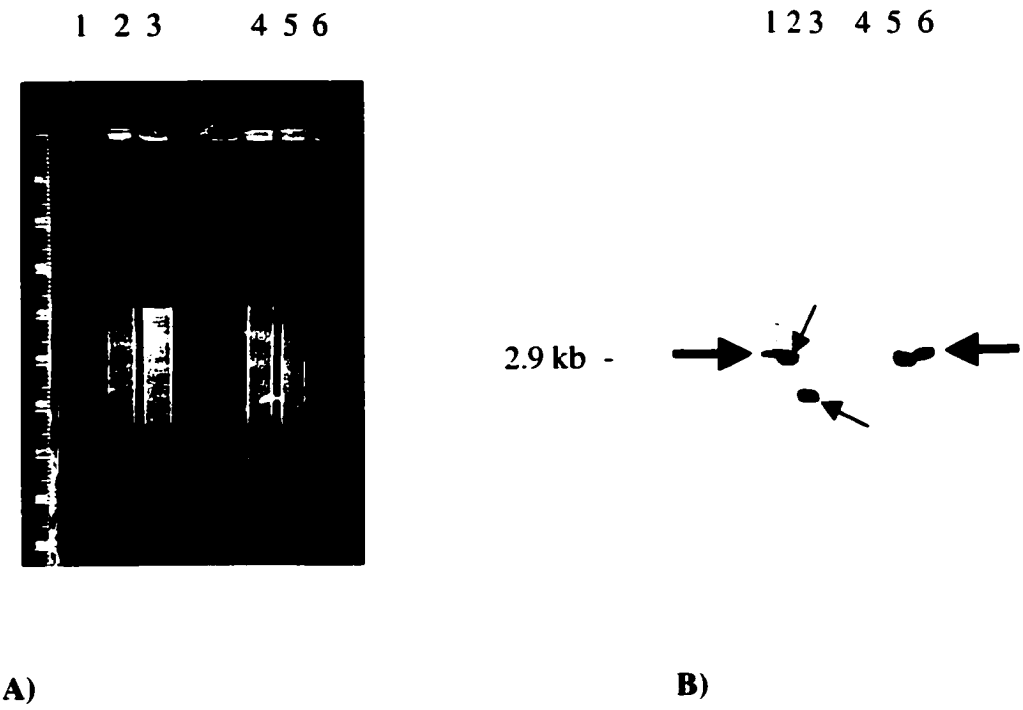


Figure 41. Evidence from Southern hybridization for a double cross-over in mutant strain VS1103.
A) Gel electrophoresis of genomic DNA samples digested with *Sma*I.
B) Hybridization with an ORF13 probe of the Southern blots showing the increase in size of the *Sma*I fragment from VS1103 due to insertion of the 1.5-kb apramycin resistance gene into the wild-type genome (difference between lanes 2 and 3 shown with thin arrows). Thick arrows mark 3.28-kb. No hybridization occurred when wild-type DNA was probed with the Am^R gene (lane 4), while VS1103 gave a signal at about 2.9-kb (lane 5).

Lanes: 2 and 5, mutant VS1103; 3 and 4, *S. venezuelae* ISP5230; 1 and 6, DNA size marker.

III. Extension of sequence beyond ORF13

To extend the cloned *S. venezuelae* sequence beyond the 3'-end of the 3.8-kb *Not*I-*Sac*I chromosomal fragment (see Fig 23), an 850-bp *Sac*I-*Sal*I fragment was excised from its 3'-end and used in a chromosome-walking procedure to probe the restriction fragments of recombinant lambda phage MP30 DNA. The probe hybridized with a 7.5-kb *Pst*I fragment of DNA from the phage (Figure 42). Since the 3.8-kb *Not*I-*Sac*I fragment has only one *Pst*I site, and the 850-bp *Sac*I-*Sal*I probe was obtained from its 3'-end, the 7.5-kb *Pst*I fragment was expected to contain a region of the chromosome downstream of ORF13. For unknown reasons, attempts to subclone the 7.5-kb *Pst*I fragment were unsuccessful. Therefore, this fragment was purified from an agarose gel and labeled for hybridization with a *Sac*I digest of phage MP30 DNA. The probe hybridized to a 1.9-kb *Sac*I fragment (Figure 43). This fragment was purified from an agarose gel and subcloned into the multicloning site of pUC18, giving pJV513. The chromosomal insert in pJV513 was sequenced and shown to contain one intact open reading frame (ORF14) and the incomplete ORF15, which was separated from ORF14 by 156 bp (Figures 44, 45). ORF14 was located 481 bp downstream of the ORF13 stop codon. The 600-bp ORF14 sequence encodes a 199-aa protein. Sequence comparisons with GenBank using BLASTP showed high similarity to a hypothetical protein (SC6D11.26; GenBank CAB76349) from *Streptomyces coelicolor* A3(2). No stop codon for ORF15 was detected in the cloned fragment, but the amino acid sequence deduced for the translated gene was similar to a hypothetical protein (SCF56.19; GenBank CAB62764) from *Streptomyces coelicolor* A3(2). Sequence comparisons with the protein database in GenBank did not allow probable functions to be assigned for ORF14 and ORF15.

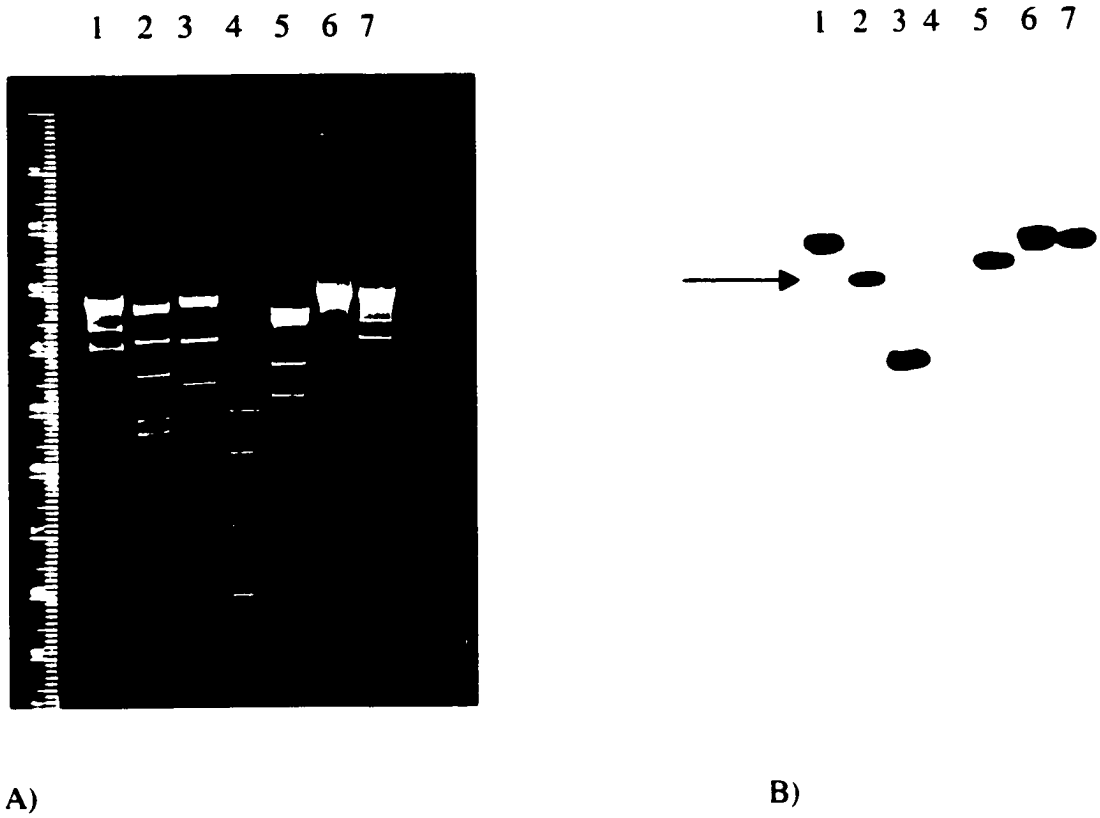


Figure 42. A) Agarose gel electrophoresis of MP30 DNA digested with various restriction enzymes, and **B)** the digests probed with an 850-bp *Sac*I-*Sal*I fragment from the 3'-end of the 3.8-kb *Not*I-*Sac*I fragment. The arrow indicates hybridization of the probe with the 7.5-kb *Pst*I fragment of MP30 DNA.

Lanes: 1, *Bgl*II; 2, *Pst*I; 3, *Nco*I; 4, DNA size marker; 5, *Mlu*I; 6, *Kpn*I; 7, *Bam*HI.

Figure 43. Hybridization of Southern blot from the *Sac*I digest of MP30 DNA with the 7.5-kb *Pst*I probe from MP30 (see Figure 42). An arrow identifies 1.9-kb *Sac*I-*Sac*I fragment containing the region downstream of ORF13.

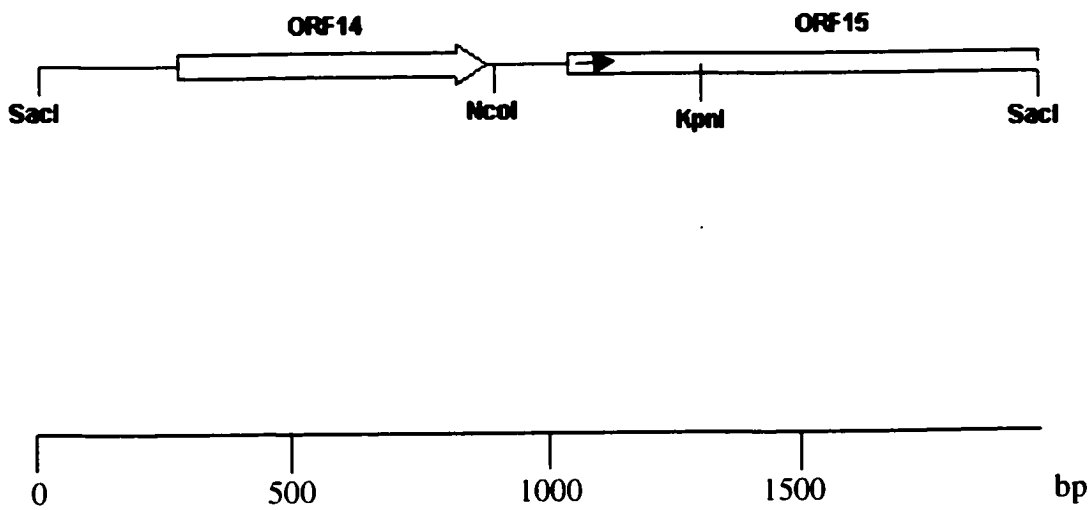


Figure 44. Restriction map of the *Sac*I-*Sac*I fragment of the *S. venezuelae* chromosome extending downstream of ORF13. The arrow indicates direction of transcription for partial ORF15

Figure 45. Nucleotide sequence of the 1.9-kb *Sac*I-*Sac*I fragment cloned in pJV513 that lies immediately beyond the 3.8-kb *Not*I-*Sac*I fragment cloned in pJV506 (see Fig 22). Arrows mark the initiation codons and indicate the direction of ORFs. The stop codon (TGA) is shown with italic-bold letters. Selected restriction enzyme sites are shown above their sequences.

CGGCGGGCACCAGGCACTGCTCAGGGCCCATCTCGGGCTCGGCATCACCG 50

AGGAACAGCGGCTGCGCTGGATGGAGCTGATGACGGCGGCCGTGGAGAAG 100

GAACTGCCGGACGACGAACTGCTGCGGCGGCGGGTGGTGGAGTACTTCGA 150

CTGGGGCACCCGGATCGCAAGGACGTCTCGGCCTCCGCTCCGGGCACGGA 200

CCTCGGCGAGCCCGGCCCCACGCCGCGCTGGGGCTGGGGCGGCCTCGCCT 250

ORF14

GACCGAGAGGGCGCCCCCTAGGGTGCGGGGGTGCCTTCGCGAACCCTCCT 300

CACCAGCGCCTCTCTCGCCGACGCGGCCCTCCTCACGCACACCGAGCAGG 350

GCCGCGAGTGGCCGCTCCTGGTGTGGGCACCGGGGCCGGGCGTCCGCATG 400

GTGTCGAGCGCGGTGCTCGGCGGCGGCATCGGCGAGCGGGGCTGGGTGCT 450

CAACGCCCAGGTCCCGCCCGGCTACGACCGGCTCGATCCCGTCGACCACC 500

TCCGCGAGCTGGCCGCCGGGCGGGGCTCGCCCGGGACGATCAACATCATC 550

GTCTCGTGCCCCGTCCCCCTGACGGACGCGGCGCTGGTCAACGCGGTGGC 600

CACGGCCACCGAGGCGAAGGTGCAGGCCCTGGTGGAACTGGGCGCGGACG 650

CCTCCGGCACCCCCACGGACGCGGTCTGCGTGGCGGCCCCGACGCCCGAA 700

GGCACCGAAGCCGAACCCTTCGCGGGCCCGCGCTCCCGGTGGGGCGCCCG 750

CCTGGCCCGGGCCGCCAGTCAGCAGAGGGCGGTGCCCGGCAGGACCCCGA 800

GACCTGGCAGACTGCGGGCGTTCACGGTCGTCTCCCCCCGCGCTGTGAGC 850

*Nco*I

GGACCGCTCCACCGGGGATGCGGCAAGTGACTCCAGGAGGCCCCCCATGG 900

CGCAGCAGCCCGGCCCCGCGGCCCGCACGAACGGGACGTGCGAGCGATTC 950

CGGTACGCGAGCGGGACGTGCGAGCGACTCCGGTACGCGAGCGGGACGTG 1000

ORF15

CGAGCGACTCCGGTACGCGAGCGGGGCGTGCGCGCGATGACGCTCTGCCC 1050

Fig. 45

Fig. 45 continued

GACCCGGCGCGCCGCACCCCGCACCCTGCTCGTCCGCAGCGCGCCCGCCG 1100

CACCCACGGCCGCCGTGCTCCTGCTGCACGGGGGACGCGAAGACGGCCCC 1150

GAGCCGCCGCCCGCACTCAACCTCCCGGCGCTGCGCATGCGTCCCTTCGC 1200

CACCGCCGTGACCCGCGCCGTCCGGGGCCGCGACGTGCTGATCGCCGAGG 1250

TGCGCTACCGCCACCGCGGCTGGAACGGACCGAACGCCGACGCCGCCCGG 1300

GTACCGACCTCGATCAACGGCCTCTGCCTGAGCCGGCGGGAGCAGATCGC 1350

GGGAGCCACGTACACCCTGCTCCTCACCCTCGGCGACGACGACCCCGGCT 1400

CCTGGGAGATCAGGACGAAGAACACCGTCCTCGACCTGCGCGACGCCACC 1450

GGAGTCCTCGACATGGACGGCGTCGTCGACCTCAACACCAACGGCGCCGA 1500

CGTGAAGACCGTCAAGGACGCGACCGGGGCGTACGTCGTCAACCCGCTCG 1550

ACAGCCCCCAGCACCGCTTCGGCATCCAGGCCCGGTACGCCAAGTTCGAC 1600

CGGATCGTCGGCACCGCCCAGGACTTCGAGGTCTTCCGCACCGCGCTCCG 1650

CGCCGCCGACGCGGACGCGGTCCGCGGCGTCGTCGCCCTCGCCGCGTGGT 1700

GCCCGCCGGGCGAGCCGGTCGCGCAGCTCCGCGACCGCGCCCTCCTCGTC 1750

CTCCACGGCTCCCGCGACCGCGTCACCGACCGCGGCCAGTCCGCAGCCTA 1800

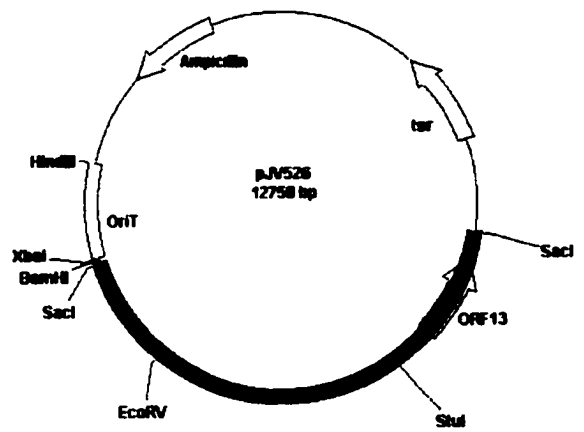
CGTCTCCCGCGCCCGCCCCGACGGCGCCCGCGCCGGCCTCCTCCACATCG 1850

AGAGCGGCGACCACGTCATGCTCCGCCACCACCCCGGCTGGCACCGCACC 1900

GCCACC 1906

IV. Introduction of ORFs 11, 12 and 13 into mutant *cml-2*

Doull *et al.* (1985) and Vats *et al.* (1987) demonstrated that cultures of *S. venezuelae* strains obtained by exposing the wild-type to various mutagens accumulated corynecins instead of Cm, and concluded that a mutation (*cml-2*) had occurred in the gene(s) responsible for biosynthesis of the dichloroacetyl moiety of the antibiotic. In an effort to complement this *cml-2* mutation, genes cloned in pJV507, pJV526, and pJV528 were introduced by transformation into a *cml-2* mutant. The conjugal plasmid pJV507 contained ORF12 (see Fig 34). pJV526 was obtained by subcloning the 6.2-kb *Sac*I-*Sac*I fragment of pJV502 in the conjugal vector pJV326 (see Figure 26), and contained ORFs 11, 12 and 13 intact. pJV528 was prepared by digesting pJV526 with *Stu*I-*Eco*RV, isolating the resulting ~10-kb linear fragment, and self-ligating it to give pJV528 (Figure 46). This plasmid carried only ORF13 as an intact open reading frame. Each plasmid was transferred conjugally into the *cml-2* mutant, and transconjugants VS1104 (containing pJV507), VS1105 (containing pJV526), and VS1106 (containing pJV528), all with a Ts^R phenotype, were isolated. The transconjugants were tested for Cm production to assess whether the *cml-2* mutation(s) had been complemented by allele exchange. Results showed that none of the plasmids was capable of complementing the mutant, indicating that *cml-2* phenotype might be due to mutations in regulatory genes or more likely is a multi-site mutation.



A 2.7-kb *Stu*I-*Eco*RV was deleted and the residual ~10-kb fragment was purified and self-ligated

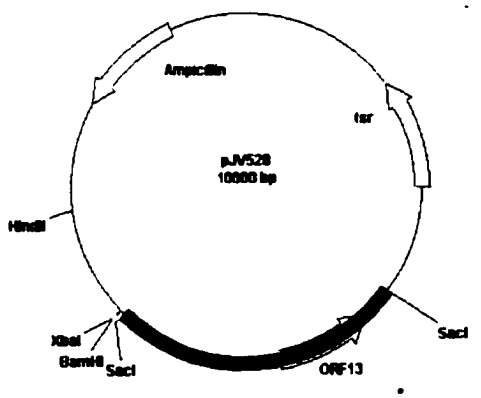


Figure 46. Construction of pJV528 by deleting a 2.7-kb *Stu*I-*Eco*RV from pJV526 and self-ligating the remaining ~10-kb linear fragment.

Discussion

Section 1. Isotopic incorporation experiments

Previous studies (Gottlieb *et al.*, 1956) showing that dichloroacetic acid is not a precursor of the dichloroacetyl substituent in Cm, have suggested that the coenzyme A ester serves as the substrate for an acyl-transfer reaction. Incorporation of ^{13}C from [6-^{13}C-glucose] into the dichloromethine and carbonyl carbons of the dichloroacetyl group in Cm is consistent with participation of acetyl CoA (Munro *et al.*, 1975) but results with [1,2-$^{13}C_2$]acetate (Simonsen *et al.*, 1978) indicated that the pathway generating dichloroacetyl CoA is probably complex.

In this thesis, where ^{13}C-labeled analogues of CoA esters were used to supplement Cm-producing cultures of *S. venezuelae*, none of the substrates appeared to be direct precursors of the dichloroacetyl moiety of chloramphenicol; non-specific incorporation of the label at relatively high isotopic dilutions indicated that the substrates fed were incorporated via indirect routes.

NAC thioesters have been successfully used as CoA-equivalent substrates for enzymes in fatty acid metabolism, and as vehicles for delivering precursors to the active sites of polyketide synthases, where they are better substrates than the corresponding free carboxylic acids (Tsantrizos *et al.*, 1995). Various NAC thioesters were successfully used in cultures of *Streptomyces longisporoflavus* to study biosynthesis of antibiotic tetronasin (Less *et al.*, 1996). Based on the fact that acetyl-NAC and acetoacetylNAC were stable in the medium, the low incorporation of isotope from the ^{13}C-labeled substrates suggested that acetoacetyl CoA is not a substrate for chlorination. However, a stability study with dichloroacetyl-NAC showed that this substrate was readily

hydrolyzed ($t_{½}$ 1-1.5 h; Lewis, 1998) so it would have been present in the culture medium for only a short time. Therefore, for a dichloroacetyl coenzyme A ester to be a possible intermediate, it would have to be generated and rapidly used at the site of incorporation. In other words, if a chlorinated derivative of acetyl coenzyme A is involved it is very unstable in the environment and could not be formed as a free intermediate. Nevertheless, such an intermediate might be stabilized by being bound to an enzyme, or formed by an enzyme complex that transfers it to the phenylpropanoid intermediate during Cm biosynthesis. This possibility raises a question about the feasibility of detecting the chlorination activity by using a free substrate in standard procedures.

Although the labeling evidence showed that acetate either via acetyl coenzyme A or acetoacetyl coenzyme A is not a direct intermediate for the dichloroacetyl group of Cm, a less direct route with free acetoacetic acid as an intermediate needs to be considered because of the evidence that C-1' of phenylalanine is a relatively specific precursor of the dichloromethine carbon of Cm. From [U-^{14}C] phenylalanine both carbons of the dichloroacetyl group were enriched at about the same specific activities, which means that these two carbons probably are derived as a unit from acetoacetate. Similar results obtained from [U-^{14}C]leucine strengthened this idea (Vining & Westlake, 1964). However, current results from feeding ethyl [2,4-$^{13}C_2$]acetoacetate do not support a role for acetoacetate as a direct precursor of the dichloroacetyl group of Cm.

It seems likely that formation of carbon-chlorine bonds in Cm biosynthesis is not a simple one-step reaction, but a process that requires a substrate formed through several metabolic pathways, and with exchange of carbon atoms among various metabolites,

including carbon dioxide. A multi-step pathway of this type has been postulated in the biosynthesis of bialaphos (Thompson & Seto, 1995).

Section 2. Possible role for 3'-O-acetylchloramphenicol in Cm biosynthesis

Efforts to detect Cm acetyl transferase (CAT) activity in *S. venezuelae* (Shaw & Hopwood, 1976) or to detect genes homologous to known CAT genes (Wynands & van Pée, unpublished) have all failed. These results have suggested that acetylation of chloramphenicol does not occur in *S. venezuelae* (Nakano *et al.*, 1977; Vining & Stuttard, 1995). It was surprising, therefore, to discover 3'-*O*-acetyl Cm accompanying Cm in some ^{13}C enrichment experiments (Groβ *et al.*, 2001). An esterase activity shown to be also present in *S. venezuelae* is suggested to be responsible for hydrolyzing 3'-O-acetyl chloramphenicol within the mycelium. The esterase activity was detected in the particulate fraction of the mycelium, and was present only when cultures were harvested in the late growth phase. Its absence during exponential growth supported the conclusion that hydrolysis of 3'-*O*-acetyl Cm is associated with antibiotic production and its efflux. It is postulated that O-acetylation of an intermediate in the biosynthesis pathway could provide metabolic shielding by creating an O-acetyl ester that will eventually lead to 3'-*O*-acetyl Cm, which lacks antibacterial activity. Subsequent hydrolysis of 3'-*O*-acetyl Cm during efflux generates the active antibiotic outside the mycelium where it can be useful in supporting competition without harming the host. These events are suggested to be part of a self-resistance mechanism (Mosher *et al.*, 1990; Mosher *et al.*, 1995) that has evolved in several antibiotic producers. In the biosynthesis of streptomycin (SM) by *S. griseus*, the antibiotic is produced as SM-6-phosphate, which lacks antibacterial activity.

However, this intermediate is the substrate for a specific SM-phosphate phosphatase that liberates the free antibiotic (Piepersberg, 1995).

Section 3. Search for acyltransferase activity associated with chloramphenicol biosynthesis

After it was established that none of the NAC analogues of acyl-CoA thioesters tested was used directly as a substrate in reactions introducing chlorine into Cm, and given the uncertainty about natural substrates used in the formation of a dichloroacetyl precursor, efforts to identify a chlorinating enzyme were deferred until features of the reaction introducing the dichloroacetyl group into Cm had been investigated further.

Standard procedures of protein purification were utilized to detect and partially purify an N-acetyltransferase that used *p*-nitrophenylserinol (*p*-NPS) as a substrate. However, similar methods were unable to detect N-acetyltransferase activity using *p*-aminophenylserine (*p*-APS) as the substrate in mycelial extracts from Cm-producing *S. venezuelae* cultures.

It was postulated that formation of dichloroacetyl coenzyme A and transfer of the dichloroacetyl group to *p*-APS might be a concerted activity that takes place on a single enzyme complex, which may be unstable or difficult to purify. The evidence (He *et al.*, 2001) that a nonribosomal peptide synthetase (NRPS) is required for activation by adenylation and thioester formation of pathway intermediates, may explain the failure to detect the acetyltransferase activity. The transferase function will be detected only under the particular conditions necessary for activation of the substrate by the NRPS (i.e., using an adenylated derivative of the substrate).

Several other N-acetyltransferases have been reported from streptomycete strains. Among these are puromycin N-acetyltransferase from *Streptomyces alboniger* (Lacalle *et al.*, 1989) and phosphinothricin N-acetyltransferase from *Streptomyces hygroscopicus*. The activity of the latter enzyme confers resistance to the antibiotic phosphinothricin in the producer (Wohlleben *et al.*, 1988), by a mechanism similar to that postulated to shield *S. venezuelae* by acetylation of an intermediate in Cm biosynthesis (see above).

Section 4. PCR amplification of the halogenase gene fragment

Amplification of a halogenase gene fragment from genomic DNA of *S. venezuelae* ISP5230 by touchdown-PCR (TD-PCR) implied the presence of a halogenase gene, and hybridization of the amplified fragment at high stringency to fragments of various sizes from genomic digests provided the necessary evidence that *S. venezuelae* was indeed the source of the DNA. Genes for Cm biosynthesis are expected to be clustered in *S. venezuelae* (Vats *et al.*, 1987), and since the hybridizing fragments from the genome were relatively large (e.g., the 8-kb *Pst*I fragment; Figure 10), they were expected to contain several of the chloramphenicol biosynthesis genes linked to the halogenase gene.

Using pairs of degenerate primers to amplify an intervening gene fragment by PCR has proven to be a successful strategy (Mullis & Faloona, 1987; Turgay & Marahiel, 1994). The method is useful when the sequences of proteins from homologous genes are available, and a high degree of consensus allows degenerate oligonucleotide primers to be designed. Non-ribosomal peptide synthase genes have been cloned from *Pseudomonas syringae* and *Bacillus licheniformis* by this technique (Turgay & Marahiel, 1994).

The TD-PCR method was used in this project in the expectation that it would improve the likelihood of amplifying a halogenase gene fragment. The touch down method increases the chance of specific amplification when primer pools with high degeneracy have to be used. In TD-PCR, lowering the annealing temperature after a set of cycles gives less chance of spurious priming and a lower level of false product amplification (Don *et al.*, 1991); thus the specific product will constitute the major component of amplicons. In other optimization methods the annealing temperature is increased stepwise until the optimum temperature for amplification of a single specific product is achieved; this is a lengthy optimization process, and TD-PCR has been shown to give better results in less time.

In addition to being effective in *S. venezuelae*, primers MPF1 and MPR2 successfully amplified putative halogenase genes from the genomic DNA of five other streptomycetes: *S. armentosus* UC2862, which produces the chlorinated secondary metabolite armentomycin (Liu *et al.*, 1995), *S. griseoviridis*, which produces the chlorinated antitumor antibiotic roseophilin (Hayakawa *et al.*, 1992), *S. aureofaciens* NRRL2209, which produces chlortetracycline (Dairi *et al.*, 1995), *S. glaucescens* and *S. coelicolor* A3(2). Amplification of products from the latter two species with our primers could not be foreseen since these organisms are not reported to produce halogenated metabolites (Gribble, 1996). Consistent with the fact that so far no halogenated metabolites have been reported from *S. parvulus* and *S. akiyoshiensis* (Gribble, 1996), no products were amplified from their DNA. The results suggest that TD-PCR with primers MPF1 and MPR2 can be regarded as a useful indicator of halogenase genes in a

streptomycete genome. However, to reach a final conclusion on primer specificity, extensive optimization of PCR reactions may be needed.

Successful amplification of halogenase fragments from *S. venezuelae* and other streptomycetes is noteworthy, because design of the primers MPF1 and MPR2 was based on two consensus regions in five halogenases, each of which acts on a different substrate with high levels of expected substrate-specificity (van Pée & Holzer, 1999). Moreover, despite the high degeneracy of the primer pools, the halogenase gene fragment was the only product amplified from the *S. venezuelae* genome, and the PCR products amplified from other streptomycetes were all similar in size to the *S. venezuelae* fragment. It is also important to notice that these five halogenases, used for design of primers, catalyze chlorination of an aromatic carbon, while in Cm the dichloroacetyl group is non-aromatic.

This is the first time that amplification by PCR of a halogenase gene fragment has been reported, suggesting that the primers might be useful for amplifying halogenase gene fragments from other *Streptomyces* species. For amplifying halogenase fragments from other bacteria, the primers can be redesigned to improve their specificity and selectivity by taking into account the specific codon usage of organisms other than streptomycetes. The PCR method is more direct and should be faster than procedures such as Tn5 mutagenesis or shotgun cloning currently used for gene cloning.

The halogenase PCR product amplified from *S. venezuelae* ISP5230 genomic DNA (Piraee & Vining, *submitted for publication in J Indust Microbiol Biotech*) was successfully used in this study to screen 6000 phages from a genomic DNA library of the organism, and to isolate a DNA cluster containing genes involved in the chlorination step

of Cm. Sequence analyses revealed the presence of two complete ORFs located in this cluster, and a 5'-truncated ORF. The sequence of the truncated ORF11 was completed by adding 430 nucleotides from an adjacent 7.5-kb *Bam*HI-*Bam*HI chromosomal fragment which was previously isolated from *S. venezuelae* ISP5230 by He *et al.* (2001), and was shown to be involved in Cm biosynthesis. Database similarity comparisons showed that ORF11 encoded an AMP-binding enzyme, and that ORF12 and ORF13 encoded a halogenase and an aldo/keto reductase, respectively. Insertional inactivation of ORF11 and ORF12 abolished Cm production, and caused the mutant strains to accumulate non-chlorinated congeners of Cm (predominantly corynecin II) in the production medium, confirming a role for each of these genes in chlorination of the antibiotic. Southern hybridization analyses of the genomic DNA of each mutant strain confirmed that the native gene had been replaced with a disrupted copy containing the Am^R gene used for insertional inactivation. Corynecin II is a non-chlorinated propionyl homologue of Cm. It is produced along with other corynecins by *S. venezuelae* ISP5230 when the culture is deprived of chloride anions. Cultures of a chloramphenicol-nonproducing *S. venezuelae* mutant (*cml-2*) obtained by chemical mutagenesis also accumulate corynecin II, and were postulated (Doull *et al.*, 1985) to be blocked in the chlorination step. Accumulation of the same non-chlorinated congener of Cm by the strains (VS1101, VS1102) prepared in this study by disrupting the halogenase (*cmlS*) and adjacent gene (*cmlK*) is presumably due to inactivation of enzymes involved in a complex chlorination reaction for Cm biosynthesis. Disruption of the aldo-keto reductase CmlT had no effect on biosynthesis of Cm, and the normal antibiotic titre ruled out its participation in the biosynthetic pathway.

I. Functions of gene products

a. CmlK: the AMP-ligase

The ORF11-encoded protein (CmlK) shares sequence similarity with adenylating enzymes, such as PchD. This enzyme in *P. aeruginosa* is responsible for activating salicylate by adenylation, and forming the tightly bound acyl-AMP-enzyme intermediate prior to its covalent thioesterification with the aryl carrier protein domain of the peptide synthetase PchE during biosynthesis of pyochelin (Quadri *et al.*, 1999). CmlK also showed similarity to DhbE of *B. subtilis*, which is responsible for activating 2,3-dihydroxybenzoate by adenylation prior to its attachment with glycine for biosynthesis of the siderophore dihydroxybenzoyl glycine (Rowland *et al.*, 1996). Therefore, CmlK was expected to function as an AMP-ligase.

The adenylation domain core sequences (Conti *et al.*, 1996) identified in CmlK are (the initial number is the position of the first amino acid):

1) 114-SSGSTGAPK resembling the consensus sequence [STG][STG]-G-[ST] [TSE]-[GS]-x-[PALIVM]-K (Conti *et al.*, 1996), which is the signature sequence for the enzyme superfamily, in which G and K are invariant residues.

2) 252-FGTSE, similar to the consensus sequence [YFW]-[GASW]-x-[TSA]-E in which the glutamic acid (E) is invariant.

3) 297-TGD, resembling the consensus [STA]-[GRK]-D in which aspartic acid (D) is invariant (Conti *et al.*, 1996).

CmlK lacks the 4'-phosphopantetheine-binding site present in non-ribosomal peptide synthetases (Konz & Marahiel, 1999). Common to adenylating enzymes like CmlK (e.g., PchD and EntA) is acyl transfer to the thiol group of an acceptor molecule such as a

second protein or coenzyme A (Adams & Schumann, 1993; Quadri *et al.*, 1999). Based on this feature He *et al.* (2001) suggested that CmlK catalyzes activation of *p*-APA by adenylation to generate the corresponding thioester with a yet-unknown acyl-carrier protein. The thioester-bound *p*-APA then becomes a substrate for a yet-unknown β-hydroxylase enzyme. This proposed mechanism for hydroxylation of *p*-APA during Cm biosynthesis would be similar to the β-hydroxylation of tyrosine by adenylation and formation of an enzyme-bound thioester with NovH during coumarin formation, which takes place during novobiocin biosynthesis in *Streptomyces spheroides*. The NovH-bound tyrosine is the substrate for β-hydroxylation by a cytochrome P450 monooxygenase (Chen & Walsh, 2001). Chen and Walsh (2001) have shown that NovI would not oxygenate free tyrosine or the soluble thioester surrogate L-tyrosyl-*S*-*N*-acetylcysteamine, and recognizes only tyrosine thioesterified to the pantetheinyl arm of the PCP domain of NovH.

Other examples of enzyme pairs of this type involved in adenylation and hydroxylation are CumC/CumD for β-hydroxylation of tyrosine in the coumermycin gene cluster (Wang *et al.*, 2000), and NikP1/NikQ for β-hydroxylation of histidine in the nikkomycin gene cluster (Lauer *et al.*, 2000). Chen and Walsh (2001) have suggested that the β-hydroxylation of a tethered aminoacyl-*S*-PCP might be a general mechanism for controlling the fraction of the cellular pool of the amino acid substrate designated for downstream reactions in an antibiotic biosynthetic pathway. However, CmlK itself lacks the thiolation site, and no other protein that might provide the thiolation site, or partner protein with monooxygenase activity function as a β-hydroxylase has been identified so far in the Cm biosynthetic gene cluster.

Another possible function postulated for CmlK is activation of a substrate for chlorination. Evidence for involvement of CmlK directly in the chlorination reaction of Cm biosynthesis is provided by the result of its disruption, which not only halted Cm production, but also caused accumulation of the non-chlorinated analogue corynecin II in the culture, implicating its involvement in the complicated chlorination reaction. Evidence supporting involvement of CmlK in chlorination rather than β-hydroxylation was obtained from a comparison of its aa sequence with the NRPS (CmlP) in the Cm biosynthesis pathway. CmlP is a mono-modular NRPS suggested to be involved in activating one of the intermediates in the pathway, likely to be *p*-aminophenylserine (He *et al.*, 2001). This conclusion was based on the similarity of the aa sequence of CmlP to sequences of those NRPSs that catalyze activation of phenylalanine, in the biosynthesis of gramicidin (Conti *et al.*, 1997) and tyrocidin (Mootz *et al.*, 1997). Although, CmlK was first suggested to be an activating enzyme involved in adenylation of *p*-aminophenylalanine for β-hydroxylation, its aa sequence lacks significant similarity to CmlP or to other phenylalanine-like activating enzymes. Moreover, the sequence of CmlK showed similarity to enzymes that catalyze adenylation of a different class of aromatic substrates, such as salicylate or 2,3-dihydroxybenzoate. None of these compounds are amino acids, nor are they similar to the phenylpropanoid moiety of Cm. Therefore, CmlK and CmlP may not utilize structurally-related intermediates such as *p*-aminophenylalanine or *p*-aminophenylserine as their substrates, and it seems more likely that the NRPS activates *p*-aminophenylalanine for β-hydroxylation and carries it as a thioester for downstream reactions including dichloroacetyl transfer, while CmlK activates a yet-unknown precursor for chlorination. This precursor may well be an

aromatic compound as in the case of substrates for known halogenases, which after chlorination converts to a dichloroacyl group. For example pyrrolnitrin, pyoluteorin, vancomycin, and chlortetracycline are all chlorinated on their aromatic carbons. An alternative function that could be envisioned for CmlK is the activation of a dichloroacylating agent.

b. CmlS: the halogenase

The protein encoded by ORF12 belongs to the $FADH_2$-dependent halogenases that their encoding genes have been recently isolated from a variety of secondary metabolite producers. CmlS contains a putative NAD(P)H binding site (10-GGGPAGSVAGLTLHKLGH) resembling the consensus **GxGx2(G/A)x3(G/A)x6G**, which is conserved at the N-terminal end of most of halogenases.

CmlS resembles other halogenases in exhibiting similarity to certain FAD-dependent monooxygenases. Similar results were obtained when the sequences of halogenases PrnC, Chl, PltD, PltA or PltM were used as the query sequence (data not shown). Pfam searches to detect conserved domains identified in CmlS a conserved FAD-binding site similar to that of the well-characterized *p*-hydroxybenzoate hydroxylases (Wierenga *et al.*, 1979; Schreuder *et al.*, 1989). Pelzer *et al.* (1999) identified a region with weak similarity to the FAD-binding motif near the C-terminus of halogenase BhaA in the balhimycin biosynthesis gene cluster. A conserved aspartic acid (D) at position 304 inside the putative flavin-binding site of BhaA was suggested to form hydrogen bonds with the flavin moiety of FAD (Pelzer *et al.*, 1999). CmlS contains this conserved Asp in its putative FAD binding site at amino acid 259.

The presence of putative flavin- and nicotinamide-binding sites in CmlS and other halogenases, in addition to the sequence similarity between halogenases and certain FAD-dependent monooxygenases, imply that the two enzyme classes have related catalytic activities. In FAD-dependent hydroxylations the flavin cofactor is reduced in the presence of NAD(P)H and substrate to form the 4(a)-flavin hydroperoxide, which then hydroxylates the substrate (Entsch *et al.*, 1976; Gatti *et al.*, 1994). The catalytic mechanism suggested by Hohaus *et al.* (1997) for halogenation appears to be similar to the mechanism proposed for hydroxylation by PobA enzymes (Dimarco *et al.*, 1993) in which the flavin hydroperoxide initially formed oxidizes the organic substrate to an epoxide, which then undergoes nucleophilic attack. In halogenase-catalyzed reactions the nucleophile is a halide ion, and specific removal of water gives the halogenated product. Sequence and mechanistic similarities between halogenases and FAD-dependent monooxygenases suggest an evolutionary relationship between these enzymes. Figure 36 shows the relatedness of halogenases to certain monooxygenases.

Identification of the putative FAD binding site, and particularly, the similarity of the N-terminal region of CmlS, including its putative NAD(P)H binding site, to a number of reductases may suggest that the reduction of FAD in the presence of NADPH is catalyzed by the halogenase itself. This could occur in a fashion similar to that explained for the FAD-dependent *p*-hydroxybenzoate hydroxylase (Entsch *et al.*, 1976; Gatti *et al.*, 1994). The N-terminal regions of halogenases PrnA, PltA, the putative halogenase AdpC from *Anabaena* sp. 90 (accession number CAC01605) and the one from *Amycolatopsis orientalis* also exhibited similarity to certain oxidoredutases (data not shown). Whether CmlS is similar to PrnA and PrnC in requiring a separate flavin reductase (Keller *et al.*,

2000), or the flavin reductase activity is associated with its putative FAD binding site is not yet known. However, Keller *et al.* (2000) have suggested that a separate FAD reductase is required.

Finally it can be assumed that crystallization and X-ray analysis will give opportunities for comparative studies of halogenases and other flavin-dependent monooxygenases that have already been crystallized, and from which preliminary X-ray data are available (Schreuder *et al.*, 1989). Halogenases show substrate- and regio-selectivity (van Pée & Holzer, 1999) and could potentially be useful for the production of fine chemicals as an alternative to currently used chemical halogenation processes that produce unacceptable amounts of by-products

c. CmlT: the aldo/keto reductase

CmlT is similar to enzymes of the aldo-keto reductase (AKR) family that catalyze reduction of specific carbonyl carbons in the presence of NADPH (Jez *et al.*, 1997; Khurana *et al.*, 1998). Members of the AKR family have about 300 amino acids in parallel $(\beta/\alpha/\beta)_8$-barrel structures containing a novel NADP-binding motif located in the C-terminal region (Jez *et al.*, 1997). Two motifs matching those conserved in members of the AKR family were identified in CmlT (Table 6). Information obtained by aligning the sequences of a large number of enzymes from the AKR family, and examining the crystal structures of the enzymes has been used to determine the general features of the AKR family. For example, it has been found that the binding pocket is hydrophobic in nature, favoring aromatic and non-polar substrates over highly polar ones (Jez *et al.*, 1997; Khurana *et al.*, 1998), and that binding of the nicotinamide cofactor does not require a Rossman fold motif. The latter structural feature differentiates AKR from other

oxidoreductases (Khurana *et al.*, 1998). The catalytic mechanism common to members of the AKR superfamily involves stereospecific transfer of a hydride ion (H^-) from NADPH to the substrate carbonyl atom, followed by the transfer of a proton (H^+), most likely from

Table 6. Aldo/keto-reductase motifs* identified in CmlT

Sequence in CmlT	Consensus for aldo/keto reductase
46-GVTFLDTAEIYGPYTNER	G-[FY]-R-[HSAL]-[LIVMF]-D-[STAGC]-[AS]-x(5)-E- x(2)-[LIVM]-G
130-LAELVAEGKIRHIGLSEA	[LIVMFY]-x(9)-[KREQ]-x-[LIVM]-G-[LIVM]-[SC]-N-[FY]

- The number preceding the sequences in CmlT indicates the position of the first motif amino acid.

* Based on the aldo/keto reductase motifs reported by Jez *et al.* (1997)

a conserved tyrosine (i.e., Y-55), to the substrate oxygen (Khurana *et al.*, 1998). The amino acids D-51, A-53, Y-56, K-81, and H-117 in CmlT correspond to the conserved amino acids D-50, A-52, Y-55, K-84, and H-117 identified in the active sites of AKRs by examination of their crystal structures and comparisons of their amino acid sequences (Jez *et al.*, 1997).

A possible role for CmlT in reducing the aldehyde carbonyl of N-dichloroacetyl-*p*-aminophenylserinal to the corresponding primary alcohol group during Cm biosynthesis was refuted by the results of gene knock-out experiments. The putative reductase domain identified in CmlP is expected to catalyze the reductive reaction. This type of reductive release appears to be a general termination mechanism in nonribosomal peptide biosynthesis. The enzyme uses NADPH to release the peptidyl-carrier protein-bound thioester as an aldehyde, which is then further reduced to an alcohol. Several NRPSs with the reductase domain can be found in GenBank. The MxcG, which is involved in biosynthesis of the iron-chelating compound myxochelin in *Stigmatella aurantiaca* Sg a15 (Gaitatzis *et al.*, 2001), and the saframycin Mx1 synthetase A that participates in biosynthesis of the antibiotic saframycin by *Myxococcus xanthus* (Pospiech *et al.*, 1996) are known to possess reductase domains.

II. Analysis of the sequence downstream of ORF13

Chromosome walking downstream of ORF13 identified two potential ORFs. The two ORFs read in the same direction and are separated by short non-translated gap of 156 bp. No function could be postulated for the products of these two genes since similarity comparisons identified proteins from *S. coelicolor* with unknown functions. It appears

likely that the putative ORFs in this part of the *S. venezuelae* chromosome are not involved in Cm biosynthesis.

III. Complementation of the *cml-2* mutant with cloned genes

Conjugal transfer to the *cml-2* mutant of the cloned genes provided by plasmids pJV507, pJV526 and pJV528 that contain genes involved in Cm biosynthesis failed to complement the mutation(s) causing blockage of Cm biosynthesis. All transconjugants produced corynecin II at the same level produced by the parent *cml-2* strain. Expectedly, transconjugants carrying segments of chromosomal DNA containing only ORF13, provided by plasmid pJV528, or ORFs 12 and 13, provided by plasmid pJV507, were also unable to produce Cm. Consistent with this result, Doull *et al.* (1985) did not detect cosynthesis of Cm by growing mixtures of the mutant with other blocked mutants of *S. venezuelae* ISP5230. Moreover, an earlier attempt in which J. Y. He and I (unpublished) tried to complement the *cml-2* mutant by conjugally transferring into it a genomic DNA library of *S. venezuelae* prepared in plasmid pKC1218 failed to detect any Cm producer among 1300 transconjugants tested. "Multi-site" mutations caused by the "random hits" during the chemical mutagenesis used to obtain the mutant may explain the results.

Section 5. The Cm biosynthesis gene cluster

So far thirteen ORFs have been cloned from the Cm biosynthetic pathway, nine of them shown by knockout experiments to be involved in Cm biosynthesis. ORF8 is probably the most interesting gene in the pathway since it encodes a mono-modular nonribosomal peptide synthetase (CmlP). Cm is a relatively simple molecule and no peptide formation is apparent in the chemical structure. It appears that in *S. venezuelae* a novel tool for biosynthesis of a secondary metabolite has evolved by utilizing an NRPS.

It is likely that CmlP provides a protective and channeling function for the pathway intermediates, shielding them from other competing reactions. At some point during the pathway the NRPS activates a precursor amino acid by adenylation (He *et al.*, 2001), and protects it by thioester bond formation, at the same time delivering it to the proper downstream reactions away from metabolizing enzymes. Sequence analyses indicating the specificity of the adenylation domains, showed that CmlP has the required functional motifs for activation of a phenylalanine-like amino acid, suggested earlier (He *et al.*, 2001) to be *p*-aminophenylserine.

ORF12 of the cluster is believed to encode a halogenase (CmlS) and was shown to be involved in the halogenation step for Cm biosynthesis by a knockout experiment. However, the nature of the substrate for CmlS still remains in question. Identification of the chlorination substrate also will help to determine the mechanism for chlorination since the final product is chlorinated on an acyl carbon, not at an aromatic carbon as with other known halogenases. The mechanism of FAD-dependent halogenation, as explained for tryptophan by Hohaus *et al.* (1997), may not be applicable to Cm biosynthesis if the chlorination substrate is not an aromatic metabolite. The substrate specificity of tryptophan halogenase has been tested using a number of related structures (van Pée & Hozler, 1999). It will be interesting to find out if Cm halogenase can be exchanged with halogenases from other pathways through combinatorial biosynthesis and to determine its ability to chlorinate a set of selected substrates.

Appendix

Appendix I. ^{1}H-NMR spectra

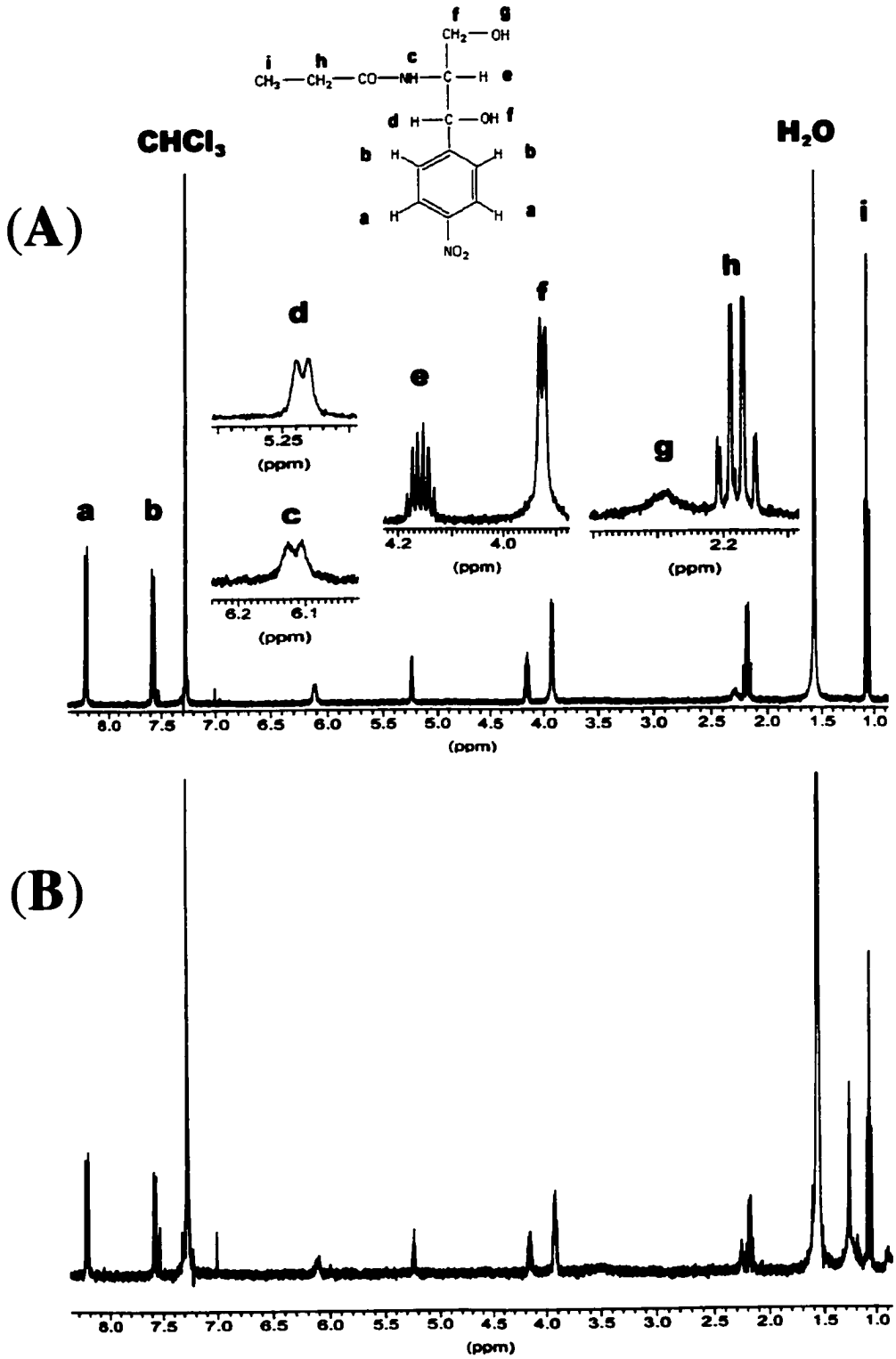


Figure 47. ^{1}H-NMR spectra of (**A**) the compound purified from cultures of mutant VS1102, and (**B**) an authentic sample of corynecin II. Assignment of signals in the spectrum of (A) was consistent with the chemical structure of corynecin II.

References

Adams, R., & Schumann, W. (1993). Cloning and mapping of the *Bacillus subtilis* locus homologous to *Escherichia coli ent* genes. *Gene* **133,** 119-121.

Ahmed, Z. U., & Vining, L. C. (1983). Evidence for a chromosomal location of genes coding for chloramphenicol production in *Streptomyces venezuelae*. *J Bacteriol* **154,** 239-244.

Akagawa, H., Okanishi, M., & Umezawa, H. (1979). Genetics and biochemical studies of chloramphenicol-nonproducing mutants of *Streptomyces venezuelae* carrying plasmid. *J Antibiot* **32,** 610-620.

Amann, R., Wolfgang, L., & Schleifer, K-H. (1994). Identification of uncultured bacteria: A challenging task for molecular taxonomists. *ASM News* **60,** 360-365.

Arias, P., Fernandez-Moreno, M. A., & Malpartida, F. (1999). Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2) as a DNA-binding protein. *J Bacteriol* **181,** 6958-68.

August, P. R., Yu, T-W., & Floss, H. G. (1999). Molecular biological aspects of antibiotic biosynthesis. In: *Drug Discovery from Nature*. Grabley. S., & Thiericke, R. (eds.). Springer-Verlag, New York, pp. 215-232.

Beckwith, J. R., & Hager, L. P. (1963). Biological chlorination. VIII. Late intermediates in biosynthesis of caldariomycin. *J Biol Chem* **238,** 3091-3094.

Bongs, G., & van Pée, K-H. (1994). Enzymatic chlorination using bacterial non heme chloroperoxidases. *Enzym Microbiol Technol* **16,** 53-60.

Brown, M. P., Aidoo, K. A., & Vining, L. C. (1996). A role for *pabAB*, a *p*-aminobenzoate synthase gene of *Streptomyces venezuelae* ISP5230, in chloramphenicol biosynthesis. *Microbiol* **142,** 1345-1355.

Cane, D. E. (1995). Isoprenoid antibiotics. In: *Genetics and Biochemistry of Antibiotic Production*. Vining, L. C., & Stuttard, C. (eds.), Butterworth-Heinemann, Newton, MA. pp. 633-655.

Chang, Z. (1999). *Genes for cysteine biosynthesis and metabolism in Streptomyces venezuelae* ISP5230*: cloning, sequencing, functional analysis and relevance to chloramphenicol biosynthesis.* Ph.D. thesis, Dalhousie University, NS, Canada.

Chang, Z., Sun, Y., He, J., & Vining, L. C. (2001). *p*-aminobenzoic acid and chloramphenicol biosynthesis in *Streptomyces venezuelae*: gene sets for a key enzyme, 4-amino-4-deoxychorismate synthase. *Microbiol* **147,** 2113-2126.

Chatterjee, S., Vining, L. C., & Westlake, D. W. S. (1982). Nutritional requirements for chloramphenicol biosynthesis in *Streptomyces venezuelae*. *Can J Microbiol* **29,** 247-253.

Chatterjee, S., & Vining, L. C. (1981). Nutrient utilization in actinomycetes. Induction of α–glucosidases in *Streptomyces venezuelae*. *Can J Microbiol* **27,** 639-645.

Chen, H, & Walsh, C. T. (2001). Coumarin formation in novobiocin biosynthesis: β-hydroxylation of the aminoacyl enzyme tyrosyl-*S*-NovH by a cytochrome P450 NovI. *Chem & Biol* **74,** 1-12.

Chiu, H. T., Hubbard, B. K., Shah, A. N., Eide, J., Fredenburg, R. A., Walsh, C. T. & Khosla, C. (2001). Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. *Proc Natl Acad Sci USA* **98,** 8548-8553.

Conti, E., Franks, N. P., & Brick, P. (1996). Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* **4,** 287-298.

Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J* **16,** 4174-4183.

Dairi, T., Nakano, T., Aisaka, K., Katsumata, R., & Hasegawa, M. (1995). Cloning and nucleotide sequence of the gene responsible for chlorination of tetracycline. *Biosci Biotech Biochem* **59,** 1099-1106.

Demain A. L. (1974). How do antibiotic-producing microorganisms avoid suicide. *Annl N Y Acad Sci* **235,** 601-612.

Demain A. L. (1998). Microbial natural products: Alive and well in 1998. *Nature Biotechnol* **16,** 3-4.

Desnottes, J-F. (1996). New targets and strategies for the development of antibacterial agents. *Trends Biotechnol* **14,** 134-140.

DiMarco, A. A., Averhoff, B. A., Kim, E. E., & Ornston, L. N. (1993). Evolutionary divergence of *pobA*, the structural gene encoding *p*-hydroxybenzoate hydroxylase in an *Acinetobacter calcoaceticus* strain well-suited for genetic analysis. *Gene* **125,** 25-33.

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., & Mattick, J. S. (1991). 'Touch down' PCR to circumvent spurious priming during gene amplification. *Nucl Acids Res* **19,** 4008.

Doull, J. L., Ahmed, Z., Stuttard, C., & Vining, L. C. (1985). Isolation and characterization of *Streptomyces venezuelae* mutants blocked in chloramphenicol biosynthesis. *J Gen Microbiol* **131,** 97-104.

Doull, J. L., Vats, S., Chaliciopoulos, M., Stuttard, C., Wong, K., & Vining, L. C. (1986). Conjugational fertility and location of chloramphenicol biosynthesis genes on the chromosomal linkage map of *Streptomyces venezuelae*. *J Gen Microbiol* **132,** 1327-1338.

Dunford H. B., Lambeir, A., Kashem, M., & Pickard, M. (1987). On the mechanism of chlorination by chloroperoxidase *Arch Biochem Biophys* **252,** 292-302.

Entsch, B., Ballou, D. P., & Massey, V. (1976). Flavin-oxygen derivatives involved in hydroxylation by *p*-hydroxybenzoate hydroxylase. *J Biol Chem* **251,** 2550-2563.

Facey, S. J., Groβ, F., Vining, L. C., Yang, K., & van Pée, K-H. (1996). Cloning, sequencing and disruption of a bromoperoxidase-catalase gene in *Streptomyces venezuelae*: evidence that it is not required for chlorination in chloramphenicol biosynthesis. *Microbiol* **142,** 657-66.

Flett, F., Mersinias, V., & Smith, C. P. (1997). High efficiency intergenic conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* **155,** 223-229.

Francis, M. M., Vining, L. C., & Westlake, D. W. S. (1978). Characterization and regulation of anthranilate synthase from a chloramphenicol-producing *Streptomyces*. *J Bacteriol* **134,** 10-16.

Gaitatzis, N., Kunze, B., & Müller, R. (2001). In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. *Proc Natl Acad Sci USA* **98,** 11136-11141.

Gatti, D. L., Palfey, B. A., Lah, M. S., Entsch, B., Massey, V., Ballou, D. P., & Ludwig, M. L. (1994). The mobile flavin of 4-OH benzoate hydroxylase. *Science* **266,** 110-114.

Genetics Computer Group. 1994. Sequence analysis software package, University of Wisconsin Genetics Computer Group, Madison, Wis.

Gladek, A., & Zakrzewska, J. (1984). Genome size of *Streptomyces*. *FEMS Microbiol Lett* **24,** 73-76.

Glazer, A. N., & Nikaido, H. (1995). Microbial biotechnology: fundamentals of applied microbiology. W. H. Freeman & Co. New York, pp. 431-515.

Gottlieb, D., Carter, H. E., Robbins, P. W., & Burg, R. W. (1962). Biosynthesis of chloramphenicol IV. Incorporation of carbon^{14}-labeled precursors. *J Bacteriol* **84,** 888-894.

Gottlieb, D., Robbins, P. W., & Carter, H. E. (1956). The biosynthesis of chloramphenicol. II. Acetylation of *p*-nitrophenylserinol. *J Bacteriol* **72,** 153-156.

Gribble, G. W. (1996). Naturally occurring organohalogen compounds-A comprehensive survey. Fortschr Chem Org Naturst 68, 1-423.

Groβ, F., Lewis, E. A., Piraee, M., van Pée, K-H, Vining, L. C., & White, R. L. (2001). Isolation of 3'-*O*-acetylchloramphenicol: A possible intermediate in chloramphenicol biosynthesis. *Bioorg Med Chem Lett. In Press.*

Hager, L. P., Morris, D. R., Brown, F. S., & Eberwein, H. (1966). Chloroperoxidase. II. Utilization of halogen anions. *J Biol Chem* **241,**1769-77.

Hammer, P. E., Burd, W., Hill, D. S., Ligon, J. M., & van Pée, K-H. (1999). Conservation of the pyrrolnitrin biosynthetic gene cluster among six pyrrolnitrin-producing strains. *FEMS Microbiol Lett* **180,** 39-44.

Hammer, P. E., Hill, D. S., Lam, S. T., van Pée, K-H, & Ligon, J. M. (1997). Four genes from *Psuedomonas fluorescens* that encode the biosynthesis of pyrrolnitrin. *Appl Environ Microbiol* **63,** 2147-2154.

Hansch, C., Nakamoto, K., Gorin, M., & Denisevich, P. (1973). Structure-activity relationship of chloramphenicols. *J Med Chem* **16,** 917-922.

Hayakawa, Y., Kawakami, K., & Seto, H. (1992). Structure of a new antibiotic, roseophilin. *Tetrahedron Lett* **33,** 2701-2704.

He, J-Y., Magarvey, N., Piraee, M., & Vining, L. C. (2001). The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular nonribosomal peptide synthetase gene. *Microbiol* **147,** 2817-2829.

Hecht, H. J., Sobek, H., Haag, T., Pfeifer, O., & van Pée, K-H. (1994). The metal-ion-free oxidoreductase from *Streptomyces aureofaciens* has an α/β hydrolase fold. *Nat Struct Biol* **1,** 532-537.

Hofmann, B., Tolzer, S., Pelletier, I., Altenbuchner J., van Pée, K-H., & Hecht, H. J. (1998). Structural investigation of the cofactor-free chloroperoxidases. *J Mol Biol* **279,** 889-900.

Hohaus, K., Altmann, A., Burd, W., Fischer, I., Hammer, P. E., Hill, D. S., Ligon, J. M., & van Pée, K-H. (1997). NADH-dependent halogenases are more likely to be involved in halometabolite biosynthesis than haloperoxidases. *Angew Chem Int Ed Engl* **36,** 2012-2013.

Hopwood D. A., Bibb, M. J., Bruton, C. J., Chater, K. F., Feitelson, J. S., & Gil, J. A. (1983). Cloning *Streptomyces* genes for antibiotic production. *Trends Biotech* **1,** 42-48.

Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, C. P., Smith, C. P., Ward, J. M., & Schrempf, H. (1985a). Genetic Manipulation of *Streptomyces*. A Laboratory Manual. John Innes Foundation. Norwich.

Hopwood, D. A., Malpartida, F., Kieser, H. M., Ikeda, H., Duncan, J., Fuji, I., Rudd, B. A. M., Floss, H. G., & Omura, S. (1985b). Production of 'hybrid' antibiotics by genetic engineering. *Nature* **314,** 642-644.

Horinouchi, S., & Beppu, T. (1995). Autoregulators. In: *Genetics and Biochemistry of Antibiotic Production.* Vining, L. C., & Stuttard, C. (eds.), Butterworth-Heinemann, Newton, MA. pp. 103-119.

Hotta, K., Davies, J., & Yagisawa, M. (1995). Aminoglycosides and Aminocyclitols. In: *Genetics and Biochemistry of Antibiotic Production.* Vining, L. C., & Stuttard, C. (eds.), Butterworth-Heinemann, Newton, MA. pp. 571-595.

Hotta, K., & Okami, Y. (1996). Diversity in aminoglycoside antibiotic resistance of actinomycetes and its exploitation in the search for novel antibiotics. *J Indust Microbiol* **17,** 352-358.

Hotta, K., Takahashi, A., Okami, Y., & Umezawa, H. (1983). Relationship between antibiotic resistance and antibiotic productivity in actinomycetes which produce aminoglycoside antibiotics, *J Antibiot* **36,** 1789-1791.

Huang, L-M., & Jeang, K-T. (1994). PCR primers containing an inosine triplet to complement a variable codon within a conserved protein-coding region. *Biotechniques* **16,** 246- 250.

Hutchinson, C. R. (1999). Combinatorial biosynthesis of antibiotics. In: *Drug Discovery from Nature.* Grabley. S., & Thiericke, R. (eds.). Springer-Verlag, New York, pp. 233-254.

Ishikawa, J., & Hotta, K. (1999). Frameplot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. *FEMS Microbiol Lett* **17,** 251-253.

Jez, J. M., Bennet, M. K. J., Schlegel, B. P., Lewis, M., & Penning, T. M. (1997). Comparative anatomy of the aldo–keto reductase superfamily. *Biochem J* **326,** 625–636.

Jones, A., & Vining, L. C. (1976). Biosynthesis of chloramphenicol in *Streptomyces* sp. 3022a. Identification of *p*-amino-L-phenylalanine as a product from the action of arylamine synthetase on chorismic acid. *Can J Microbiol* **22,** 237-244.

Jones, A., & Westlake, D. W. S. (1974). Regulation of chloramphenicol synthesis in *Streptomyces* sp.3022a. Properties of arylamine synthetase, an enzyme involved in antibiotic biosynthesis *Can J Microbiol* **20,** 1599-1611.

Kaneko, T., Nakamura, Y., Sato, S. & 21 other authors (2001). Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti. DNA Res* **7,** 331-338.

Keller, S., Wage, T., Hohaus, K., Holzer, M., Eichhorn, E., & van Pée, K-H. (2000). Purification and partial characterization of tryptophan 7-halogenase (PrnA) from *Psuedomonas fluorescens. Angew Chem Int Ed* **39,** 2300-2302

Khurana , S., Powers, D. B., Anderson, S., & Blaber, M. (1998). Crystal structure of 2,5-diketo-D-gluconic acid reductase A complexed with NADPH at 2.1-Å resolution. *Proc Natl Acad Sci USA* **95,** 6768-6773.

Kirner, S., Hammer, P. E., Hill, D. S. Altmann, A., Fischer, I., Weislo, L. J., Lanahan, M., van Pée, K-H., & Ligon, J. M. (1998). Functions encoded by pyrrolnitrin biosynthetic genes from *Pseudomonas fluorescens. J Bacteriol* **180,** 1939-1943.

Knoch, M., van Pée, K-H., Vining, L. C., & Lingens, F. (1989). Purification, properties and immunological detection of a bromoperoxidase-catalase from *Streptomyces venezuelae* and from a chloramphenicol-nonproducing mutant. *J Gen Microbiol* **135,** 2493-2502.

Konz, D., & Marahiel, M. A. (1999). How do peptide synthetases generate structural diversity? *Chem & Biol* **6,** R39-R48.

Kuntz, I. D. (1992). Structure-based strategies for drug design and discovery. *Science* **257**, 1078-1082.

Lacalle, R. A., Pulido, D., Vara, J., Zalacain, M. & Jimenez, A. (1989). Molecular analysis of the *pac* gene encoding a puromycin N-acetyltransferase from *Streptomyces alboniger. Gene* **79,** 375-380.

Lancini, G., & Lorenzetti, R. (1993). Biotechnology of antibiotics and other bioactive microbial metabolites. Plenum Press. New York, pp. 145-160.

Larson, J. L., & Hershberger, C. L. (1986). The minimal replicon of a streptomycete plasmid produces an ultrahigh level of plasmid DNA. *Plasmid* **15,** 199-209.

Lauer, B., Russwurm, R. & Bormann, C. (2000). Molecular characterization of two genes from *Streptomyces tendae* Tu901 required for the formation of the 4-formyl-4-imidazolin-2-one-containing nucleoside moiety of the peptidyl nucleoside antibiotic nikkomycin. *Eur J Biochem* **267,** 1698-1706.

Less, S. L., Handa, S., Leadlay, P. F., Dutton, C. J., & Staunton, J. (1996). Biosynthesis of tetronasin: Part 5. Novel fluorinated and non-fluorinated analogues of tetronasin *via* intact incorporation of di-, tri- and tetraketide analogue precursors. *Tetrahedron Lett* 37, 3511-3514.

Lewis, E. A. (1998). *The biosynthesis of chloramphenicol. M.Sc. Thesis.* Chemistry Department, Dalhousie University, Canada.

Libby, R. D., & Rotberg, N. S. (1990). Compound I formation is a partially rate-limiting process in chloroperoxidase-catalyzed bromination reactions. *J Biol Chem* **265,** 14808-14811.

Libby, R. D., Thomas, J. A., Kaiser, L. W., & Hager, L. P. (1982). Chloroperoxidase halogenation reactions; chemical versus enzymic halogenating intermediates. *J Biol Chem* **257,** 5030-5037.

Liu, K., White, R. L., He, J-Y., & Vining, L. C. (1995). Biosynthesis of armentomycin: a chlorinated nonprotein amino acid. *J Antibiot* **48,** 347-348.

Long, L. M., & Troutman, H. D. (1949). Chloramphenicol (Chloromycetin). VII. Synthesis through *p*-nitroacetophenone. *J Am Chem Soc* **71,** 2473-2475.

Lowe, D. A., & Westlake D. W. S. (1971). Regulation of chloramphenicol synthesis in *Streptomyces* sp. 3022a. 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthetase. *Can J Microbiol* **49,** 448-455.

Lowe, D. A., & Westlake D. W. S. (1972). Regulation of chloramphenicol synthesis in *Streptomyces* sp. 3022a. Branch-point enzymes of the shikimic acid pathway. *Can J Biochem* **50,** 1064-1073.

Malik, V. S., & Vining, L. C. (1970). Metabolism of chloramphenicol by the producing organism. *Can J Microbiol* **16,** 173-179.

Malik. V. S., & Vining, L. C. (1971). Metabolism of chloramphenicol by the producing organism. Some properties of chloramphenicol hydrolase. *Can J Microbiol* **17,** 1287-1290.

Malik. V. S., & Vining, L. C. (1972). Effect of chloramphenicol on its biosynthesis by *Streptomyces* species 3022a. *Can J Microbiol* **18,** 137-143.

Malmberg, L. H., HU, W-S., & Sherman, D. H. (1993). Precursor flux control through targeted chromosomal insertion of the lysine ε-aminotransferase (*lat*) gene in cephamycin C biosynthesis. *J Bacteriol* **175,** 6916-6924.

Malpartida, F., Hallam, S. E., Kieser, H. M., Motamedi, H., Hutchinson, C. R., Butler, M. J., & Hopwood, D. A. (1987). Homology between *Streptomyces* genes coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes. *Nature* **325,** 818-821.

Malpartida, F., & Hopwood, D. A. (1984). Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature* **309,** 462- 464.

Maplestone, R. A., Stone, M. J., & Williams, D. H. (1992). The evolutionary role of secondary metabolites – a review. *Gene* **115,** 151-157.

Martin, A. R. (1998). Antibacterial antibiotics. In: *Wilson and Gisvold's Textbook of Organic Medicinal and Pharmaceutical Chemistry.* 10th edition. Delgado, J. N. & Remers, W. A. (eds.), Lippincott-Raven Publishers, Philadelphia, pp. 253-325.

Mazodier, P., Petter, R., & Thompson, C. (1989). Intergenic conjugation between *Escherichia coli* and *Streptomyces* species. *J Bacteriol* **171,** 3583-3585.

McGrath, R., Vining, L. C., Sala, F., & Westlake, D. W. S. (1968). Biosynthesis of chloramphenicol. III. Phenylpropanoid intermediates. *Can J Biochem* **46,** 587-594.

McVeigh, H. P., Munro, J., & Embley, T. M. (1996). Molecular evidence for the presence of novel actinomycete lineages in a temperate forest soil. *J Indust Microbiol* **17,** 197-204.

Mitscher, L. A. (1995). Antibiotics and antimicrobial agents. In: *Principles of Medicinal Chemistry.* 4th edition. Foye, W. O., Lemke, T. L., & Williams, D. A. (eds.), Williams & Wilkins, New York. 759-802.

Mootz, H. D. & Marahiel, M. A. (1997). The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J Bacteriol* **179,** 6843-6850.

Morris, D. R., & Hager, L. P. (1966). Chloroperoxidase I. Isolation and properties of the crystalline glycoprotein *J Biol Chem* **241,** 1763-1768.

Mosher, R. H., Camp, D. J., Yang, K., Brown, M. P., Shaw, W. V., & Vining, L. C. (1995). Inactivation of chloramphenicol by *O*-phosphorylation. *J Biol Chem* **10,** 27000-27006.

Mosher, R. H., Paradkar, A. S., Anders, C., Barton, B. & Jensen, S. E. (1999). Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminate synthase 1 in *Streptomyces clavuligerus. Antimicrob Agents Chemother* **43,** 1215-1224.

Mosher, R. H., Ranade, N. P., Schrempf., H., & Vining, L. C. (1990). Chloramphenicol resistance in *Streptomyces*: cloning and characterization of a chloramphenicol hydrolase gene from *Streptomyces venezuelae*. *J Gen Microbiol* **136,** 293-301.

Mullis, K. B., & Faloona, F. (1987). Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Methods Enzymol* **155,** 335-350.

Munro, M. H. G., Taniguchi, M., Rinehart, K. L., & Gottlieb, D. (1975). A CMR study of the biosynthesis of chloramphenicol. *Tetrahedron Lett* **31,** 2659-2662.

Nakano, H., Matsuhashi, Y., Takeuchi, T., & Umezawa, H. (1977). Distribution of chloramphenicol acetyltransferase and chloramphenicol-3-acetate esterase among *Streptomyces* and *Corynebacterium*. *J Antibiot* **XXX,** 76-82.

Neidleman, S. L., (1975). Microbial halogenation. *CRC Crit Rev Microbiol* **5,** 333-358.

Neidleman, S. L., & Geigert, J. (1986). Biohalogenation: Principles, basic roles and applications. Ellis Horwood Ltd., Chichester.

Nowak-Thompson, B. , Chaney, N., Wing, J. S., Gould, S. J., & Loper, J. E. (1999). Characterization of the pyoluteorin biosynthesis gene cluster of *Pseudomonas fluorescens* Pf-5. *J Bacteriol* **181,** 2166-2174.

Omura, S. (1986). Philosophy of new drug discovery. *Microbiol Rev* **50,** 259-279.

Omura, S. (1992). Trends in the search for bioactive microbial metabolites. *J Ind Microbiol* **10,** 135-156.

O'Neill, W. P., Nstrom, R. F., Rinehart, K. L., & Gottlieb, D. (1973). Biosynthesis of chloramphenicol. Origin and degradation of the aromatic ring. *Biochem* **12,** 4775-4784.

Page, R. M. (1996). TreeView: An application to display phylogenetic trees on personal computers. *Comp App Biosci* **12,** 357-358.

Paradkar, A. S., & Jensen, S. E. (1995). Functional analysis of the gene encoding the clavaminate synthase 2 isoenzyme involved in clavulanic biosynthesis in *Streptomyces clavuligerus*. *J Bacteriol* **177,** 1307-1314.

Pelzer, S., Suβmuth, R., Heckmann, S., Recktenwald, J., Huber, P., Jung, G., & Wohlleben, W. (1999). Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908. *Antimicrob Agents Chemother* **43,** 1565-1573.

Piepersberg, W. (1995). Streptomycin and related aminoglycosides: In: *Genetics and Biochemistry of Antibiotic Production*. Vining, L. C., & Stuttard, C., (eds.), Newton, MA: Butterworth-Heinemann, pp. 531-570.

Pospiech, A., Bietenhader, J., & Schupp, T. (1996). Two multifunctional peptide synthetases and an O-methyltransferase are involved in the biosynthesis of the DNA-binding antibiotic and antitumour agent saframycin Mx1 from *Myxococcus xanthus*. *Microbiol* **142,** 741-746.

Prutz, W. A. (1996). Hypochlorous acid interactions with thiols, nucleotides, DNA, and other biological substrates. *Arch Biochem Biophys* **332,** 110-120.

Quadri, L. E., Keating, T. A., Patel, H. M. & Walsh, C. T. (1999). Assembly of the *Pseudomonas aeruginosa* nonribosomal peptide siderophore pyochelin: *In vitro* reconstitution of aryl-4,2-bisthiazoline synthetase activity from PchD, PchE, and PchF. *Biochem* **38,** 14941-14954.

Ramakrishnan, K., Oppenhuizen, M. K., & Fisher, J. (1983). Stereoselectivity of chloroperoxidase-dependent halogenation. *Biochem* **22,** 3271-3277.

Redenbach, M., Kieser, H. M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H., & Hopwood, D. A. (1996). A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome. *Mol Microbiol* **21,** 77-96.

Retzlaff, L., Mayer, G., Beyer, S., Ahlert, J., Verseck, S., Distler, J., & Piepersberg, W. (1993). Streptomycin production in streptomycetes: a progress report. In: *Industrial Microorganisms: Basic and Applied Molecular Genetics*. Baltz, R. H., Hegeman, G. D., & Skatrud, P. L. (eds.), American Society for Microbiology, Washington D.C. pp.183-194.

Rohner, J., Susanne, E., & Zeek, A. (1988). Metabolic products of microorganisms. tetracenomycins B3 and D3 key intermediates of the elloramycin and tetracenomycin C biosynthesis. *J Antibiot* **XII,** 1066-1073.

Rowland, B. M., Grissman, T. H., Osburne, M. S., & Taber, H. W. (1996). Sequence and genetic organization of a *Bacillus subtilis* operon encoding 2,3-dihydroxybenzoate biosynthetic enzymes. *Gene* **178,** 119-123.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S., & Coulson A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74,** 5463-5467.

Schreuder, H. A., Prick, P. A. J., Wierenga, R. K., Vriend, G., Wilson, K. S., Hol, W. G. J., & Drenth, J. (1989). Crystal structure of the *p*-hydroxybenzoate hydroxylase-substrate complex refined at 1.9 Å resolution. *J Mol Biol* **208,** 679-696.

Scrutton, N. S., Berry, A., & Perham, R. N. (1990). Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* **343**, 38-43.

Shapiro, S., & Vining, L. C. (1983). Nitrogen metabolism and chloramphenicol production in *Streptomyces venezuelae*. *Can J Microbiol* **29,** 1706-1714.

Shaw, W. V. (1983). Chloramphenicol acetyltransferase: enzymology and molecular biology. *CRC Crit Rev Biochem* **14,** 1-46.

Shaw, W. V., & Hopwood, D. A. (1976). Chloramphenicol acetylation in *Streptomyces*. *J Gen Microbiol* **94,** 159-166.

Sherman, D. H., Malpartida, F., Bibb, M. J., Kieser, H. M., & Hopwood, D. A. (1989). Structure and deduced function of the granaticin-producing polyketide synthase gene cluster of *Streptomyces violaceoruber* Tu22. *EMBO J* **8,** 2717-25.

Simonsen, J. N., Paramasigamani, R., Vining, L. C., McInnes, A. G., Walter, J. A., & Wright, L. C. (1978). Biosynthesis of chloramphenicol. Studies on the origin of the dichloroacetyl moiety. *Can J Microbiol* **24,** 136-142.

Stemma, W. P. C. (1994). DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. *Proc Natl Acad Sci USA* **91,** 1074-1075.

Strohl, W. (1997). Industrial Antibiotics: Today and the Future. In: *Biotechnology of Antibiotics*, Marcel Dekker, NY, pp. 1-49.

Stuttard, C. (1982). Temperate phages of *Streptomyces venezuelae*: lysogeny and host specificity shown by SV1 and SV2. *J Gen Microbiol* **128,** 115-121.

Stutzman-Engwall, K. J., Otten, S. L., & Hutchinson, C. R. (1992). Regulation of secondary metabolism in *Streptomyces* spp. and overproduction of daunorubicin in *Streptomyces peucetius*. *J Bacteriol* **174,** 144-54.

Suyes, R. B., Sykes, R. B., Cimarusti, C. M., Bonner, D. P., Bush, K., Floyd, D. M., Georgopapadakou, N. H., Koster, W. M., Liu, W. C., Parker, W. L., Principe, P. A., Rathnum, M. L., Slusarchyk, W. A., Trejo, W. H., & Wells, J. S. (1981). Monocyclic β-lactam antibiotics produced by bacteria. *Nature* **291,** 489-491.

Thomas, J. A., Morris, D. R., & Hager, L. P. (1970). Chloroperoxidase VII. Classical peroxidatic, and halogenating forms of the enzyme. *J Biol Chem* **245,** 3129-3134.

Thompson, C. J., & Seto, H. (1995). Bialaphos. In: *Genetics and Biochemistry of Antibiotic Production*. Vining, L. C., & Stuttard, C., (eds.), Butterworth-Heinemann, Newton, MA. pp. 197-222.

Tomich, P. K. (1988). *Streptomyces* cloning: Useful recombinant DNA systems and a summation of cloned genes. *Antimicrob Agents Chemother* **32,** 1465-1471.

Tsantrizos, Y. S., Zhou, F., & Yang, X. (1995). Biosynthesis of the hypotensive metabolite oudenone by *Oudemansiella radicata*. 1. Intact incorporation of a tetraketide chain elongation intermediate. *J Org Chem* **60,** 6922-6929.

Turgay, K., & Marahiel, M. A. (1994). A general approach for identifying and cloning peptide synthetase genes. *Peptide Res* **7,** 238-241.

van Pée, K-H. (1988). Molecular cloning and high-level expression of a bromoperoxidase gene from *Streptomyces aureofaciens* Tu24. *J Bacteriol* **170,** 5890-5894.

van Pée, K-H. (1996). Biosynthesis of halogenated metabolites by bacteria. *Annu Rev Microbiol* **50,** 375-99.

van Pée, K-H., & Holzer, M. (1999). Specific enzymatic chlorination of tryptophan and tryptophan derivatives. *Adv Exp Med Biol* **467,** 603-609.

van Pée, K-H., Keller, S., Wage, T., Wynands, I., Schnerr, H., & Zehner, S. (2000). Enzymatic halogenation catalyzed via a catalytic triad and by oxidoreductases. *Biol Chem* **381,** 1-5.

van Wageningen, A., Kirkpatrick, P., Williams, D., Harris, B., Kershaw, J., Lennard, N., Jones, M., Jones, S., & Solenberg, P. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem & Biol* **3,** 155-162.

Vats, S., Stuttard, C., & Vining, L. C. (1987). Transductional analysis of chloramphenicol biosynthesis genes in *Streptomyces venezuelae*. *J Bacteriol* **169,** 3809-3813.

Vilter, H. (1984). Peroxidase from Phaeophyceae: a vanadium (V)-dependent peroxidase from *Ascophyllum nodosum*. *Phytochem* **23,** 1387-90.

Vining, L. C. (1979). Antibiotic tolerance in producer organisms. *Adv Appl Microbiol* **25,** 147-168.

Vining, L. C. (1992). Secondary metabolism, inventive evolution and biochemical diversity-a review. *Gene* **115,** 135-140.

Vining, L. C., & Chatterjee, S. (1982). Catabolic repression and the control of secondary metabolism. In: *Overproduction of Microbial products*. Krumphanzel, V., Sikyta. B., & Vanek, Z. (eds.), Czechoslovak Academy of Sciences, Prague, pp. 35-46.

Vining, L. C. & Hopwood, D. A. (1995). Polyketide-derived Antibiotics. In: *Genetics and Biochemistry of Antibiotic Production*. Vining, L. C., & Stuttard, C., (eds.), Butterworth-Heinemann, Newton, MA. pp. 323-330.

Vining, L. C., & Stuttard, C. (1995). Chloramphenicol. In: *Genetics and Biochemistry of Antibiotic Production*. Vining, L. C., & Stuttard, C. (eds.), Butterworth-Heinemann, Newton, MA. pp 505-530.

Vining, L. C. & von Dohren, H. (1995). Peptides and peptide-derived antibiotics. In: *Genetics and Biochemistry of Antibiotic Production*. Vining, L. C., & Stuttard, C., (eds.), Butterworth-Heinemann, Newton, MA. pp. 121-129.

Vining, L. C., & Westlake, D. W. S. (1964). Biosynthesis of the phenylpropanoid moiety of chloramphenicol. *Can J Microbiol* **10,** 705-716.

Wang, Z. X., Li, S. M., & Heide, L. (2000). Identification of the coumermycin A1 biosynthetic gene cluster of *Streptomyces rishiriensis* DSM 40489. *Antimicrob Agents Chemother* **44,** 3040-3048.

Weisner, W., van Pée, K.-H, & Lingens, F. (1988). Purification and characterization of a novel bacterial non-heme chloroperoxidase from *Pseudomonas pyrrocinia. J Biol Chem* **263,** 13725-13732.

White, O., Eisen, J. A., Heidelberg J. F. & 29 other authors (1999). Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* **286,** 1571-1577.

White, R. J., Maiese, W. M., & Greenstien, M. (1988). Screening for new products from microorganisms, In: *Manual of Industrial Microbiology*. Demain A. L., Solomon, N. A. (eds.). American Society for Microbiology. pp. 24-31.

Wierenga, R. K., De Jong, R. J., Kalk, K. H., Hol, W. G. J., & Drenth, J. (1979). Crystal structure of *p*-hydroxybenzoate hydroxylase. *J Mol Biol* **131,** 55-73.

Wohlleben, W., Arnold, W., Broer, I. O., Hillemann, D., Strauch, E. & Puhler, A. (1988). Nucleotide sequence of the phosphinothricin N-acetyltransferase gene from *Streptomyces viridochromogenes* Tu494 and its expression in *Nicotiana tobacum. Gene* **70,** 25-37.

Woodruff, H. B. (1996). Impact of microbial diversity on antibiotic discovery, a personal history. *J Indust Microbiol* **17,** 323-327.

Wright, F., & Bibb, M. J. (1992). Codon usage in the G+C rich *Streptomyces* genome. *Gene* **113,** 55-65.

Wuosma, A. M., & Hager, L. P. (1990). Methyl chloride transferase: A carbocation route for biosynthesis of halometabolites. *Science* **249,**160-162.

Yang, K., Han, L., & Vining, L. C. (1995). Regulation of jadomycin B production in *Streptomyces venezuelae*ISP5230: Involvement of a repressor gene, *jad*R_2. *J Bacteriol* **177,** 6111-6117.