JOURNAL OF BACTERIOLOGY, Jan. 1995, p. 252–258 Vol. 177, No. 1
0021-9193/95/$04.00+0

Nucleotide Sequence of *pvdD*, a Pyoverdine Biosynthetic Gene from *Pseudomonas aeruginosa*: PvdD Has Similarity to Peptide Synthetases

TONY R. MERRIMAN,† MARILYN E. MERRIMAN,† AND IAIN L. LAMONT*

Department of Biochemistry and Centre for Gene Research, University of Otago, Dunedin, New Zealand

Received 8 August 1994/Accepted 9 October 1994

Pseudomonas aeruginosa **secretes a fluorescent siderophore, pyoverdine, when grown under iron-deficient conditions. Pyoverdine consists of a chromophoric group bound to a partly cyclic octapeptide. As a step toward understanding the molecular events involved in pyoverdine synthesis, we have sequenced a gene, *pvdD*, required for this process. The gene encodes a 2,448-residue protein, PvdD, which has a predicted molecular mass of 273,061 Da and contains two highly similar domains of about 1,000 amino acids each. The protein is similar to peptide synthetases from a range of bacterial and fungal species, indicating that synthesis of the peptide moiety of pyoverdine proceeds by a nonribosomal mechanism. The *pvdD* gene is adjacent to a gene, *fpvA*, which encodes an outer membrane receptor protein required for uptake of ferripyoverdine.**

Pseudomonas aeruginosa is an opportunistic pathogen which infects injured, immunodeficient, or otherwise compromised patients. The bacteria secrete a siderophore, pyoverdine, which is likely to play an important role in infection by competing with transferrin for iron in order to overcome the iron-withholding defense mechanism present in mammals (3, 11, 42). Almost all isolates of *P. aeruginosa* from infected patients secrete pyoverdine (9, 22, 29), and it has been shown that pyoverdine is present in the sputa of cystic fibrosis patients infected with *P. aeruginosa* (21). A mutant of *P. aeruginosa* which is unable to synthesize pyoverdine showed reduced virulence in an animal model of infection (24).

Pyoverdine from *P. aeruginosa* PAO is a water-soluble, yellow-green fluorescent compound. It consists of a dihydroxyquinoline chromophoric group linked to an eight-residue partly cyclic peptide (D-Ser–L-Arg–D-Ser–L-N^5-OH-Orn–L-Lys–L-N^5-OH-Orn–L-Thr–L-Thr) via the N-terminal serine, with a small dicarboxylic acid attached to the chromophore. Iron complexation is thought to occur through oxygen atoms present on the dihydroxyquinoline and two hydroxamate units supplied by the L-N^5-OH-Orn residues. Several *Pseudomonas* species produce similar compounds, variously called pyoverdines or pseudobactins, all of which have the same dihydroxyquinoline group but differ in the nature of the attached peptide (reviewed in references 1 and 7), and it is likely that all of these are synthesized by similar mechanisms. Synthesis of the chromophoric group begins with condensation of D-tyrosine and L-2,4-diaminobutyric acid (7), with glutamic acid being the precursor of the amide-linked dicarboxylic acid (50). It has been suggested that biosynthesis of the peptide moiety of pyoverdine occurs by a nonribosomal mechanism (30).

Little is known about the molecular nature of enzymes involved in the biosynthesis of pyoverdine in *P. aeruginosa* or any other pseudomonad. Recently, *pvdA*, which encodes an enzyme (L-ornithine N^5-oxygenase) responsible for catalyzing the hydroxylation of ornithine, an early step in pyoverdine biosynthesis in *P. aeruginosa*, has been characterized at the molecular level (55). The *pbsC* gene, which is involved in synthesis of pseudobactin by *Pseudomonas* sp. strain M114, has also been sequenced (2). To gain further insight into the molecular mechanism of pyoverdine biosynthesis, we have determined the nucleotide sequence of the *pvdD* (pyoverdine synthetase D) gene from *P. aeruginosa.*

Nucleotide sequence of pSOT1 and identification of *pvdD*. We have previously shown that restriction fragments D and E of pSOT1 (Fig. 1) are part of a locus required for synthesis of pyoverdine (34, 35). To gain insight into the nature of the corresponding pyoverdine biosynthetic gene(s), the DNA sequence of fragments B through F of pSOT1 was determined. DNA fragments to be sequenced were subcloned into M13mp18 or mp19 (60), using standard procedures (4), with *Escherichia coli* TG1 (16) used as the host strain. Double-strande M13 DNA was prepared (25), and partial deletions of the cloned fragments were generated by using exonuclease III (Erase-a-Base system; Promega) or by using restriction enzymes which cleaved the DNA within the insert and within the M13 polylinker. Oligonucleotide primers were used to sequence regions not covered by these strategies; these were synthesized by using a model 180B DNA synthesizer (Applied Biosystems, Inc., Foster City, Calif.) or were purchased from Macromolecular Resources (Colorado State University). Single-stranded DNA of M13 subclones was prepared (5) and sequenced, using a *Taq* DNA polymerase sequencing kit (Amersham) or Sequenase Version 2.0 (United States Biochemical, Cleveland, Ohio), both with 7-deaza-dGTP. Sequence ambiguities were resolved by using the Sequenase Version 2.0 kit with dITP instead of 7-deaza-dGTP. A model 373A automated DNA sequencer (Applied Biosystems) was used to sequence across all of the cloning sites in pSOT1. Ninety-four percent of the sequence data were obtained from both strands. The remaining single-stranded data were obtained from at least two gels and were clear and unambiguous. The DNA sequence was analyzed on a MicroVaxII computer system by using various softwares: Ssedit (48), Vtutin (47), Codonuse (45), NLDNA (48), Map_Zap (41), Diagon (44), Dbextract (46), and the

* Corresponding author. Mailing address: Department of Biochemistry, University of Otago, P.O. Box 56, Dunedin, New Zealand. Phone: 64-3-479-9869. Fax: 64-3-479-7866. Electronic mail address: ilamont@sanger.otago.ac.nz.

† Present address: Wellcome Centre for Human Genetics, Nuffield Orthopaedic Centre, Oxford OX3 7BN, United Kingdom.

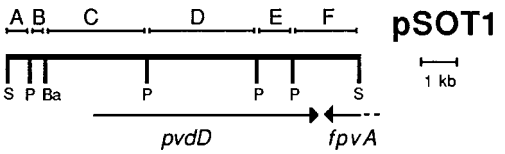


FIG. 1. (A) Clone pSOT1. pSOT1 is the largest (~9.6-kb) *Sal*I fragment of λOT1 (35) cloned into pUC9 (54) and contains DNA required for pyoverdine synthesis (34, 35). Restriction fragments are labelled by using the same nomenclature as previously (36). The locations and orientations of two genes (*pvdD* and *fpvA*) sequenced in this study are shown. Ba, *Bam*HI; P, *Pst*I; S, *Sal*I.

FoldRNA program from the version 7 UNIX of the Genetics Computer Group package (12).

The sequenced DNA contained a large open reading frame (ORF) of 7,374 bp (Fig. 2). Downstream from this ORF and oriented in the opposite direction was a partial ORF of 792 bp. Both ORFs had patterns of codon usage which are very similar to those of other genes from *P. aeruginosa* (58), indicating that they are likely to encode proteins.

Mutations mapping to restriction fragments D and E, which are entirely contained within the 7.4-kb ORF, prevent synthesis of pyoverdine by *P. aeruginosa* (34, 35). This finding shows that this gene is required for synthesis of pyoverdine, and it has been named *pvdD*, following the assignment of *pvdA*, *pvdB*, and *pvdC* to other genes involved in synthesis of pyoverdine (55). Analysis of codon usage shows that translation is likely to start at a GTG codon (Fig. 2), which nonetheless results in methionine being incorporated as the first amino acid in the protein (18). This is six bases downstream from a Shine-Dalgarno sequence which is complementary to seven of the eight bases at the 3′ end of 16S rRNA. The *pvdD* gene encodes a protein of 2,448 residues with a predicted molecular mass of 273,061 Da.

Two transcriptional start sites have been identified upstream from *pvdD* (36) (Fig. 2). Two inverted repeat sequences downstream from *pvdD* could function as transcriptional terminators for the gene. Both are capable of forming stem-loop structures; one is present 21 bp downstream from the predicted TGA termination codon (Fig. 2) and comprises an 8-bp stem and 4-bp loop ($\Delta G = -14.8$ kJ/mol, as calculated by using the FoldRNA program), and the second is 122 bp downstream from the termination codon (Fig. 2) and consists of a 12-bp stem and 3-bp loop ($\Delta G = -19.1$ kJ/mol). Both inverted repeats are followed by multiple thymidine nucleotides; this is characteristic of transcription termination sequences which function independently of the Rho termination factor (59). Transcription of *pvdD* could be terminated at either of these sites, giving a transcript of about 7.5 kb. The size of the transcript corresponding to *pvdD* was previously estimated to be 5 kb following Northern (RNA) analysis (36), but the sequence of the gene indicates that this is an underestimate, presumably due to the high degree of instability of this transcript (36).

Direct repeats within *pvdD*. Diagon analysis indicated that the *pvdD* gene contains two direct repeats of about 3 kb each. The repeats span bases 139 to 3180 and 3331 to 6363 of the sequence presented in Fig. 2 and are 91.8% identical. The deduced amino acid sequences of the repeats are 90.7% identical and 94.7% similar (Fig. 3), with one three-amino-acid insertion in the first repeat 83 residues from the N-terminal end. The C-terminal regions of the repeats (amino acids 407 to 1040) were different at only seven positions, with four of the differences representing conserved changes. In contrast, there were tracts of relatively different sequence in the N-terminal parts of the repeats.

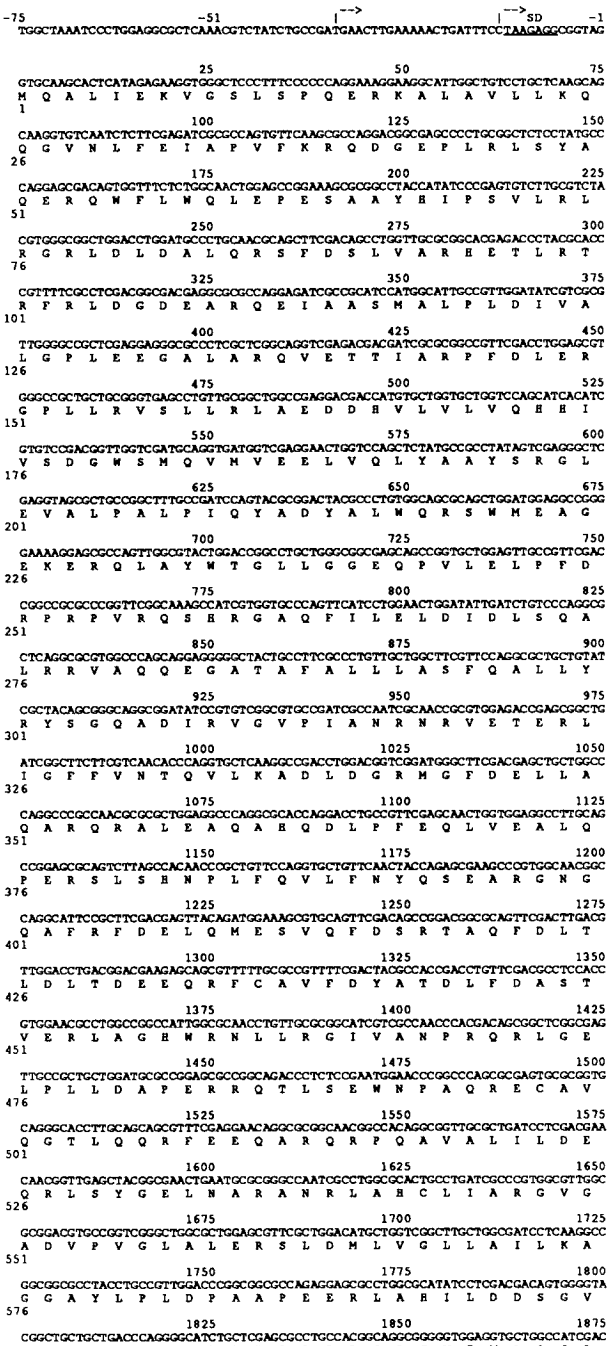
Identification of proteins similar to PvdD. The GenBank translated nucleic acid database (version 83) was searched by using FastA (32) for proteins showing sequence similarity to the first repeat of PvdD (1,013 amino acids). The protein showing greatest similarity was the product of the *pbsC* gene from *Pseudomonas* sp. strain M114, and this is discussed below. Significant similarity was also found to a family of enzymes including GrsB (53), TycA (57), SrfA (10), AcvA, PcbA, and PcbB (8, 27, 40), EntF (37), and AngR (14, 52). These enzymes are peptide synthetases, with the last two being involved in siderophore synthesis. After PbsC, the protein showing greatest similarity was GrsB, which is a well-characterized peptide synthetase from *Bacillus brevis*. This protein contains four domains, each responsible for the adenylation of a specific amino acid and its subsequent incorporation into the peptide antibiotic gramicidin (53). Each domain was similar to each of the repeats within PvdD, and an alignment of the PvdD repeats with the domains of GrsB is shown in Fig. 3.

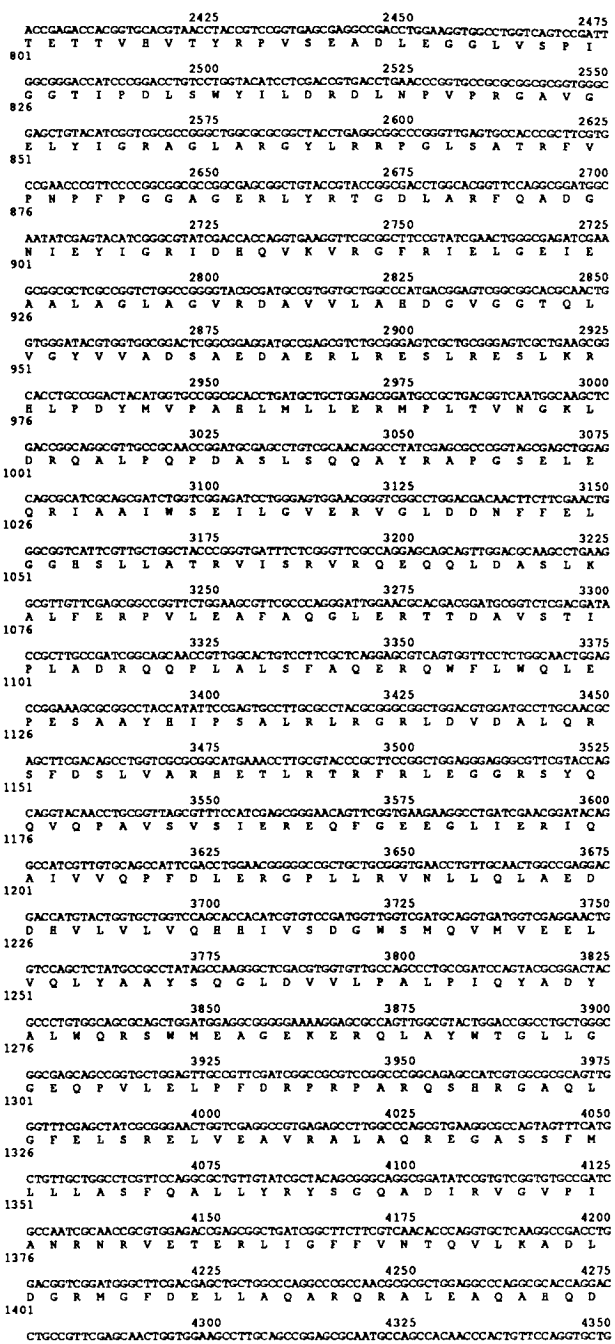
The mechanism of nonribosomal peptide synthesis by GrsB and other peptide synthetases has been studied (reviewed in references 26 and 56). Each domain of a peptide synthetase recognizes a specific amino acid and activates this substrate by adenylation. It is thought that the amino acyladenylate is then covalently attached to the enzyme by a thioester bond between the carboxyl group of the amino acid and a thiol group, which is probably supplied by a 4′-phosphopantetheine cofactor. The amino acid is then incorporated into the peptide by the enzymatic formation of a peptide bond which joins it to the amino acid bound to the adjacent domain. This is coupled with scission of the thioester bond between the cofactor and the amino acid, regenerating the enzyme (26).

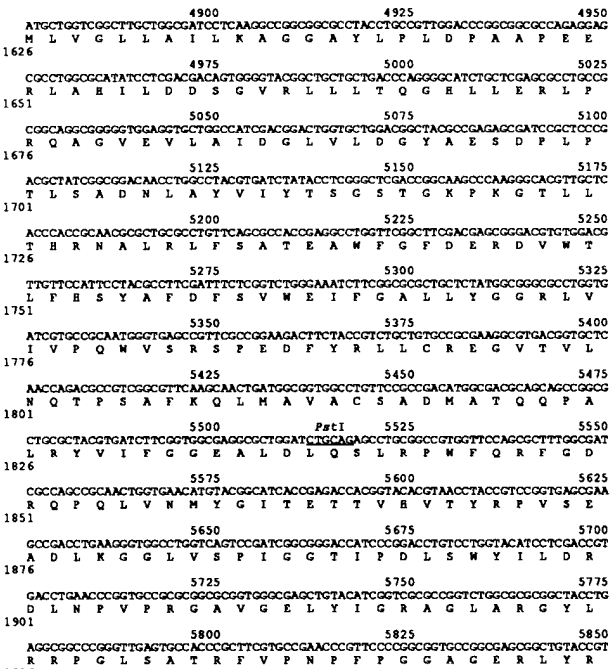
Several consensus motifs associated with these activities are present in different peptide synthetases (19), and all of these are present in PvdD (Fig. 3). From the extensive similarities found with peptide synthetases, including the presence of consensus sequence motifs (Fig. 3), it is very likely that each domain of PvdD recognizes, activates, and incorporates a specific amino acid into the peptide moiety of pyoverdine.

PvdD also contains a thioesterase active-site motif. The C-terminal segment of PvdD (336 amino acids), which is not part of either repeat sequence, was independently compared with sequences in the translated GenBank database. The only significant similarities were to EntF from *E. coli* (37) and the AcvA protein from *Penicillium chrysogenum* (40), both of which are peptide synthetases. In each case, similarity was centered around a thioesterase active-site motif (GXSXG [20], where X is any amino acid) which is present at residues 2240 to 2244 in PvdD. Similar sequence motifs are also present downstream from the domains of gramicidin S synthetase and surfactin synthetase which incorporate the C-terminal amino acids into the peptides (10, 53). Motifs of this sort are thought to be involved in termination of peptide synthesis by cleavage of the thioester bond between the 4′-phosphopantetheine cofactor of the final amino acid-activating domain and the last amino acid incorporated into the peptide (reference 10 and references therein). The presence of this motif indicates that PvdD is likely to have a thioesterase activity associated with release of the completed pyoverdine peptide, and this implies that the protein incorporates the two C-terminal threonines into the peptide. Further experiments will be required to investigate the possible thioesterase activity and to determine the amino acid substrates of PvdD.

Similarity of PvdD to PbsC. The protein found during database searching which was most similar to the first repeat of PvdD was the product of the *pbsC* gene of *Pseudomonas* sp. strain M114. This protein is predicted to contain 803 amino


-75 -51 --> --> SD -1
TGGCTAAATCCCTGGAGGCGCTCAAACGTCTATCTGCCGATGAACTTGAAAAACTGATTTCCTAAGAGGCGGTAG
25 50 75
GTGCAAGCACTCATAGAGAAGGTGGGCTCCCTTTCCCCCCAGGAAAGGAAGGCATTGGCTGTCCTGCTCAAGCAG
M Q A L I E K V G S L S P Q E R K A L A V L L K Q
1
100 125 150
CAAGGTGTCAATCTCTTCGAGATCGCGCCCAGTGTTCAAGCGCCAGGACGGCGAGCCCCTGCGGCTCTCCTATGCC
Q G V N L F E I A P V F K R Q D G E P L R L S Y A
26
175 200 225
CAGGAGCGACAGTGGTTTCTCTGGCAACTGGAGCCGGAAAGCGGCGCCTACCATATCCCGAGTGTCTTGCGTCTA
Q E R Q W F L W Q L E P E S A A Y H I P S V L R L
51
250 275 300
CGTGGGCGGCTGGACCTGGATGCCCTGCAACGCAGCTTCGACAGCCTGGTTGCGCGGCACGAGACCCTACGCACC
R G R L D L D A L Q R S F D S L V A R H E T L R T
76
325 350 375
CGTTTTCGCCTCGACGGCGACGAGGCGGCGCCAGGAGATCGCCGCATCCATGGCATTGCCGTTGGATATCGTCGCG
R F R L D G D E A R Q E I A A S M A L P L D I V A
101
400 425 450
TTGGGGCCGCTCGAGGAGGGCGCCCTCGCTCGGCAGGTCGAGACGACGATCGCGCGGCCGTTCGACCTGGAGCGT
L G P L E E G A L A R Q V E T T I A R P F D L E R
126
475 500 525
GGGCCGCTGCTGCGGGTGAGCCTGTTGCGGCTGGCCGAGGACGACCATGTGCTGGTGCTGGTCCAGCATCACATC
G P L L R V S L L R L A E D D H V L V L V Q H H I
151
550 575 600
GTGTCCGACGGTTGGTCGATGCAGGTGATGGTCGAGGAACTGGTCCAGCTCTATGCCGCCTATAGTCGAGGGCTC
V S D G W S M Q V M V E E L V Q L Y A A Y S R G L
176
625 650 675
GAGGTAGCGCTGCCGGCTTTGCCGATCCAGTACGCGGACTACGCCCTGTGGCAGCGCAGCTGGATGGAGGCCGGG
E V A L P A L P I Q Y A D Y A L W Q R S W M E A G
201
700 725 750
GAAAAGGAGCGCCAGTTGGCGTACTGGACCGGCCTGCTGGGCGGCGAGCAGCCGGTGCTGGAGTTGCCGTTCGAC
E K E R Q L A Y W T G L L G G E Q P V L E L P F D
226
775 800 825
CGGCCGCGCCCGGTTCGGCAAAGCCATCGTGGTGCCCAGTTCATCCTGGAACTGGATATTGATCTGTCCCAGGCG
R P R P V R Q S H R G A Q F I L E L D I D L S Q A
251
850 875 900
CTCAGGCGCGTGGCCCAGCAGGAGGGGGCTACTGCCTTCGCCCTGTTGCTGGCTTCGTTCCAGGCGCTGCTGTAT
L R R V A Q Q E G A T A F A L L L A S F Q A L L Y
276
925 950 975
CGCTACAGCGGGCAGGCGGATATCCGTGTCGGCGTGCCGATCGCCAATCGCAACCGCGTGGAGACCGAGCGGCTG
R Y S G Q A D I R V G V P I A N R N R V E T E R L
301
1000 1025 1050
ATCGGCTTCTTCGTCAACACCCAGGTGCTCAAGGCCGACCTGGACGGTCGGATGGGCTTCGACGAGCTGCTGGCC
I G F F V N T Q V L K A D L D G R M G F D E L L A
326
1075 1100 1125
CAGGCCCGCCAACGCGCGCTGGAGGCCCAGGCGCACCAGGACCTGCCGTTCGAGCAACTGGTGGAGGCCTTGCAG
Q A R Q R A L E A Q A H Q D L P F E Q L V E A L Q
351
1150 1175 1200
CCGGAGCGCAGTCTTAGCCACAACCCGCTGTTCCAGGTGCTGTTCAACTACCAGAGCGAAGCCCGTGGCAACGGC
P E R S L S H N P L F Q V L F N Y Q S E A R G N G
376
1225 1250 1275
CAGGCATTCCGCTTCGACGAGTTACAGATGGAAAGCGTGCAGTTCGACAGCCGGACGGCGCAGTTCGACTTGACG
Q A F R F D E L Q M E S V Q F D S R T A Q F D L T
401
1300 1325 1350
TTGGACCTGACGGACGAAGAGCAGCGTTTTTGCGCCGTTTTCGACTACGCCACCGACCTGTTCGACGCCTCCACC
L D L T D E E Q R F C A V F D Y A T D L F D A S T
426
1375 1400 1425
GTGGAACGCCTGGCCGGCCATTGGCGCAACCTGTTGCGCGGCATCGTCGCCAACCCACGACAGCGGCTCGGCGAG
V E R L A G H W R N L L R G I V A N P R Q R L G E
451
1450 1475 1500
TTGCCGCTGCTGGATGCGCCGGAGCGCCGGCAGACCCTCTCCGAATGGAACCCGCCCCAGCGCGAGTGCGCGGTG
L P L L D A P E R R Q T L S E W N P A Q R E C A V
476
1525 1550 1575
CAGGGCACCTTGCAGCAGCGTTTCGAGGAACAGGCGCGGCAACGGCCACAGGCGGTTGCGCTGATCCTCGACGAA
Q G T L Q Q R F E E Q A R Q R P Q A V A L I L D E
501
1600 1625 1650
CAACGGTTGAGCTACGGCGAACTGAATGCGCGGGCCAATCGCCTGGCGCACTGCCTGATCGCCCGTGGCGTTGGC
Q R L S Y G E L N A R A N R L A H C L I A R G V G
526
1675 1700 1725
GCGGACGTGCCGGTCGGGCTGGCGCTGGAGCGTTCGCTGGACATGCTGTCGGCTTGCTGGCGATCCTCAAGGCC
A D V P V G L A L E R S L D M L V G L L A I L K A
551
1750 1775 1800
GGCGGCGCCTACCTGCCGTTGGACCCGGCGGCGCCAGAGGAGCGCCTGGCGCATATCCTCGACGACAGTGGGGTA
G G A Y L P L D P A A P E E R L A H I L D D S G V
576
1825 1850 1875
CGGCTGCTGCTGACCCAGGGGCATCTGCTCGAGCGCCTGCCACGGCAGGCGGGGGTGGAGGTGCTGGCCATCGAC
R L L L T Q G H L L E R L P R Q A G V E V L A I D
601
1900 1925 1950
GGACTGGTGCTGGACGGCTACGCCGAGAGCGATCCGCTCCCGACGCTATCGGCGGACAACCTGGCCTACGTGATC
G L V L D G Y A E S D P L P T L S A D N L A Y V I
626
1975 2000 2025
TATACCTCGGGCTCGACCGGCAAGCCCAAGGGCACATTGCTCACCCACCGCAACGCGCTGCGCCTGTTCAGCGCC
Y T S G S T G K P K G T L L T H R N A L R L F S A
651
2050 2075 2100
ACCGAGGCCTGGTTCGGCTTCGACGAGCGGGACGTGTGGACATTGTTCCATTCCTACGCCTTCGATTTCTCGGTC
T E A W F G F D E R D V W T L F H S Y A F D F S V
676
2125 2150 2175
TGGGAAATCTTCGGCGCGCTGCTCTATGGCGGGTGCCTGGTGATTGTGCCGCAATGGGTGAGCCGTTCGCCGGAA
W E I F G A L L Y G G C L V I V P Q W V S R S P E
701
2200 2225 2250
GACTTCTACCGTCTGCTGTGCCGCGAAGGCGTGACGGTGCTCAACCAGACGCCGTCGGCGTTCAAGCAACTGATG
D F Y R L L C R E G V T V L N Q T P S A F K Q L M
726
2275 2300 2325
GCGGTGGCCTGTTCCGCCGACATGGCGACGCAGCAGCCGGCGCTGCGCTACGTGATCTTCGGTGGCGAGGCGCTG
A V A C S A D M A T Q Q P A L R Y V I F G G E A L
751
PstI 2350 2375 2400
GATCTGCAGAGCCTGCGGCCGTGGTTCCAGCGCTTCGGCGATCGCCAGCCGCAACTGGTGAACATGTACGGCATC
D L Q S L R P W F Q R F G D R Q P Q L V N M Y G I
776


2425 2450 2475
ACCGAGACCACGGTGCACGTAACCTACCGTCCGGTGAGCGAGGCCGACCTGGAAGGTGGCCTGGTCAGTCCGATT
T E T T V H V T Y R P V S E A D L E G G L V S P I
801
2500 2525 2550
GGCGGGACCATCCCGGACCTGTCCTGGTACATCCTCGACGGTGACCTGAACCCGGTGCCGCGCCGCGGTGGGC
G G T I P D L S W Y I L D R D L N P V P R G A V G
826
2575 2600 2625
GAGCTGTACATCGGTCGCGCCGGGCTGGCGCGCGGCTACCTGAGGCGGCCCGGGTTGAGTGCCACCCGCTTCGTG
E L Y I G R A G L A R G Y L R R P G L S A T R F V
851
2650 2675 2700
CCGAACCCGTTCCCCGGCCGCGCCGGCGAGCGGCTGTACCGTACCGGCGACCTGGCACGGTTCCAGGCGGATGGC
P N P F P G G A G E R L Y R T G D L A R F Q A D G
876
2725 2750 2725
AATATCGAGTACATCGGGCGTATCGACCACCAGGTGAAGGTTCGCGGCTTCCGTATCGAACTGGGCGAGATCGAA
N I E Y I G R I D H Q V K V R G F R I E L G E I E
901
2800 2825 2850
GCGGCCCTCGCCGGTCTGCCGGGGTACGCGATGCCGTGGTCGCTGGCCCATGACGGAGTCGGCGGCACGCAACTG
A A L A G L A G V R D A V V L A H D G V G G T Q L
926
2875 2900 2925
GTGGGATACGTGGTGGCGGACTCGGCGGAGGATGCCGAGCGTCTGCGGGAGTCGCTGCGGGAGTCGCTGAAGCGG
V G Y V V A D S A E D A E R L R E S L R E S L K R
951
2950 2975 3000
CACCTGCCGGACTACATGGTGCCGGCGCACCTGATGCTGCTGGAGCGGATGCCGCTGACGGTCAATGGCAAGCTC
H L P D Y M V P A H L M L L E R M P L T V N G K L
976
3025 3050 3075
GACCGGCAGGCGTTGCCGCAACCGGATGCGAGCCTGTCGCAACAGGCCTATCGAGCGCCGGTAGCGAGCTGGAG
D R Q A L P Q P D A S L S Q Q A Y R A P G S E L E
1001
3100 3125 3150
CAGCGCATCGCAGCGATCTGGTCGGAGATCCTGGGAGTGGAACGGGTCGGCCTGGACGACAACTTCTTCGAACTG
Q R I A A I W S E I L G V E R V G L D D N F F E L
1026
3175 3200 3225
GGCGGTCATTCGTTGCTGGCTACCCGGGTGATTTCTCGGGTTCGCCAGGAGCAGCAGTTGGACGCAAGCCTGAAG
G G H S L L A T R V I S R V R Q E Q Q L D A S L K
1051
3250 3275 3300
GCGTTGTTCGAGCGGCCGGTTCTGGAAGCGTTCGCCCAGGGATTGGAACGCACGACGGATGCGGTCTCGACGATA
A L F E R P V L E A F A Q G L E R T T D A V S T I
1076
3325 3350 3375
CCGCTTGCCGATCGGCAGCAACCGTTGGCACTGTCCTTCGCTCAGGAGCGTCAGTGGTTCCTCTGGCAACTGGAG
P L A D R Q Q P L A L S F A Q E R Q W F L W Q L E
1101
3400 3425 3450
CCGGAAAGCGCGGCCTACCATATTCCGAGTGCCTTGCGCCTACGCGGGCGGCTGGACGTGGATGCCTTGCAACGC
P E S A A Y H I P S A L R L R G R L D V D A L Q R
1126
3475 3500 3525
AGCTTCGACAGCCTGGTCGCGGGCATGAAACCTTGCGTACCCGCTTCCGGCTGGAGGAGGGCGTTCGTACCAG
S F D S L V A R H E T L R T R F R L E G G R S Y Q
1151
3550 3575 3600
CAGGTACAACCTGCGGTTAGCGTTTCCATCGAGCGGGAACAGTTCGGTGAAGAAGGCCTGATCGAACGGATACAG
Q V Q P A V S V S I E R E Q F G E E G L I E R I Q
1176
3625 3650 3675
GCCATCGTTGTGCAGCCATTCGACCTGGAACGGGGCCGCTGCTGCGGGTGAACCTGTTGCAACTGGCCCAGGAC
A I V V Q P F D L E R G P L L R V N L L Q L A E D
1201
3700 3725 3750
GACCATGTACTGGTGCTGGTCCAGCACCACATCGTGTCCGATGGTTGGTCGATGCAGGTGATGGTCGAGGAACTG
D H V L V L V Q H H I V S D G W S M Q V M V E E L
1226
3775 3800 3825
GTCCAGCTCTATGCCGCCTATAGCCAAGGGCTCGACGTGGTGTTGCCAGCCCTGCCGATCCAGTACGCGGACTAC
V Q L Y A A Y S Q G L D V V L P A L P I Q Y A D Y
1251
3850 3875 3900
GCCCTGTGGCAGCGCAGCTGGATGGAGGCGGGGAAAAGGAGCGCCAGTTGGCGTACTGGACCGGCCTGCTGGGC
A L W Q R S W M E A G E K E R Q L A Y W T G L L G
1276
3925 3950 3975
GGCGAGCAGCCGGTGCTGGAGTTGCCGTTCGATCGGCCGCGTCCGGCCCGGCAGAGCCATCGTGCCGCCCAGTTG
G E Q P V L E L P F D R P R P A R Q S H R G A Q L
1301
4000 4025 4050
GGTTTCGAGCTATCGCGGGAACTGGTCGAGGCCGTGAGAGCCTTGGCCCAGCGTGAAGGCGCCAGTAGTTTCATG
G F E L S R E L V E A V R A L A Q R E G A S S F M
1326
4075 4100 4125
CTGTTGCTGGCCTCGTTCCAGGCGCTGTTGTATCGCTACAGCGGGCAGGCGGATATCCGTGTCGGTGTGCCGATC
L L L A S F Q A L L Y R Y S G Q A D I R V G V P I
1351
4150 4175 4200
GCCAATCGCAACCGCGTGGAGACCGAGCGGCTGATCGGCTTCTTCGTCAACACCCAGGTGCTCAAGGCCGACCTG
A N R N R V E T E R L I G F F V N T Q V L K A D L
1376
4225 4250 4275
GACGGTCGGATGGGCTTCGACGAGCTGCTGGCCCAGGCCCGCCAACGCGCGCTGGAGGCCCAGGCGCACCAGGAC
D G R M G F D E L L A Q A R Q R A L E A Q A H Q D
1401
4300 4325 4350
CTGCCGTTCGAGCAACTGGTGGAAGCCTTGCAGCCGGAGCGCAATGCCAGCCACAACCCACTGTTCCAGGTGCTG
L P F E Q L V E A L Q P E R N A S H N P L F Q V L
1426
4375 4400 4425
TTCAACCATCAGAGCGAGATACGCTCGGTGACGCCCGAGGTTCAGTTGGAGGACCTGCGTCTGGAAGGCCTGGCC
F N H Q S E I R S V T P E V Q L E D L R L E G L A
1451
4450 4475 4500
TGGGACGGCCAGACTGCGCAGTTCGACCTGACGCTGGATATTCAGGAAGACGAAAACGGCATCTGGGCCTCCTTC
W D G Q T A Q F D L T L D I Q E D E N G I W A S F
1476
4525 4550 4575
GACTATGCCACCGATCTGTTCGACGCCTCCACCGTGGAAGCCCTGGCCGGCCATTGGCGCAACCTGTTGCGGGGC
D Y A T D L F D A S T V E R L A G H W R N L L R G
1501
4600 4625 4650
ATCGTCGCCAACCCACGACAGCGGCTCGGCGAGTTGCCGCTGCTGGATGCGCCGGAGCGCCGGCAGACCCTCTCC
I V A N P R Q R L G E L P L L D A P E R R Q T L S
1526
4675 4700 4725
GAATGGAACCCGGCCCAGCGCGAGTGCGCGGTGCAGGGCACCTTGCAGCAGCGTTTCGAGGAGCAGGCGCGGCAA
E W N P A Q R E C A V Q G T L Q Q R F E E Q A R Q
1551
4750 4775 4800
CGGCCACAGGCGGTTGCGCTGATCCTCGACGAACAACGGTTGAGCTACGGCGAACTGAATGCGCGGGCCAATCGC
R P Q A V A L I L D E Q R L S Y G E L N A R A N R
1576
4825 4850 4875
CTGGCGCACTGCCTGATCGCTCGCGGCGTTGGCGCGGACGTGCCGGTCGGGCTGGCGCTGGAGCGTTCGCTGGAC
L A H C L I A R G V G A D V P V G L A L E R S L D
1601



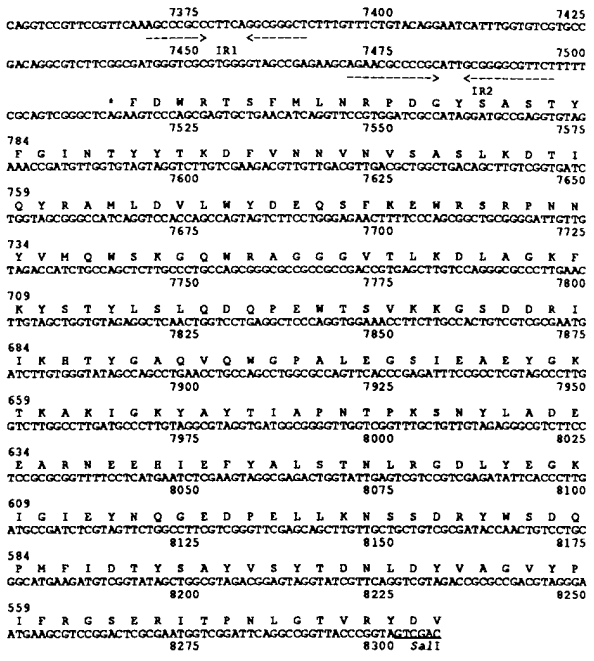


FIG. 2. Nucleotide sequences of the *pvdD* gene and part of the *fpvA* gene. The predicted amino acid sequences of the proteins are also shown. Nucleotide numbering is relative to the likely translational start site of *pvdD*. Transcription start sites for *pvdD* ($|^{-->}$), a Shine-Dalgarno sequence (SD), and inverted repeats downstream from *pvdD* (IR1 and IR2) are shown. *Bam*HI, *Pst*I, and *Sal*I restriction sites are indicated. Part of the sequence (nucleotides −75 to 390) has been published elsewhere (36) and is included here for completeness.

acids and to have a molecular mass of 89 kDa (2). However, when the translated *pbsC* sequence was aligned with that of PvdD by using the GAP program (12), similarity extended over the first 988 amino acids of PvdD, with the start of PbsC predicted by Adams et al. (2) corresponding to amino acid 185 of PvdD. It seems likely that the GTG codon in the *pbsC* sequence which corresponds to the proposed GTG start site in *pvdD* is the start of the *pbsC* gene.

Mutations of *pbsC* in *Pseudomonas* sp. strain M114 prevent the synthesis of pseudobactin, and it has been suggested that PbsC may be a member of the peptide synthetase family which is involved in the biosynthesis of M114 pseudobactin (2). PbsC contains most of the sequence motifs present in peptide synthetases, but motifs VI and X (Fig. 3) are missing. The exact relationship of PbsC with members of the peptide synthetase family remains to be determined, but its striking similarity to PvdD indicates that the two proteins are likely to have some functions in common in siderophore synthesis. DNA hybridization experiments have shown that genes similar to *pvdD* are also present in siderophore synthesis DNA from other fluorescent pseudomonads (34, 35), indicating that proteins similar to PvdD are also present in these species.

***pvdD* is adjacent to a gene encoding a ferripyoverdine uptake protein.** DNA sequencing also revealed a partial ORF of 792 bp downstream from *pvdD* and orientated in the opposite direction (Fig. 2). The 264-residue predicted protein product of the partial ORF downstream from *pvdD* was 99.2% identical to amino acids 552 to 814 of FpvA, the outer membrane receptor for ferripyoverdine (33). There was a two-amino-acid insertion and one amino acid difference when the translated sequence of this ORF was compared with that of Poole et al. (33). Further sequencing and genetic analysis has shown that this ORF does indeed encode the ferripyoverdine outer membrane receptor (51). The differences between our *fpvA* se-

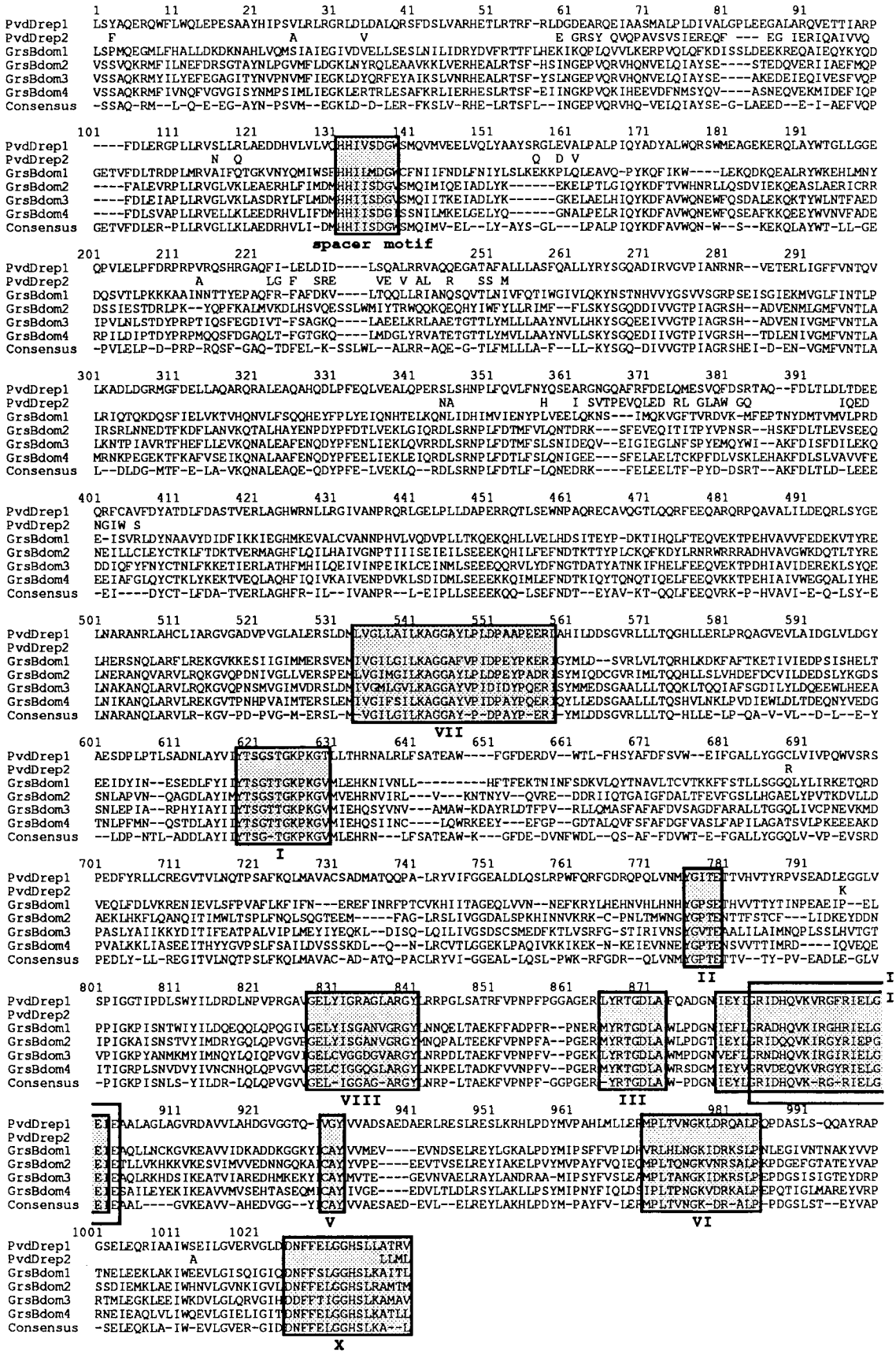

spacer motif
I
II
III
IV
V
VI
VII
VIII
IX
X

FIG. 3. Alignment of the predicted peptide synthetase domains of PvdD with those of GrsB. The two repeats of PvdD (PvdDrep1 and PvdDrep2) were aligned with the four domains of gramicidin S synthetase B (GrsBdom1 to GrsBdom4) (53), using Homed (49) and CLUSTAL V (23). The two repeats of PvdD were identical except at sites where a residue is shown for PvdDrep2. Boxes denote consensus sequence motifs which have been identified for peptide synthetase-type enzymes; the consensus sequences are listed in reference 19. Motifs I to VI are conserved among the superfamily of adenylate-forming enzymes, including nonpeptide synthetases. Motifs VII to X are conserved among a subfamily of the adenylate-forming superfamily which includes all known peptide synthetases and also EntF and AngR. Motifs I and III have been associated with ATP-binding and hydrolysis activities, respectively (15, 17, 38). Motif X contains a conserved serine (position 1034 in the alignment) to which a 4′-phosphopantetheine cofactor is very probably attached (13, 39). A so-called spacer motif present at the N-terminal end of peptide synthetase domains involved in elongation, but not initiation, of peptide synthesis (15) is also indicated. It has been suggested that amino acids between 627 and 829 in the alignment are involved in substrate recognition by peptide synthetase domains (10).

quence and that of Poole et al. (33) probably reflect minor differences between the *P. aeruginosa* strains used in the respective studies. The presence of *fpvA* downstream from *pvdD* indicates that ferripyoverdine synthesis and uptake genes are clustered, as is the case in other fluorescent pseudomonads (6, 28, 31). In addition, this allows the *fpvA* gene to be located at 47 min on the genetic map of *P. aeruginosa*, as we have previously shown that clone pSOT1, which includes this DNA, is derived from that part of the chromosome (35).

Nucleotide sequence accession number. The sequence shown in Fig. 2 has been assigned accession number U07359 in the GenBank and EMBL libraries.

We are very grateful to Irene Rombel for providing some of the plasmid constructs used in this study and to Anthony Smith for critical reading of the manuscript.

T.R.M. was supported by a postgraduate studentship from the Health Research Council of New Zealand. M.E.M. was supported by the Health Research Council of New Zealand.

REFERENCES

1. **Abdallah, M. A.** 1991. Pyoverdins and pseudobactins, p. 139–154. *In* G. Winkelmann (ed.), CRC handbook of microbial iron chelates. CRC Press, Fla.
2. **Adams, C., D. N. Dowling, D. J. O'Sullivan, and F. O'Gara.** 1994. Isolation of a gene (*pbsC*) required for siderophore biosynthesis in fluorescent *Pseudomonas* sp. strain M114. Mol. Gen. Genet. **243:**515–524.
3. **Ankenbauer, R., S. Sriyosachati, and C. D. Cox.** 1985. Effects of siderophores on the growth of *Pseudomonas aeruginosa* in human serum and transferrin. Infect. Immun. **49:**132–140.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. Greene Publishing Associates and Wiley Interscience, New York.
5. **Bankier, A. T., K. M. Weston, and B. G. Barrell.** 1987. Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. Methods Enzymol. **155:**51–93.
6. **Bitter, W., J. D. Marugg, L. A. de Weger, J. Tommassen, and P. J. Weisbeek.** 1991. The ferric-pseudobactin receptor PupA of *Pseudomonas putida* WCS358: homology to TonB-dependent *Escherichia coli* receptors and specificity of the protein. Mol. Microbiol. **5:**647–655.
7. **Budzikiewicz, H.** 1993. Secondary metabolites from fluorescent pseudomonads. FEMS Microbiol. Rev. **104:**209–228.
8. **Coque, J. J. R., J. F. Martín, J. G. Calzada, and P. Liras.** 1991. The cephamycin biosynthetic genes *pcbAB*, encoding a large multidomain peptide synthetase, and *pcbC* of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *Acremonium chrysogenum* and *Penicillium chrysogenum*. Mol. Microbiol. **5:**1125–1133.
9. **Cornelis, P., D. Hohnadel, and J.-M. Meyer.** 1989. Evidence for different pyoverdine-mediated iron uptake systems among *Pseudomonas aeruginosa* strains. Infect. Immun. **57:**3491–3497.
10. **Cosmina, P., F. Rodrigeuz, F. de Ferra, G. Grandi, M. Perego, G. Venema, and D. van Sinderen.** 1993. Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. Mol. Microbiol. **8:**821–831.
11. **Cox, C. D., and P. Adams.** 1985. Siderophore activity of pyoverdin for *Pseudomonas aeruginosa*. Infect. Immun. **48:**130–138.
12. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
13. **D'Souza, C., M. N. Makano, N. Corbell, and P. Zuber.** 1993. Amino-acylation site mutations in amino acid-activating domains of surfactin synthetase: effects on surfactin production and competence development in *Bacillus subtilis*. J. Bacteriol. **175:**3502–3510.
14. **Farrell, D. H., P. Mikesell, L. A. Actis, and J. H. Crosa.** 1990. A regulatory gene, *angR*, of the iron uptake system of *Vibrio anguillarum*: similarity with phage P22 *cro* and regulation by iron. Gene **86:**45–51.
15. **Fuma, S., Y. Fujishima, N. Corbell, C. D'Souza, M. M. Nakano, P. Zuber, and K. Yamane.** 1993. Nucleotide sequence of 5′ portion of *srfA* that contains the region required for competence establishment in *Bacillus subtilis*. Nucleic Acids Res. **21:**93–97.
16. **Gibson, T. J.** 1984. Ph.D. thesis. University of Cambridge, Cambridge, England.
17. **Gocht, M., and M. A. Marahiel.** 1994. Analysis of core sequences in the D-Phe activating domain of the multifunctional peptide synthesis TycA by site-directed mutagenesis. J. Bacteriol. **176:**2654–2662.
18. **Gold, L., and G. Stormo.** 1987. Translational initiation, p. 1302–1307. *In* J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
19. **Guilvout, I., O. Mercereau-Puijalon, S. Bonnefoy, A. P. Pugsley, and E. Carniel.** 1993. High-molecular-weight protein 2 of *Yersinia enterocolitica* is homologous to AngR of *Vibrio anguillarum* and belongs to a family of proteins involved in nonribosomal peptide synthesis. J. Bacteriol. **175:**5488–5504.
20. **Gutiérrez, S., B. Díez, E. Montenegro, and J. F. Martin.** 1991. Characterization of the *Cephalosporium acremonium pcbAB* gene encoding α-aminoadipyl-cysteinyl-valine synthetase, a large multidomain peptide synthetase: linkage to the *pcbC* gene as a cluster of early cephalosporin biosynthetic genes and evidence of multiple functional domains. J. Bacteriol. **173:**2354–2365.
21. **Haas, B., J. Kraut, J. Marks, S. C. Zanker, and D. Castignetti.** 1991. Siderophore presence in sputa of cystic fibrosis patients. Infect. Immun. **59:**3997–4000.
22. **Haas, B., E. Murphy, and D. Castignetti.** 1991. Siderophore synthesis by mucoid *Pseudomonas aeruginosa* strains isolated from cystic fibrosis patients. Can. J. Microbiol. **37:**654–657.
23. **Higgins, D. G., A. J. Bleasby, and R. Fuchs.** 1992. CLUSTAL V: improved software for multiple sequence alignment. Computer Appl. Biosci. **8:**189–191.
24. **Holder, I. A., Z. Akchich, and J.-M. Meyer.** 1994. Pyoverdin iron binding and virulence of *Pseudomonas aeruginosa*, abstr. B-49, p. 37. *In* Abstracts of the 94th General Meeting of the American Society for Microbiology 1994. American Society for Microbiology, Washington, D.C.
25. **Ish-Horowicz, D., and J. F. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9:**2989–2998.
26. **Kleinkauf, H., and H. von Döhren.** 1990. Nonribosomal biosynthesis of peptide antibiotics. Eur. J. Biochem. **192:**1–15.
27. **MacCabe, A. P., H. van Liempt, H. Palissa, S. E. Unkles, M. B. R. Riach, E. Pfeifer, H. von Döhren, and J. R. Kinghorn.** 1991. δ-(L-α-Aminoadipyl)-L-cysteinyl-D-valine synthetase from *Aspergillus nidulans*. J. Biol. Chem. **266:**12646–12654.
28. **Magazin, M. D., J. C. Moores, and J. Leong.** 1986. Cloning of the gene coding for the outer membrane receptor protein for ferric pseudobactin, a siderophore from a plant growth-promoting *Pseudomonas* strain. J. Biol. Chem. **261:**795–799.
29. **McMorran, B., and I. Lamont.** 1992. Production of pyoverdine by isolates of *Pseudomonas aeruginosa* from patients with cystic fibrosis. Proc. Univ. Otago Med. School **70:**11–12.
30. **Moores, J. C., M. Magazin, G. S. Ditta, and J. Leong.** 1984. Cloning of genes involved in the biosynthesis of pseudobactin, a high-affinity iron transport agent of a plant growth-promoting *Pseudomonas* strain. J. Bacteriol. **157:**53–58.
31. **Morris, J., D. J. O'Sullivan, M. Koster, J. Leong, P. J. Weisbeek, and F. O'Gara.** 1992. Characterization of fluorescent siderophore-mediated iron uptake in *Pseudomonas* sp. strain M114: evidence for the existence of an additional ferric siderophore receptor. Appl. Environ. Microbiol. **58:**630–635.
32. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
33. **Poole, K., S. Neshat, K. Krebes, and D. E. Heinrichs.** 1993. Cloning and nucleotide analysis of the ferripyoverdine receptor gene *fpvA* of *Pseudomonas aeruginosa*. J. Bacteriol. **175:**4597–4604.
34. **Rombel, I. T.** 1992. Ph.D. thesis. University of Otago, Dunedin, New Zealand.
35. **Rombel, I. T., and I. L. Lamont.** 1992. DNA homology between siderophore

genes from fluorescent pseudomonads. J. Gen. Microbiol. **138:**181–187.
36. **Rombel, I. T., B. J. McMorran, and I. L. Lamont.** Identification of a DNA sequence motif required for expression of iron-regulated genes in pseudomonads. Mol. Gen. Genet., in press.
37. **Rusnak, F., M. Sakaitani, D. Drueckhammer, J. Reichert, and C. T. Walsh.** 1991. Biosynthesis of the *Escherichia coli* siderophore enterobactin: sequence of the *entF* gene, expression and purification of EntF, and analysis of covalent phosphopantetheine. Biochemistry **30:**2916–2927.
38. **Saraste, M., P. R. Sibbald, and A. Wittinghofer.** 1990. The P-loop—a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. **15:**430–434.
39. **Schlumbohm, W., T. Stein, C. Ullrich, J. Vater, M. Krause, M. A. Marahiel, V. Kruft, and B. Wittmann-Liebold.** 1991. An active serine is involved in covalent substrate amino acid binding at each reaction center of gramicidin S synthetase. J. Biol. Chem. **266:**23135–23141.
40. **Smith, D. J., A. J. Earl, and G. Turner.** 1990. The multifunctional peptide synthetase performing the first step of penicillin synthesis in *Penicillium chrysogenum* is a 421 073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases. EMBO J. **9:**2743–2750.
41. **Smith, R.** 1988. A finite state machine algorithm for finding restriction sites and other pattern matching applications. Computer Appl. Biosci. **4:**459–465.
42. **Sriyosachati, S., and C. D. Cox.** 1986. Siderophore-mediated iron acquisition from transferrin by *Pseudomonas aeruginosa*. Infect. Immun. **52:**885–891.
43. **Staab, J. F., M. F. Elkins, and C. F. Earhart.** 1989. Nucleotide sequence of the *Escherichia coli entE* gene. FEMS Microbiol. Lett. **59:**15–20.
44. **Staden, R.** 1982. An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucleic Acids Res. **10:**2951–2961.
45. **Staden, R., and A. D. McLachlan.** 1982. Codon preference and its use in identifying protein coding regions in long DNA sequences. Nucleic Acids Res. **10:**141–156.
46. **Stockwell, P. A. (University of Otago, Dunedin, New Zealand).** 1988. Personal communication.
47. **Stockwell, P. A.** 1985. VTUTIN: a full screen gel management editor. Computer Appl. Biosci. **1:**253–259.
48. **Stockwell, P. A.** 1987. DNA sequence analysis software, p. 19–45. *In* M. J. Bishop and C. J. Rawlings (ed.), Nucleic acid and protein sequence analysis. IRL Press, Oxford.
49. **Stockwell, P. A.** 1988. HOMED: a homologous sequence editor. Trends Biochem. Sci. **13:**322–324.
50. **Taraz, K., R. Tappe, H. Schröder, U. Hohlneicher, I. Gwose, H. Budzikiewicz, G. Mohn, and J. F. Lefèvre.** 1991. Ferribactins—the biogenetic precursors of pyoverdins [1]. Z. Naturforsch. **46c:**527–533.
51. **Thorn, R. L., M. E. Merriman, and I. L. Lamont.** Unpublished data.
52. **Tolmasky, M. E., L. A. Actis, and J. H. Crosa.** 1993. A single amino acid change in AngR, a protein encoded by pJM1-like virulence plasmids, results in hyperproduction of anguibactin. Infect. Immun. **61:**3228–3233.
53. **Turgay, K., M. Krause, and M. A. Marahiel.** 1992. Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes. Mol. Microbiol. **6:**529–546.
54. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
55. **Visca, P., A. Ciervo, and N. Orsi.** 1994. Cloning and nucleotide sequence of the *pvdA* gene encoding the pyoverdin biosynthetic enzyme L-ornithine N^5-oxygenase in *Pseudomonas aeruginosa*. J. Bacteriol. **176:**1128–1140.
56. **Von Döhren, H.** 1993. The organization of multifunctional peptide and depsipeptide synthetases. Biochem. Soc. Trans. **21:**214–217.
57. **Weckermann, R., R. Fürbaβ, and M. A. Marahiel.** 1988. Complete nucleotide sequence of the *tycA* gene encoding the tyrocidine synthetase I from *Bacillus brevis*. Nucleic Acids Res. **16:**11841.
58. **West, S. E. H., and B. H. Iglewski.** 1988. Codon usage in *Pseudomonas aeruginosa*. Nucleic Acids Res. **16:**9323–9335.
59. **Yager, T. D., and P. H. von Hippel.** 1987. Transcript elongation and termination in *Escherichia coli*, p. 1241–1275. *In* J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
60. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.