The EMBO Journal vol.9 no.9 pp.2743-2750, 1990

The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *Penicillium chrysogenum* is a 421 073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases

David J.Smith[1], Alison J.Earl[2] and Geoffrey Turner[3]

Department of Microbiology, University of Bristol, University Walk, Bristol BS8 1TD, UK and [2]Fermentation Development, SmithKline Beecham, Clarendon Road, Worthing, W.Sussex BN14 8QH, UK

[1]Present address: The Glaxo Institute for Molecular Biology SA, Route des Acacias 46, CH-1211, Geneva 24, Switzerland
[3]Present address: Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, UK

Communicated by J.R.S.Fincham

The nucleotide sequence of the *Penicillium chrysogenum* Oli13 *acvA* gene encoding δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase, which performs the first step in penicillin biosynthesis, has been determined. The *acvA* gene contains an open reading frame of 11 238 bp encoding a protein of 3746 amino acids with a predicted mol. wt of 421 073 dalton. Three domains within the protein of ~570 amino acids have between 38% and 43% identity with each other and share similarity with two antibiotic peptide synthetases from *Bacillus brevis* as well as two other enzymes capable of performing ATP-pyrophosphate exchange reactions. The *acvA* gene is located close to the *pcbC* gene encoding isopenicillin N synthetase, the enzyme for the second step of β-lactam biosynthesis, and is transcribed in the opposite orientation to it. The intergenic region of 1107 bp from which the *acvA* and *pcbC* genes are divergently transcribed has also been sequenced.

Key words: acvA gene/ACV synthetase/β-lactam antibiotics/*Penicillium chrysogenum*/peptide synthetase

Introduction

The formation of the tripeptide δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV) from the L-isomers of the constitutent amino acids is the first step in the pathway of β-lactam antibiotic biosynthesis in all producers so far studied. ACV is then cyclized to give isopenicillin N (IPN) which is the precursor for the divergent pathways of penicillin biosynthesis in filamentous fungi or cephalosporin and cephamycin biosynthesis in filamentous fungi, actinomycetes and some eubacteria (Nuesch *et al.*, 1987).

Early studies of ACV synthetase (ACVS) using cell-free extracts of *Penicillium chrysogenum* (Lara *et al.*, 1982) and *Cephalosporium acremonium* (Banko *et al.*, 1986) suggested that two separate enzymes may be involved in the formation of ACV. However, based on the observation that less ACV is formed from the dipeptide L-(α-aminoadipyl)-L-cysteine and L-valine than the free amino acids, Banko *et al.* (1987) proposed that a single multifunctional enzyme may be responsible for ACV biosynthesis. A similar observation was made with ACVS from *Streptomyces clavuligerus* and an analogous multifunctional enzyme was proposed (Jensen *et al.*, 1988).

Recently the ACVS of *Aspergillus nidulans* has been purified and shown to be a single multifunctional peptide synthetase (van Liempt *et al.*, 1989). The purified *A.nidulans* ACVS had an estimated molecular weight of 230 kd and formed ACV from the constituent L-amino acids. The enzyme could be amino-acylated with L-[^{14}C]valine and catalysed ATP-pyrophosphate exchange which was dependent on the presence of the constituent amino acids of ACV. Based on these observations the authors proposed a multi-enzyme thiotemplate mechanism in which each of the constituent amino acids of ACV are activated as aminoacyl adenylates with peptide bonds formed through the participation of amino acid thiolester intermediates. The multi-enzyme thiotemplate mechanism is also found in bacterial peptide synthetases, such as those involved in the synthesis of the *Bacillus brevis* peptide antibiotics tyrocidine and gramicidin S, but is different from that of glutathione biosynthesis (reviewed by Kleinkauf and von Dohren, 1987).

The genes encoding the enzymes involved in β-lactam antibiotic biosynthesis have been the subject of much recent investigation (reviewed by Ingolia and Queener, 1989) and genes for several of the biosynthetic steps from a number of producing species have been cloned (Samson *et al.*, 1985; Carr *et al.*, 1986; Ramon *et al.*, 1987; Samson *et al.*, 1987; Barredo *et al.*, 1989a; MacCabe *et al.*, 1990). Smith *et al.* (1990a) showed that all the genes essential for the biosynthesis of penicillin from amino acid precursors are closely linked in *P.chrysogenum*. Subsequently, conserved gene clusters for β-lactam biosynthesis which contained the gene encoding ACVS were isolated from two filamentous fungi, *A.nidulans* and *P.chrysogenum*, the Gram-positive prokaryote *S.clavuligerus* and the Gram-negative prokaryote *Flavobacterium* sp. SC 12,154 (Smith *et al.*, 1990b).

We have determined the nucleotide sequence of the *P.chrysogenum* Oli13 *acvA* gene, which encodes ACVS, and predict that it encodes a massive protein containing three large, homologous domains which are similar to antibiotic peptide synthetases from a bacillus.

Results

The penicillin biosynthetic gene cluster of *P.chrysogenum* Oli13, isolated as described in Smith *et al.* (1990b), is shown in Figure 1. The DNA sequence of the region corresponding to the *acvA* gene was determined as described in Materials and methods (Figure 2).

An open reading frame (ORF) of 11 238 bp from a translation initiation codon situated 1107 bp upstream of the initiation codon of the *pcbC* gene and encoding a protein with a predicted molecular weight of 421 073 daltons was identified. The ORF indicates that the direction of transcription of the *acvA* gene is opposite to that of the *pcbC*

gene which correlates with the detection of *acvA* mRNA using single stranded RNA probes (Smith *et al.*, 1990b).

The acvA – pcbC intergenic region

The nucleotide sequence of the 1107 bp intergenic region between the *pcbC* and *acvA* genes was determined and is likely to contain all the information necessary for the divergent transcription and regulation of the two genes. The first 418 bp of the region 5′ to the *pcbC* gene of *P.chrysogenum* AS-P-78 has previously been described by Barredo *et al.* (1989b) and contains sequences that may be important for *pcbC* gene expression, particularly a TATA like sequence (TATAAT) at −198 bp. Examination of the sequence 5′ of the *acvA* gene also reveals a TATA like sequence (TATATA) at −228 bp. There do not appear to be any extensive areas of similarity between the non-coding regions immediately upsteam of the *acvA* and *pcbC* genes. Comparison with the 5′ non-coding region of the third gene (*aat*) of the penicillin biosynthetic gene cluster, which encodes acylCoA:6-aminopenicillanic acid acyltransferase (Barredo *et al.*, 1989a), also does not reveal any extensive similarity.

Several attempts were made to prepare intact *acvA* mRNA from *P.chrysogenum* Oli13 in order to map the 5′ ends of the *acvA* transcript. A number of different preparation methods were tried but Northern blot analysis revealed that the *acvA* mRNA was always substantially degraded and not of high enough quality for transcript mapping experiments. In these preparations the *pcbC* and *aat* gene transcripts were unaffected (results not shown). The start of the 'smear' of degraded *acvA* mRNA was >9.5 kb (the size of the largest marker used) which correlates with the reported size of *acvA* mRNA detected in the high-penicillin-producing strain, *P.chrysogenum* BW1901 (Smith *et al.*, 1990b) and *A.nidulans* (MacCabe *et al.*, 1990).

acvA coding sequence

In the absence of any transcript mapping or N-terminal amino acid sequence data it was assumed that the first ATG of the ORF is used for translation initiation. However, we note that there are two further in-frame ATGs located close to the first, one of which (at +54 bp) is in a 10 bp segment of DNA which is identical to the region containing the *pcbC* translation initiation codon (Figure 2).

The CodonPreference and TestCode programs supplied as part of the UWGCG sequence analysis package (Devereux *et al.*, 1984) were used to determine the coding potential of the predicted ORF. The sequence of the *P.chrysogenum*, *pyrG*, *pcbC*, *trpC* and 3-phospho-D-glycerate kinase genes were used to construct a codon frequency table (using CodonFrequency) for input to CodonPreference. In this manner it was shown that the ORF exhibited *P.chrysogenum* codon bias, codon third position GC bias (Bibb *et al.*, 1984) and compositional bias (Fickett, 1982) throughout the 11 238 bp (results not shown). Outside the ORF no such bias was found, confirming the position of the predicted initiation and termination signals. No evidence was found to support the presence of introns in the *acvA* gene.

The 3′ non-coding region

The *acvA* ORF terminates at a TGA codon and several other termination codons in all three possible reading frames are located in the subsequent 150 bp. A putative eukaryotic polyadenylation signal (AATAAA) is found 129 bp down from the predicted termination codon (Figure 2).

The ACVS protein

The polypeptide predicted to be encoded by the *acvA* gene contains 3746 amino acids and has a molecular weight of 421 073 daltons; this is considerably larger than the estimated size of purified ACVS from *A.nidulans* (230 kd).

We have previously shown that a 2.2 kb fragment of the *Flavobacterium* sp. SC 12,154 gene encoding ACVS hybridized to three separate regions of the *P.chrysogenum* Oli13 *acvA* gene (Smith *et al.*, 1990b). This indicated either that there are repeated domains within the *P.chrysogenum* Oli13 ACVS or that regions of the *Flavobacterium* sp. SC 12,154 gene had been rearranged relative to the *P.chrysogenum* Oli13 gene. To investigate these possibilities further we compared the amino acid sequence from regions of the *P.chrysogenum* Oli13 ACVS with the remainder of the sequence.

The first 1200 amino acid residues of the *P.chrysogenum* Oli13 ACVS were compared with the following 2100 residues using UWGCG Compare with a window of 30 and a match stringency of 70%. This showed that a region of ~650 amino acids in the first 1200 shared similarity with two distinct regions of approximately equivalent size in the remainder of the protein. The results were plotted graphically using DotPlot (Figure 3). The three similar regions were termed domains A, B and C.

With the assistance of UWGCG BestFit and Gap the core regions of similarity between domain A, B and C (~570

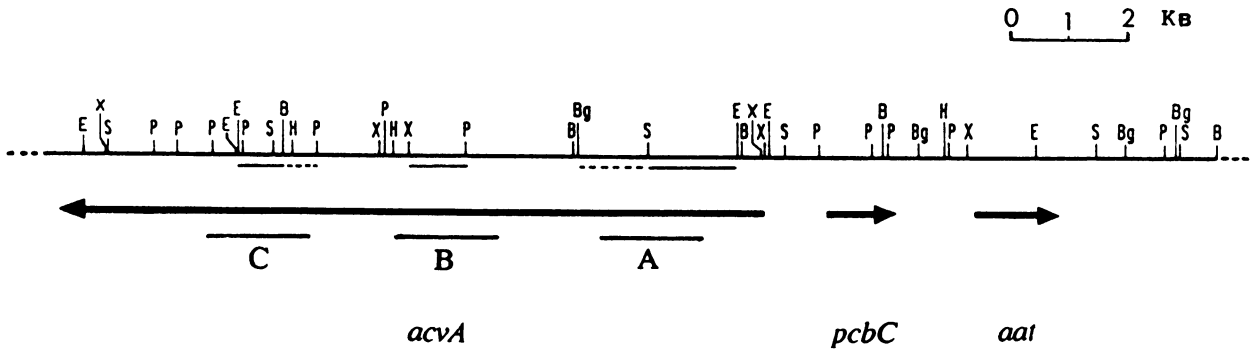

Fig. 1. Restriction map of a portion of cosmid clone pCX3.2 (Smith *et al.*, 1990b) showing the penicillin biosynthetic gene cluster of *P.chrysogenum* Oli13. The location of the *acvA*, *pcbC* and *aat* genes are shown with bold arrows indicating the direction of transcription. Lines shown underneath the arrow representing the *acvA* gene indicate the location of the conserved A, B and C domains in the ACVS protein. Lines shown above the arrow indicate the DNA fragments to which a 2.2 kb fragment of the *Flavobacterium* sp. SC 12,154 gene encoding ACVS hybridized (Smith *et al.*, 1990b). Dotted lines indicate less strongly hybridizing regions. E, *Eco*RI; X, *Xho*I; S, *Sal*I; P, *Pst*I; B, *Bam*HI; H, *Hin*dIII; Bg, *Bgl*II.

amino acids) were identified and aligned (Figure 4) and their position and extent within the *acvA* gene are marked on Figure 1. Optimal alignment of the three domains necessitated the insertion of gaps, none of these extensive (2.1 insertion/deletion events per 100 residues). Within the domains are several areas where there is continuous identity over a number of amino acids (Figure 4). Using UWGCG Gap the % identities and % similarities between the domains were determined (Table I). The core (as defined in Figure 4) of domain A has 577 amino acids, B has 568 and C 572.

The location within the *P.chrysogenum* Oli13 *acvA* gene of the DNA encoding the A, B and C domains is shown in Figure 1. Examination of the location of these DNA regions shows that they correspond with areas identified as having homology with a 2.2 kb DNA fragment of the ACVS gene from *Flavobacterium* sp. SC 12,154 (Smith *et al.*, 1990b).

The core of each ACVS domain, as shown in Figure 4, was used to search sequence databases using UWGCG FastA. Four proteins were identified as having some

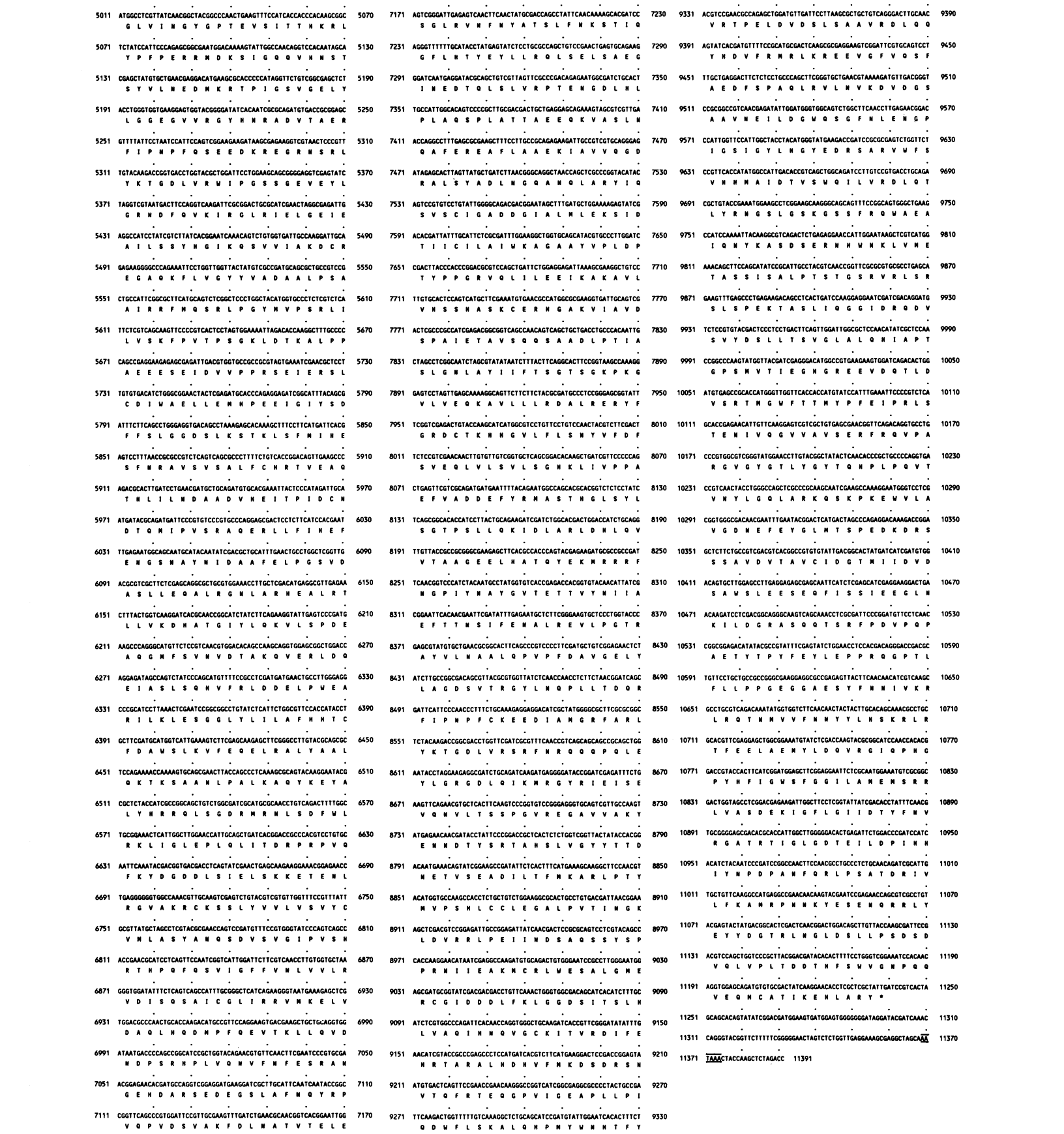

Fig. 2. DNA sequence of the *acvA* gene of *P.chrysogenum* Oli13 including 5′ and 3′ untranslated regions and the predicted amino acid sequence of the protein. The translation start site of the *pcbC* gene (at −1107) is underlined with an arrow. TATA like elements in the intergenic region between the *pcbC* and *acvA* genes are overlined and the putative *acvA* polyadenylation signal at +11369 is over and underlined.

similarity with each of the three conserved domains of ACVS. These were: *Photinus pyralis* (firefly) luciferase (de Wet *et al.*, 1987), *Petroselinum crispum* (parsley) 4-coumarate-CoA ligase (Lozoya *et al.*, 1988), *B.brevis* tyrocidine synthetase 1 (Weckermann *et al.*, 1988) and gramicidin S synthetase 1 (Kraetzschmar *et al.*, 1989). The % identity and % similarity between each of the ACVS domains and these proteins was determined with UWGCG Gap (Table I). One region was particularly conserved in the three ACVS A, B, C domains and the luciferase, 4-coumarate-CoA ligase, tyrocidine synthetase 1 and gramicidin S synthetase 1 proteins (Figure 5). As predicted from the above results, the luciferase, 4-coumarate-CoA ligase, tyrocidine synthetase 1 and gramicidin S synthetase 1 proteins also share similarity (Table I). Randomizing the sequences and then performing the comparisons again showed that the similarities found between the proteins were significant (results not shown).

A region of ~355 amino acids following ACVS domain C had some similarity with the sequence of gramicidin S synthetase 1 and tyrocidine synthetase 1 immediately following the region of these proteins identified as being similar to the A, B and C domains (results not shown). All other comparisons of the ACVS sequence outside the A, B and C domains with protein sequence databases and the proteins identified above did not reveal any significant similarities.

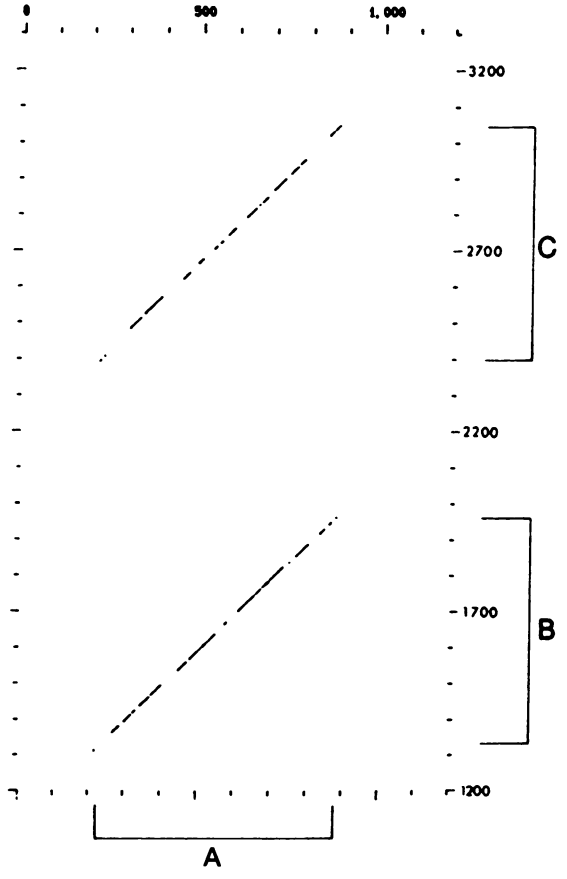

Fig. 3. Dot (diagon) plot comparing the initial 1200 amino acids of ACVS (*x*-axis) with the following 2100 residues (*y*-axis) using UWGCG Compare with a match window of 30 and a match stringency of 70%. Lines on the diagonal indicate areas meeting the match criteria and are termed domain A, B and C. The numbers on the axes correspond to amino acid positions in the protein.

Discussion

The sequencing of the *acvA* gene of *P.chrysogenum* Oli13 completes the sequence analysis of the three genes thought to be necessary for the biosynthesis of penicillin from amino acid precursors (Smith *et al.*, 1990a; Barredo *et al.*, 1989a; Carr *et al.*, 1986). The close linkage of penicillin biosynthetic genes suggests that they may be coordinately regulated and the divergent transcription of the *acvA* and *pcbC* genes from a 1107 bp intergenic region should allow simple functional analyses to be performed on this DNA to locate any elements affecting their transcription and regulation. Although sequences important for gene expression in filamentous fungi have not been well defined, potential transcription signals such as TATA like motifs, which are sometimes found upstream of the transcription initiation sites of a number of filamentous fungal genes (Gurr *et al.*, 1987), have been identified in the upstream sequence of both the *acvA* and *pcbC* genes. A eukaryotic polyadenylation signal, AATAAA, which is also sometimes found at the 3′ end of filamentous fungal genes (Gurr *et al.*, 1987), is found downstream of the *acvA* gene although its functional significance has not been demonstrated.

It is unclear whether the *P.chrysogenum* Oli13 *acvA* gene contains any introns. Introns in filamentous fungal genes are generally small, averaging <100 bp and several genes lack introns completely, including the *pcbC* gene (Carr *et al.*, 1986; Gurr *et al.*, 1987). The presence of many introns in the *acvA* gene appears unlikely because no breaks in the *acvA* ORF were identified, analysis of the sequence did not reveal any regions having a low coding potential and the estimated size of *P.chrysogenum acvA* mRNA is >9.5 kb (Smith *et al.*, 1990b). MacCabe *et al.* (1990) also estimated *A.nidulans acvA* mRNA to be >9.5 kb in size. However, the presence of small in-frame introns containing no termination signals cannot be excluded.

It is likely that the ACVS of *P.chrysogenum* closely resembles the *A.nidulans* ACVS (van Liempt *et al.*, 1989) and possesses the same multifunctional properties. The large protein predicted to be encoded by the *acvA* gene is perhaps then not surprising when the multifunctionality of the enzyme is considered. The specific recognition and adenylation of three amino acids, the formation of peptide bonds between

Table I. Similarities between *P.chrysogenum* Oli13 ACVS domains and other proteins

% Similarity[b]	% Identity[a]						
	A	B	C	T	G	P	L
A		43.1	38.0	33.5	34.6	26.7	21.9
B	64.2		39.3	37.2	36.8	22.6	21.5
C	59.9	59.8		33.6	33.3	20.4	21.5
T	57.6	61.6	59.3		56.5	21.6	23.6
G	55.1	58.8	56.4	72.0		23.3	22.5
P	50.3	48.2	48.2	48.1	46.9		35.0
L	47.3	46.4	49.0	44.5	46.4	57.1	

Key: A, B, C; *P.chrysogenum* Oli13 ACVS domains. T, *B.brevis* tyrocidine synthetase 1; G, *B.brevis* gramicidin S synthetase 1; P, parsley 4-coumarate-CoA ligase; L, firefly luciferase.
[a]Determined using UWGCG Gap (gap weight = 3.0, gap length weight = 0.1).
[b]Determined using UWGCG Gap (gap and gap length weight as above) where mismatches are weighted according to the evolutionary distance between the amino acids compared (Gribskov and Burgess, 1986).

them via enzyme bound intermediates and the epimerization of one amino acid finally followed by cleavage of the complete tripeptide from the enzyme must require a complex protein capable of recognizing and assembling the components in the correct order.

The reason for the discrepancy between the estimated size of purified *A.nidulans* ACVS of 230 kd and the predicted size for the *P.chrysogenum* Oli13 ACVS of 421 kd, based on the gene sequence, is unclear. It may simply reflect a genuine major difference between the enzymes, which is probably unlikely given the close evolutionary relationship between the two fungi and the intergeneric hybridization pattern (Smith *et al.*, 1990b), or indicate the presence of many introns in the *P.chrysogenum acvA* gene. As discussed above, this is also unlikely. Alternatively the size of the *A.nidulans* ACVS may have been wrongly estimated or the protein is post-translationally processed.

The identification of three extensive regions with high similarity in the *P.chrysogenum* Oli13 ACVS protein supports the data obtained earlier by DNA homology between the *P.chrysogenum* Oli13 *acvA* gene and the *Flavobacterium* sp. SC 12,154 gene encoding ACVS (Smith *et al.*, 1990b), which indicated that such repeated regions may be present and are homologous. The function of these domains

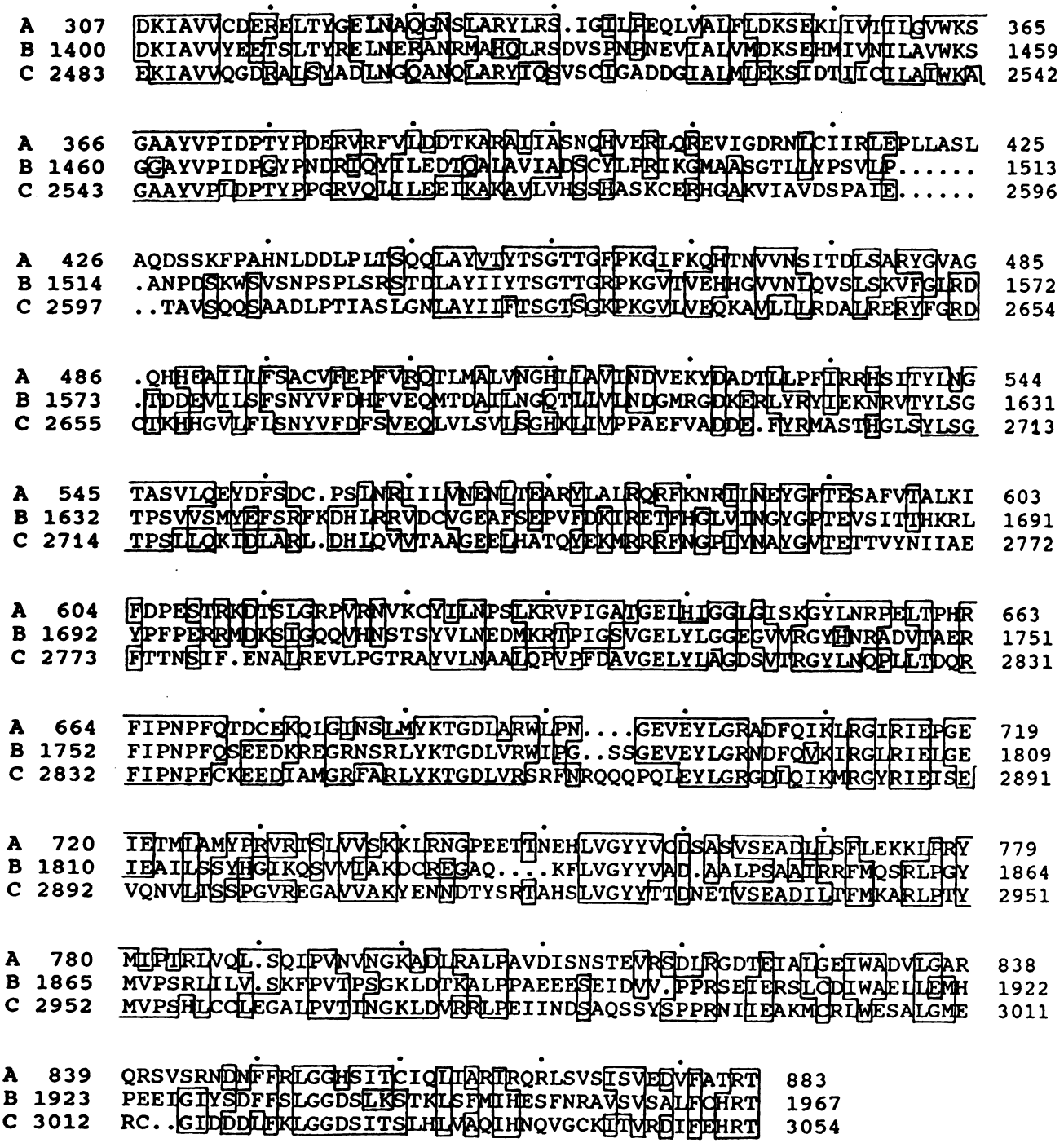

Fig. 4. Alignment (assisted by UWGCG Bestfit and Gap) of the core region of homology between the A, B and C domains of the *P.chrysogenum* Oli13 ACVS. Identical amino acids are boxed. Numbers refer to amino acid position within the protein.

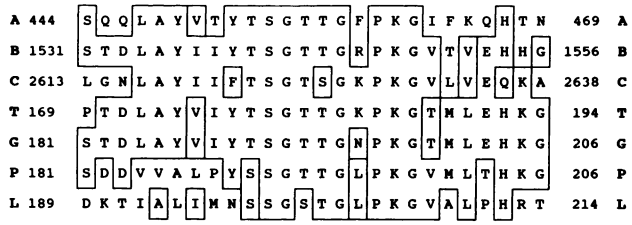

Fig. 5. Alignment (assisted by UWGCG Gap) of the region of high similarity between the A, B and C domains of the *P.chrysogenum* Oli13 ACVS, *B.brevis* tyrocidine synthetase 1 (T), *B.brevis* gramicidin S synthetase 1 (G), parsley 4-coumarate-CoA ligase (P) and firefly luciferase (L) proteins. Numbers refer to amino acid position within the respective proteins.

is unknown. However, the discovery of similarity between each of the ACVS domains and four other proteins, all of which recognize and activate by adenylation an amino acid or perform ATP–pyrophosphate exchange reactions, is interesting. It has been suggested (van Liempt *et al.*, 1989), that ACV biosynthesis proceeds via a multi-enzyme thio-template mechanism (Kleinkauf and von Dohren, 1987) and it is tempting to speculate that each of the three ACVS A, B and C domains may recognize and adenylate one of the constituent amino acids of ACV. Separation and expression of the DNA encoding each of the domains individually in heterologous hosts may allow enzymic functions to be assigned to each domain. It will also be interesting to see whether changing the order of the domains has any affect on the sequence of the tripeptide produced.

The three domains of the *P.chrysogenum* Oli13 ACVS have some similarity with the N-terminal portions of the *B.brevis* tyrocidine synthetase 1 and gramicidin S synthetase 1. These are closely related bacterial multifunctional enzymes which perform the initial steps in the synthesis of the cyclic peptide antibiotics tyrocidine and gramicidin S by activating phenylalanine and epimerizing it to the D-isomer (Marahiel *et al.*, 1985; Kraetzschmar *et al.*, 1989). A region following the ACVS C domain also has some similarity with the C-terminal portions of tyrocidine synthase 1 and gramicidin S synthetase 1. It therefore appears that the A and B domains are truncated versions of a more extensive similarity between the C-terminal region of ACVS (including the C domain) and the major part of tyrocidine synthetase 1 and gramicidin S synthetase 1.

Small peptides with antibiotic/toxic activity are made by a wide range of organisms (Kleinkauf and von Dohren, 1987; Kurahashi, 1974) and the identification of similarity between the *P.chrysogenum* ACVS and two bacterial enzymes involved in antibiotic peptide synthesis indicates that some of the bacterial and fungal multifunctional peptide synthetases may have a common ancestral origin.

The DNA encoding the A, B and C domains identified in the *P.chrysogenum* Oli13 ACVS align with *acvA* gene DNA fragments which hybridized to a 2.2 kb fragment of the *Flavobacterium* sp. SC 12,154 ACVS gene. It therefore appears that the *Flavobacterium* sp. SC 12,154 ACVS may also contain these domains, one of which was probably encoded by the DNA within the 2.2 kb fragment used as a probe to construct the intergeneric hybridization pattern (Smith *et al.*, 1990b).

There are precedents for multifunctional enzymes in filamentous fungi existing as separate monofunctional enzymes in prokaryotes. The pentafunctional *AROM* locus (for aromatic amino acid biosynthesis) of *A.nidulans* is related to the monofunctional bacterial *aro* loci (Charles *et al.*, 1986). As appears to be the case with the *P.chrysogenum acvA* gene, the *A.nidulans AROM* gene does not contain any introns and it has been suggested (Hawkins, 1987) that it may have arisen from multiple gene fusion of the bacterial genes. Perhaps a more interesting example is found in cephalosporin biosynthesis. The filamentous fungus *C.acremonium*, which produces cephalosporins, possesses a dual function enzyme which performs both ring expansion of penicillin N and the subsequent hydroxylation step (Samson *et al.*, 1987). In contrast, in the cephalosporin synthesizing prokaryote *S.clavuligerus*, these activities are located on diferent enzymes (Jensen *et al.*, 1985).

It has been suggested (Ramon *et al.*, 1987; Weigel *et al.*, 1988) that some β-lactam biosynthetic genes arose in prokaryotes and were transferred to a progenitor of the filamentous fungi synthesizing β-lactam antibiotics. It will therefore be interesting to determine whether *Flavobacterium* sp. SC 12,154 contains one ACVS gene, encoding a large multifunctional protein, or is composed of separate genes for each of the domains identified in the *P.chrysogenum* ACVS. The sequencing of the *Flavobacterium* sp. SC 12,154 gene encoding ACVS will resolve this question.

The availability of the *acvA* gene sequence should facilitate studies of the *P.chrysogenum* ACVS, a complex, multi-functional and commercially important peptide synthetase.

Materials and methods

Strains and plasmids

The isolation of the *acvA* gene of *P.chrysogenum* Oli13 on the cosmid clone pCX3.2 has been described previously (Smith *et al.*, 1990b).

Sequencing methods

CsCl/ethidium bromide density gradient purified plasmid DNA was prepared for sequencing in the following manner: 3–5 μg of plasmid DNA in 20 μl of H_2O was denatured by the addition of 20 μl of 0.4 M NaOH, 0.4 mM EDTA and left at room temperature for 5 min. The denatured DNA was neutralized by the addition of 4 μl of 2 M sodium acetate (pH 4.7), precipitated by the addition of 100 μl of ethanol at −20°C followed by 15 min at −70°C and the DNA pellet collected by centrifugation, washed with 70% ethanol and dried.

Denatured plasmid DNA was sequenced using a commercially available kit (Sequenase; United States Biochemical Corporation) according to the instruction manual supplied with the kit.

DNA fragment isolation from agarose gels, subcloning and plasmid isolation was performed using standard techniques (Maniatis *et al.*, 1982).

Sequencing strategy

Fragments of the *acvA* gene contained on cosmid clone pCX3.2 were subcloned in the plasmids pUC19 (Yannisch-Perron *et al.*, 1985) and pBluescript (Stratagene, La Jolla, CA, USA) and sequenced using M13 universal and reverse oligonucleotide primers (United States Biochemical Corporation). The sequence thus generated allowed the synthesis of novel 18 or 20 bp oligonucleotide primers capable of initiating second strand synthesis on previously unsequenced DNA. In this way the complete nucleotide sequence of the region corresponding to the *P.chrysogenum* Oli13 *acvA* gene was determined for both strands using 63 oligonucleotide primers. Oligonucleotide primers were synthesized by Alta Bioscience (Department of Biochemistry, University of Birmingham, Birmingham, UK) and the Department of Biochemistry, University of Bristol, Bristol, UK.

Computing

Computer programs supplied as part of the University of Wisconsin Genetics Computer Group (UWGCG) sequence analysis package (Devereux *et al.*, 1984) and accessed through the UK Science and Engineering Research Council SEQNET system were used for all sequence analysis and comparison.

Acknowledgements

The authors wish to thank Patrick Downey and Bob Syce for assistance with DNA sequencing. This work was supported by Beecham Pharmaceuticals UK Division.

References

Banko,G., Wolfe,S. and Demain,A.L. (1986) *Biochem. Biophys. Res. Commun.*, **137**, 528–535.
Banko,G., Demain,A.L. and Wolfe,S. (1987) *J. Am. Chem. Soc.*, **109**, 2858–2860.
Barredo,J.L., van Solingen,P., Diez,B., Alvarez,E., Cantoral,J.M, Kattewilder,A., Smaal,E.B., Groenen,M.A.M., Veenstra,A.E. and Martin,J.F. (1989a) *Gene*, **59**, 291–300.
Barredo,J.L., Alvarez,E., Cantoral,J.M., Diez,B. and Martin,J.F. (1989b) *Mol. Gen. Genet.*, **216**, 91–98.
Bibb,M.J., Findlay,P.R. and Johnson,M.W. (1984) *Gene*, **30**, 157–166.
Carr,L.G., Skatrud,P.L., Scheetz,M.E., Queener,S.W. and Ingolia,T.D. (1986) *Gene*, **48**, 257–266.
Charles,I.G., Keyte,J.W., Brammar,W.J., Smith,M. and Hawkins,A.R. (1986) *Nucleic Acids Res.*, **14**, 2201–2213.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
de Wet,J.R., Wood,K.V., DeLuca,M., Helinski,D.R. and Subramani,S. (1987) *Mol. Cell. Biol.*, **7**, 725–737.
Fickett,J.W. (1982) *Nucleic Acids Res.*, **10**, 5303–5318.
Gribskov,M. and Burgess,R.R. (1986) *Nucleic Acids Res.*, **14**, 6745–6763.
Gurr,S.J., Unkles,S.E. and Kinghorn,J.R. (1987) In Kinghorn,J.R. (ed.), *Gene Structure in Eukaryotic Microbes*. IRL Press, Oxford, pp. 93–139.
Hawkins,A.R. (1987) *Curr. Genet.*, **11**, 491–498.
Ingolia,T.D. and Queener,S.W. (1989) *Med. Res. Rev.*, **9**, 245–264.
Jensen,S.E., Westlake,D.W.S. and Wolfe,S. (1985) *J. Antibiot.*, **38**, 263–265.
Jensen,S.E., Westlake,D.W.S. and Wolfe,S. (1988) *FEMS Microbiol. Lett.*, **49**, 213–218.
Kleinkauf,H. and von Dohren,H. (1987) *Annu. Rev. Microbiol.*, **41**, 259–289.
Kraetzschmar,J., Krause,M. and Marahiel,M.A. (1989) *J. Bacteriol.*, **171**, 5422–5429.
Kurahashi,K. (1974) *Annu. Rev. Biochem.*, **43**, 445–459.
Lara,F., Mateos,R.dC., Vazquez,G. and Sanchez,S. (1982) *Biochem. Biophys. Res. Commun.*, **105**, 172–178.
Lozoya,E., Hoffmann,H., Douglas,C., Schulz,W., Scheel,D. and Hahlbrock,K. (1988) *Eur. J. Biochem.*, **176**, 661–667.
MacCabe,A.P., Riach,M.B.R., Unkles,S.E. and Kinghorn,J.R. (1990) *EMBO J.*, **9**, 279–287.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Marahiel,M.A., Krause,M. and Skarpeid,H. (1985) *Mol. Gen. Genet.*, **201**, 231–236.
Nuesch,J., Heim,J. and Treichler,H.J. (1987) *Annu. Rev. Microbiol.*, **41**, 51–75.
Ramon,D., Carramolino,L., Patino,C., Sanchez,F. and Penalva,M.A. (1987) *Gene*, **57**, 171–181.
Samson,S.M., Belagaje,R., Blankenship,D.T., Chapman,J.L., Perry,D., Skatrud,P.L., Van Frank,R.M., Abraham,E.P., Baldwin,J.E., Queener,S.W. and Ingolia,T.D. (1985) *Nature*, **318**, 191–194.
Samson,S.M., Dotzlaf,J.E., Slisz,M.L., Becker,G.W., Van Frank,R.M., Veal,L.E., Yeh,W.-K., Miller,J.R., Queener,S.W. and Ingolia,T.D. (1987) *Bio/Technology*, **5**, 1207–1214.
Smith,D.J., Burnham,M.K.R., Edwards,J., Earl,A.J. and Turner,G. (1990a) *Bio/Technology*, **8**, 39–41.
Smith,D.J., Burnham,M.K.R., Bull,J.H., Hodgson,J.E., Ward,M.J., Browne,P., Brown,J., Barton,B., Earl,A.J. and Turner,G. (1990b) *EMBO J.*, **9**, 741–747.
van Liempt,H., von Dohren,H. and Kleinkauf,H. (1989) *J. Biol. Chem.*, **264**, 3680–3684.
Weckermann,R., Fuerbass,R. and Marahiel,M.A. (1988) *Nucleic Acids Res.*, **16**, 11841.
Weigel,B.J., Burgett,S.G., Chen,V.J., Skatrud,P.L., Frolik,C.A., Queener,S.W. and Ingolia,T.D. (1988) *J. Bacteriol.*, **170**, 3817–3826.
Yannisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

Received on April 6, 1990; revised on May 17, 1990