J. Mol. Microbiol. Biotechnol. (2000) 2(4): 543-550.

JMMB Research Article

Characterization and Expression of the Arginine Biosynthesis Gene Cluster of *Streptomyces clavuligerus*

Antonio Rodríguez-García, Álvaro de la Fuente, Rosario Pérez-Redondo, Juan F. Martín, and Paloma Liras*

Area of Microbiology, Faculty of Biology, University of León, 24071 León, Spain

Abstract

A cluster of genes *argCJBDRGH* containing most of the arginine biosynthesis genes has been found in *Streptomyces clavuligerus* after sequencing a 8.3 kb DNA region containing overlapping sequences of two DNA fragments known to contain arginine biosynthesis genes. Subcloning, complementation of *E. coli* arginine auxotrophic strains and enzymatic assays confirmed the identity of each gene. S1 nuclease mapping studies and Northern hybridization analysis revealed the formation of two large transcripts corresponding to *argCJBDR* and *arg*GH. The amount of each of these mRNAs is 10 to 44 times higher in a *S. clavuligerus argR*-disrupted mutant than in the wild type confirming the existence of an ArgR-mediated control of arginine biosynthesis gene expression. A low level constitutive monocistronic transcript of *argR* was observed *in S. clavuligerus* cells. Most of the *argGH* transcript initiating at an adenine 29 nt upstream of the *argG* initiation codon appears to stop at a termination stem and loop structure present downstream of the *argG* gene.

Introduction

Arginine biosynthesis genes (*arg*) and genes for carbamoyl phosphate (required for citrulline) biosynthesis (*car*) are usually scattered in the genome of most Gram negative bacteria. In *E. coli* there is an *arg*ECBH cluster, in which the *argE* and *argCBH* are transcribed divergently (Crabeel *et al*., 1979). The *carAB* genes are also clustered, while other genes (*argA, argB, argD, argI, argF, argG*) map at different positions in the *E. coli* chromosome. The *argECBH* cluster is also found in *Salmonella typhimurium* (Sanderson, 1970) while in *Proteus mirabilis* and *Serratia marcescens* includes additionally *argG* upstream of *argH* (Prozesky 1968, Matsumoto *et al*., 1975). In *Neisseria gonorrohoeae* and *Pseudomonas aeruginosa* all the *arg* genes are scattered (Picard and Dillon, 1989; Haas *et al*., 1977).

In Gram positive bacteria clustering of arginine genes appears to be common. In *Lactobacillus plantarum* a *carAargCJBDF* cluster has been described (Bringel *et al*., 1997) with divergent transcription for the *car* and the *arg* genes. The same organization for arginine genes (*argCJBDF*) was described in *Corynebacterium glutamicum* (Sakanyan *et al*., 1996); in *Mycobacterium tuberculosis* and *Mycobacterium leprae,* DNA sequences for a *argCJBDFRGH* cluster have been deposited in EMBL database (Accession numbers Z85982 and L78811, respectively). Arginine biosynthesis genes in *Bacillus subtilis* are organized in two operons, one is *argCJBD-carAB-argF* for early steps in the arginine pathway, including those involved in carbamoyl phosphate biosynthesis (Mountain *et al*., 1984; O'Reilly and Devine, 1994), and a second operon *argGH* encodes enzymes for the late steps of the pathway (Piggot and Hoch, 1985). An *argE* gene has been located by sequencing the *B. subtilis* genome (Kunst *et al*., 1997) but its relation with the cyclic pathway of arginine biosynthesis in this bacterium (Figure 1) remains to be elucidated; the regulatory gene *ahrC* encoding the arginine biosynthesis repressor (North *et al*., 1989) is located in a separate region. This organization in two separate operons may occur also in other Gram positive bacteria and could explain the lack of information on the location of *argG, argH* and *argR* in *Lactobacillus* and Corynebacteria.

Arginine defective mutants of *Streptomyces coelicolor* have been mapped as scattered at four different locations in the chromosome (Redenbach *et al*., 1996). By complementation of *E. coli* arginine defective mutants a cluster *argCJB* has been located in *S. coelicolor* (Hindle *et al*., 1994) whereas the *argG* gene is located in an unstable region at the end of the lineal chromosome of this actinomycete (Redenbach *et al*., 1996). Fragments of *S. clavuligerus* genome were also able to complement *E. coli* arginine mutants and the subsequent sequencing of those fragments revealed the presence of two subclusters of arginine genes in this actinomycete: *argCJ* (Ludovice *et al*., 1992) and *argRGH* (Rodríguez-García *et al*., 1995; 1997).

An important question from the regulation point of view is if the two clusters organization occurs in all actinomycetes and how expression of these clusters is controlled by the ArgR repressor. This information is particularly important in *S. clavuligerus*, since this actinomycete uses arginine as precursor for clavulanic acid biosynthesis (Valentine *et al*., 1993). Knowledge of the arginine gene cluster organization and their regulatory mechanisms (Rodríguez-García *et al*., 1997) is important to understand how the flow of arginine affects clavulanic acid production.

In this paper we provide evidence showing that most of the arginine genes in *S. clavuligerus* are present in a single *argCJBDRGH* organization, which includes genes for the early and late steps of the arginine pathway as well as the regulatory gene *argR*. We have also shown that *ca*-ORF4, in the clavulanic acid cluster, encodes a second functional ornithine N-acetyltransferase similar to ArgJ.

Received January 1, 2000; revised April 14, 2000; accepted April 14, 2000.
*For correspondence. Email degplp@unileon.es; Tel. +34-987-291504; Fax. 34-987-2915 06.

Results

Location of the Arginine Gene Cluster of *S. clavuligerus*

The *argC* gene was initially located by sequencing a 6.1 kb *Sau*3AI DNA fragment of *S. clavuligerus* present in plasmid pULML31, which complements the *argC* deficient mutant *E. coli* XC33 (Ludovice *et al*., 1992).

Independently, a 18 kb *Sau*3AI DNA fragment of *S. clavuligerus* was found to complement *E. coli argG* deficient mutants (Rodríguez-García *et al*., 1995). Further studies showed the presence in this DNA fragment of an *argRGH* cluster (Rodríguez-García *et al*., 1997).

In order to establish the definitive organization of the arginine biosynthesis genes, a 0.5 kb *Sac*I-*Sph*I DNA fragment (Figure 2A) located downstream of *argC* in pULML31 was mapped and partially sequenced. In parallel, a 2.1 kb *Sac*I-*Bgl*II DNA fragment from pULAR1, upstream of *argR* (Figure 2A), was also sequenced. The nucleotide sequence of both fragments was found to overlap, indicating that *argC* and *argR* (previously though to be in two different clusters) are closely located in the *S. clavuligerus* genome. Therefore, a DNA region, covering 8.3 kb was sequenced in both orientations.

Analysis of the *arg* Cluster

The G+C content of the sequenced fragment was 71.3%, with no significant differences of G+C in the intergenic regions. Rare TTA codons for leucine were absent from the coding regions, indicating that these genes are not regulated by the *bldA*-dependent translation control (Leskiw *et al*., 1991). Analysis of the nucleotide sequence of the entire fragment using the Geneplot Program of DNAstar revealed the presence of eight ORFs, oriented in the same direction. Comparison of the amino acid sequences deduced from the ORFs with the sequences of arginine metabolic enzymes shows that ORF1 to ORF7, correspond to arginine biosynthesis genes. These genes include, as expected, *argC* (nt 123-1148) which correspond to ORF1, *argR* and *argG* (nt 4600 to 6604) corresponding to ORF5 and ORF6. The proteins encoded by ORF2, ORF3 and ORF4 were homologous with ArgJ, ArgB and ArgD of other organisms while ORF7 encodes *argH*. ORF8 was truncated and its deduced amino acid sequence does not correspond to any enzyme of the arginine biosynthesis pathway.

Putative ribosome binding sites were found upstream of *argJ, argB, argD* and *argH*, but according to the *Streptomyces* promoter criteria (Strohl, 1992) no clear –10 and –35 boxes were found upstream of the newly described ORFs. Two inverted sequences able to form a stem and loop structure with free energy of –37.2 kcal/mol were found downstream of *argG* (nt 6628-6640/6650-6662). They might act as a transcription terminator of the *argG* gene. Similar putative terminator sequences were present downstream of *argH* (nt 8162-8172/8182-8192 and 8171-8178/8183-8190) giving stem and loop structures with free energies of –20.2 and –12.6 kcal/mol.

*arg*J Encodes a Functional Ornithine N-Acetyltransferase: Evidence for a Cyclic Arginine Pathway

argJ (ORF2) has 1,152 nt and encodes a protein of 383 amino acids with a deduced Mr of 39,733. The amino acid sequence of this protein showed the highest similarity

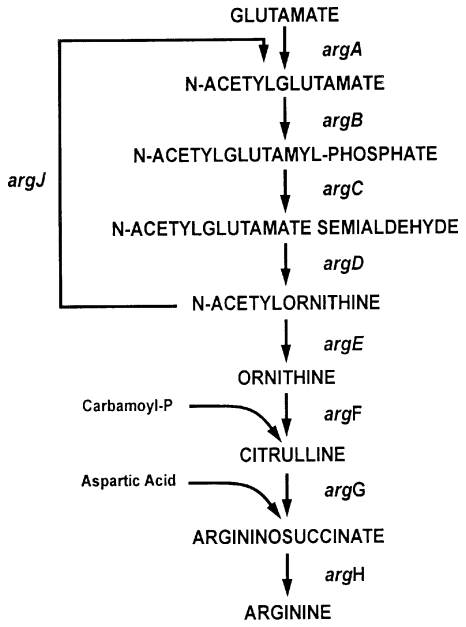

Figure 1. Pathways for arginine biosynthesis in different microorganisms.

(74.9% identical residues) with the unpublished sequence of ArgJ of *S. coelicolor* (Sanger cosmid SCL24) followed by the ornithine acetyltransferases of *M. tuberculosis* (47.8% identity) and *C. glutamicum* (40.7% identity).

E. coli contains N-acetylornithinase but not the *argJ*-encoded ornithine N-acetyltransferase. Bifunctional ornithine N-acetyltransferases (EC 2.3.1.35) with N-acetylornithinase (EC 3.5.1.16) activity have been reported in some microorganisms. In order to test whether the *argJ* gene of *S. clavuligerus* confers acetylornithinase activity a 2.1 kb *Sac*I-*Bgl*II DNA fragment containing *argJ* was subcloned in both orientations in pBSKS(+) (giving plasmids pKS21.1 and pKS21.2) and introduced into *E. coli* XS1D2, an *argE* mutant lacking acetylornithinase activity. Ornithine N-acetyltransferase activity (0.76 units/mg protein) was found in cell-free extracts of *E. coli* XS1D [pKS21.1] in which the *S. clavuligerus argJ* was expressed from the *lacZ* promoter but no activity was found in *E. coli* XS1D[pKS21.2] in which the gene was subcloned in the opposite orientation. No N-acetylornithinase activity was found in any of the transformants. These results confirm that *S. clavuligerus* contains a cyclic N-acetyltransferase pathway as reported by Hindle *et al*. (1994) in *S. coelicolor,* and not the linear (*E. coli* type) pathway.

A Second Gene Encoding an N-Acetylornithine Acetyltransferase Occurs in the Clavulanic Acid Cluster

ArgJ shows a 31.1% identity in amino acids over the entire protein sequence with the protein encoded by ORF4 of the clavulanic acid gene cluster (hereafter named *ca*-ORF4) of *S. clavuligerus* (Hodgson *et al*., 1995). The amino acid sequence of the protein encoded by *ca*-ORF4 contains most of the conserved boxes found in ArgJ proteins. However the similarity of the protein encoded by *ca*-ORF4 with the ArgJ protein of *S. clavuligerus* (31.1%) is lower than the similarity between ArgJ proteins of other Gram

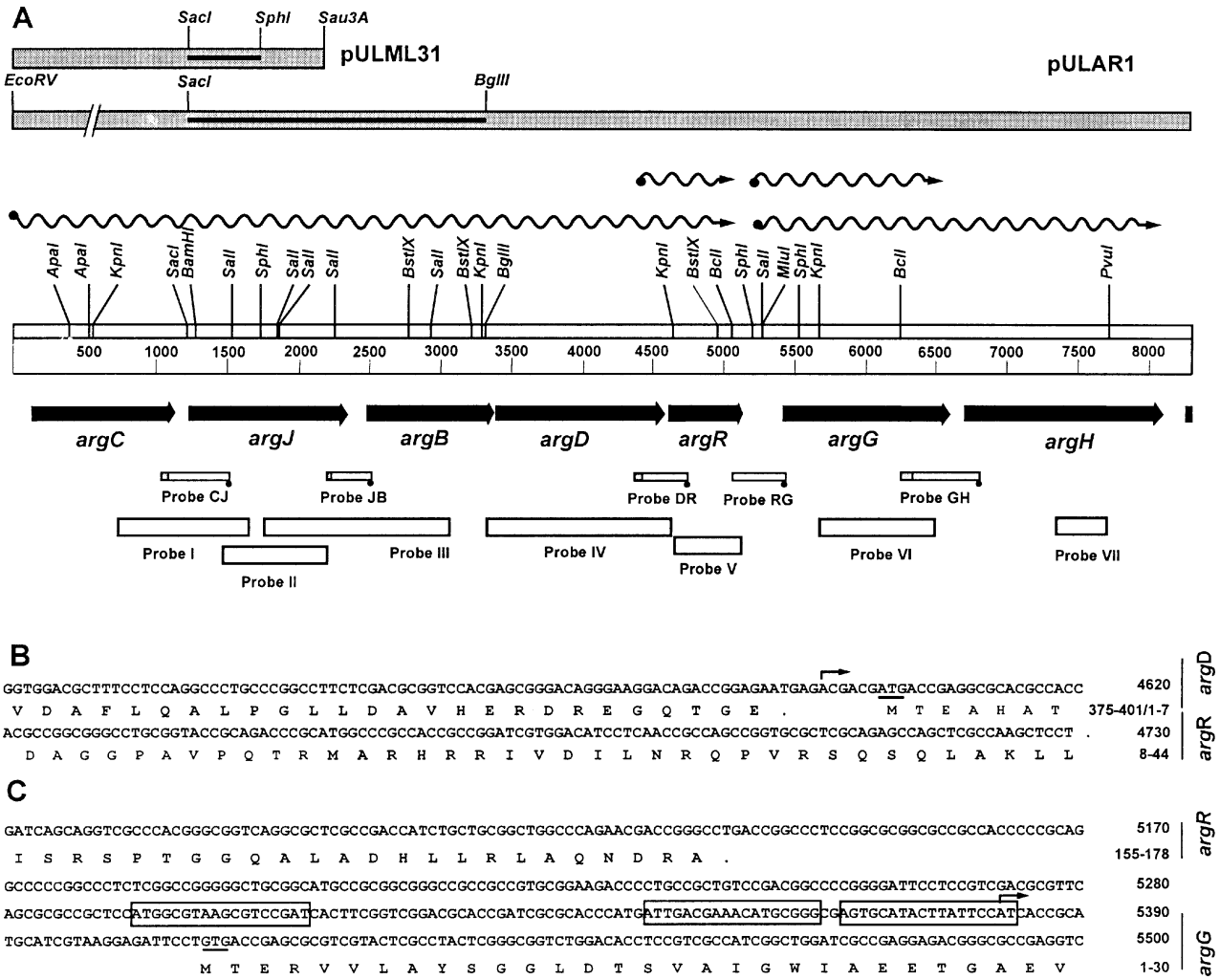

Figure 2. Organization of the arginine biosynthesis gene cluster in *S. clavuligerus*.
A. Shaded boxes at the top show plasmids pULML31 and pULAR1 and black bars indicate overlapping sequences detected initially. Wavy lanes show the transcription units originated from the cluster. The arginine biosynthesis genes are shown by black arrows. Empty boxes indicate the probes used in the S1 and Northern experiments. In probes CJ, JB, DR, RG and GH the shadowed regions indicate heterologous (*E. coli*) DNA and the dots the position of labeling. B. Intergenic region *argD-argR* showing the amino acids of the carboxyterminal-end of *argD* and the N-terminal end of *argR*. The bent arrow indicates the transcription initiation of the monocistronic *argR* mRNA. C. Intergenic region *argR-argG* showing the amino acids of the carboxyl-end of *argR* and the N-terminal end of *argG*. The regulatory Arg-sequences present upstream of the *argG* gene are boxed. The bent arrow indicates the transcription initiation of the polycistronic *argGH* mRNA.

positive bacteria. In order to test if the protein encoded by *ca*-ORF4 shows ornithine N-acetyltransferase activity, a 4.1 kb *Kpn*I-*Nru*I DNA fragment containing *ca*-ORF4 was subcloned in pBSKS(+) to give plasmid pBS41. Cell-free extracts of the transformant *E.coli* XSD1[pBS41] showed an ornithine N-acetyltransferase activity of 1.79 units/mg protein indicating that promoter sequences upstream of *ca*-ORF4 are expressed in *E. coli* and confirming that the gene encodes an ArgJ-homologous protein with ornithine N-acetyltransferase activity but no acetylornithinase activity. Therefore we confirm that in *S. clavuligerus* there are at least two genes encoding ornithine N-acetyltransferases (Eascott *et al.*, 1998). The low amino acid identity of the protein encoded by *ca*-ORF4 with other ArgJ proteins suggest that the role of this gene in *S. clavuligerus* might not be related to arginine biosynthesis, but perhaps involved in modification of arginine (e.g. acting as arginine N-acetyltransferase) as a way to commit it to clavulanic acid biosynthesis.

The Protein Encoded by *argB* Complements *argB*⁻ Mutants of *E. coli*

The *argB* (ORF3, 909 nt) gene has a G+C content of 68.5%, and encodes a protein of Mr 32,365. The N-acetylglutamate kinase protein (EC 2.7.2.8) has 302 amino acids and a sequence similar to the homologous proteins of *C. glutamicum* (50.0% amino acid identity) and *M. tuberculosis* (47.8% identity). A motif conserved in all the ArgB proteins is present in the ArgB protein of *S. clavuligerus*, ^{66}VHGGGPXI^{73}. To confirm the identity of the gene, plasmid pULB, containing the *argB* gen, was used to complement the *argB* deficient strain *E. coli* XB25. Transformant *E. coli* [pULB] was able to grow in VB minimal medium while the control strain *E. coli* XB25[pBSKS+] was not.

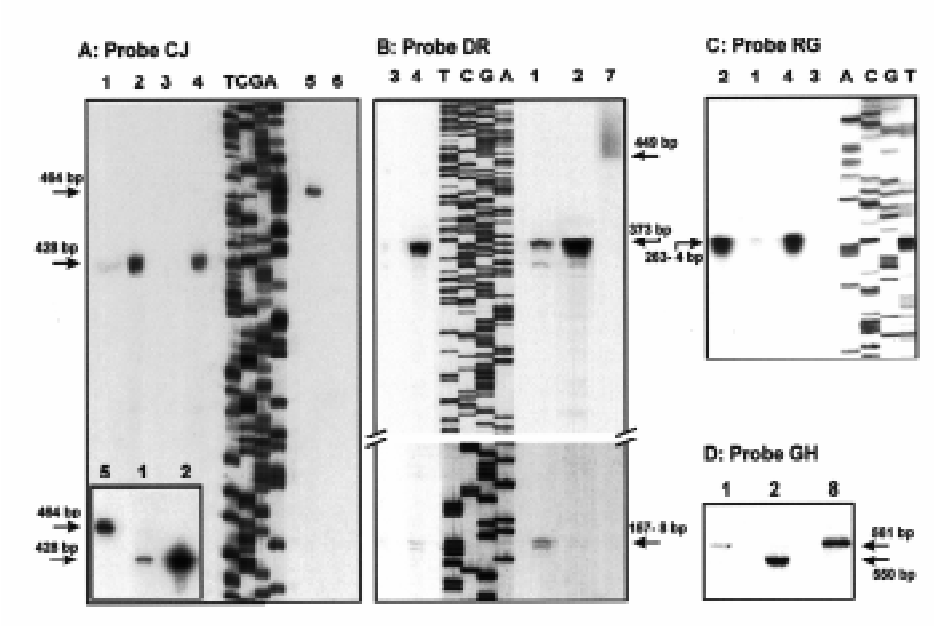

Figure 3. S1 mapping of the transcripts of the *arg* cluster using total RNA from 1,3) *S. clavuligerus* ATCC 27064 and 2,4) *S. clavuligerus argR*::*aph*. Protection of the corresponding transcript fragment was made with A) probe CJ , B) Probe DR, C) Probe RG, D) Probe GH. RNA was isolated from cells grown in GSPG medium (1,2) or TSB medium (3,4). Controls: 5) Probe CJ. 6) tRNA from *S. cerevisiae* 7) Probe DR. 8) Probe GH. The arrows show the probes used and the protected bands.

Characteristics of *argD*

ORF4 (1203 nt with a G+C content of 71.3%) encodes a 400 amino acid protein with a calculated Mr of 41,972. The argB stop codon overlaps with the initiation codon of ORF3. A search in the Swiss-Prot database with the FASTA Program revealed that the protein encoded by ORF3 has a high similarity to acetylornithine aminotransferases (ACOAT) and ornithine aminotransferases (OAT) (54.5% identity with the ArgD of *M. tuberculosis* and 47.8% with that of *C. glutamicum*).

The sequence ^{214}LVLDEVQTGIGRTGHWFAA QAEGVEADVVTLAKGLGGG^{251} corresponds to the pyridoxal phosphate (PP) binding motif, with K^{246} as the conserved lysine for covalent PP-binding (Yonaha *et al*., 1992). Multiple alignment of the ORF4-encoded protein with ten known ArgD proteins showed the presence of domains conserved in ACOAT's and OAT's, but the amino acid similarity of the ORF4-encoded protein is higher to ACOATs of *B. subtilis* and *E. coli* than to the corresponding OAT proteins. This observation is consistent with the amination of N-acetylglutamic semialdehyde to N-acetylornithine (which occurs when the intermediate is in the N-acetylated form).

In order to confirm that the cloned gene corresponds to *argD*, plasmid pULD was used to transform *E. coli* CGSC 4538, an *argD, proA* double mutant. The ampicillin resistant transformants were able to grow in VB medium supplemented with proline but also in VB medium without supplementation with either proline or arginine. Since the *proA* mutation precludes the formation of glutamate-γ-semialdehyde required for proline biosynthesis, the complementation of both *proA* and *argD* phenotypes by the *argD* gene of *S. clavuligerus* suggests an interconversion of intermediates of the proline and arginine pathways.

Characterization of *argH*

ORF7 (1,422 nt), encoding the *argH* gene (argininosuccinate lyase) was described to be located downstream of *argG* by partial sequencing of the gene (Rodríguez-García *et al.,* 1995). It is preceded by a ribosome binding motif AGGAG. The protein ArgH has a calculated Mr of 50,915. The amino acid sequence shows the highest similarity with ArgH of *M. tuberculosis* (60.4% amino acid identity) followed by the homologous human and rat proteins (43 and 42% identity). The amino acid sequence ^{282}GSSIMPQKKN^{291} found in the protein encoded by ORF7 corresponds to the GSXXMXXKXN motif characteristic of all fumarate lyases in which the central methionine residue appears to be involved in the active center of the enzyme (Woods *et al*., 1988).

The *argH* gene present in pULAR11 downstream of *argG* was used to transform *E. coli* CGSC 5359, an *argH* auxotroph. *E. coli* 5359[pULAR11] ampicillin resistant transformants grew in VB minimal medium supplemented with ampicillin, while the *E.coli* 5359 strain was unable to grow in VB medium. The growth of *E. coli* 5359[pULAR11] was slow probably due to low efficiency of transcription of *S. clavuligerus argH* gene (see below). This gene is expressed from the *S. clavuligerus argG* promoter which is known to be functional in *E. coli* (Rodríguez-García *et al*., 1995).

Location of the Arginine Genes in *Streptomyces coelicolor*

The *S. coelicolor* arginine biosynthesis genes have been reported to be scattered in the genome by classical genetic methods (Redenbach *et al*., 1996), but the relation between the locus detected and the proteins encoded has not been further elucidated. To confirm whether the arginine genes were scattered a blot of an *S. coelicolor* DNA cosmid library provided by H. Kieser (Norwich, U.K.) was hybridized at 68ºC with probes internal to *S. clavuligerus argJ, argD, argR, argG and argH* genes (probes II, IV, V, VI and VII in Figure 2). Several cosmids that gave faint to strong hybridization (6C12, 8A7, 3H2, D6, 1A4, D40, L10 and L24) and cosmids described as containing arginine-related locus (M1, M2, 4G1) were provided by H. Kieser and throughly

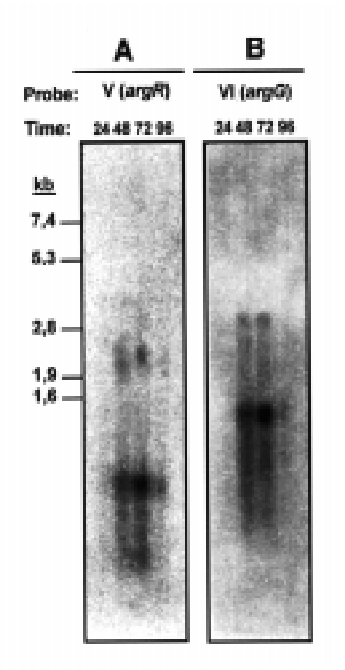

Figure 4. Time course of expression of the arginine genes as shown by Northern hybridization of total RNA from *S. clavuligerus* ATCC 27064 grown in GSPG medium for 24, 48, 72 and 96 h. The probes used were V (*argR*) (left panel) and VI (*argG*) (right panel). The standards correspond to RNA type II markers (Boehringer).

analyzed. Each cosmid was digested with *Not*I and electrophoresed in 0,8% agarose; the DNA fragments were blotted to Hybond NX membrane (Amersham) and hybridized with the different probes. Only cosmids L10 and L24 gave strong hybridization signals. A 3.2 kb *Not*I DNA fragment of cosmid L10 hybridized with probes internal to *argJ, argD* and *argR* indicating that these genes are contiguous in the *S. coelicolor* genome. A 4.9 kb *Not*I DNA band of cosmid L24 gave hybridization with the same probes suggesting that the genes *argJ*, *D* and *R* are probably located in the overlapping region between L10 and L24. Only the 4.9 kb DNA fragment of cosmid L24 gave hybridization with the *argH* probe. Cosmid 4G1, in which the *argG* gene has been located, did not gave positive hybridization with the *argG* probe as expected due to the major sequence differences between the *argG* genes of *S. clavuligerus* and *S. coelicolor* (Rodríguez-García *et al.*, 1995).

Transcription of *arg* Genes in the Wild Type *S. clavuligerus* and in a *S. clavuligerus argR*-Disrupted Mutant

In order to study the transcription of arginine biosynthesis genes, a *S. clavuligerus arg*R-disrupted mutant was obtained by transformation with plasmid pHZ*argR*-. About 2% of the transformants were mutants disrupted in *argR* by double recombination and showed the phenotype *kan*^{R} *ts*^{S}. The mutation was confirmed by hybridization of total DNA of the wild type strain and the disrupted mutant with probes internal to *argR* and *aph*II (*kan*^{R}) genes.

The transcription of the genes in the arginine cluster was studied by Northern hybridization and S1 mapping using total RNA of the wild type *S. clavuligerus* ATCC 27064 and the *argR* disrupted mutant *S. clavuligerus argR*::*aph*.

Transcription of Genes for the Early Steps of the Pathway

S1 mapping with the 464 nt CJ probe (Figure 3, lane 5) shows an hybridization band of 428 nt (Figure 3, lanes 2 and 4) with *S.clavuligerus argR*::*aph* RNA. This band, corresponding to the RNA fragment protected by the CJ probe, was found in cells grown in TSB medium (lane 4) and in GSPG medium (lane 2). S1 mapping experiments with total RNA from the wild type strain show a band of the same size but with weaker intensity (Figure 3A, lanes 1, 3 and insert). Quantification of the hybridization signal showed that the amount of RNA specific for *argCJ* is about 30 times higher in *S.clavuligerus argR*::*aph* than in the wild type strain. This result correlates well with the presence of an ArgR repressible Arg-box sequence upstream of *argC* (Rodríguez-García *et al.*, 1997).

S1 mapping with the JB probe (286 nt) showed a single weak band of 272 nt corresponding to the homologous region in the probe and the protected RNA (not shown). Probe DR (449 nt) protected a RNA fragment of 373 nt (Figure 3B, lanes 1 to 4) which indicates that *argD* and *argR* are co-transcribed. Additionally, a weak band of protection of 157-158 nt was found in the same S1 mapping experiment suggesting the presence of an additional promoter located upstream of *argR*. These results suggest that the most probable transcription initiation site of the *argR* promoter is at an adenine (nt 4594 in Figure 2) located immediately downstream of the TGA stop codon of *argD*.

Therefore, it seems that the early genes *argCJBDR* are transcribed in a single mRNA of about 5 kb; additionally *argR* is transcribed from its own promoter. To confirm these results Northern analysis of total RNA was made using as probe a PCR-amplified DNA fragment corresponding to the whole *argR* gene (Probe V). A large and diffuse hybridization band that might corresponds to the degraded *argCJBDR* transcript was observed; additionally a strong hybridization band of about 0.6 kb (Figure 4A) was found corresponding to the *argR* monocistronic transcript. The intensity of the hybridizing bands was higher in RNA preparations from cells grown for 48 and 72 h.

When the signal of the different S1 protection bands was quantified, the maximal intensity corresponds to the polycistronic *argCJBDR* RNA from *S. clavuligerus argR::aph* grown in GSPG (Figure 3, panel A, lane 2, and panel B, lane 2). By comparison with the intensity of these bands (considered as 100% in Table 1 for the derepressed mutant) the intensity of the hybridization band of the polycistronic RNA from the wild type strain grown in GSPG medium was 3,5% (using probe CJ) and 2,8% (using probe DR) indicating an increase of about 30 times in mRNA levels in the derepressed mutant. In cells grown in TSB medium the intensity of the polycistronic transcript for the wild type strain are 1.5% (probe CJ) and 0.76% (probe DR) (Table 1; Figure 3, panel A, lane 3 and panel B, lane 3) with respect to the levels of the derepressed mutant in the GSPG medium and the data for *S. clavuligerus argR*::*aph* amounts to 66.5% (probe CJ) and 30.9% (probe DR) (Table 1). Comparison of the data in GSPG and TSB media (Table 1) suggests that the complex TSB medium probably contains enough arginine to partially repress the transcription of arginine biosynthesis genes. The intensity of the signal corresponding to the monocistronic *argR* transcript is relatively similar in all the strains and growth conditions used. Therefore it appears that the polycistronic mRNA accounts for most of the ArgR repressor protein

Table 1. Relative Percentage of Expression of the Early and Late *arg* Genes as Shown by S1 Protection Studies

		Strain				Transcription Increase* (fold)	
		S. clavuligerus 27064		*S. clavuligerus argR*::*aph*			
		TSB	GSPG	TSB	GSPG	TSB	GSPG
Probe CJ	mRNA *arg*CJBDR	1.5	3.5	66.5	100.0[1]	44.3	28.5
Probe DR	mRNA *argCJBDR*	0.7	2.8	30.9	100.0[2]	44.1	35.7
	mRNA *argR*	0.2	0.5	0.5	0.4	2.5	0.8
Probe GH	mRNA *argGH*	3.0	10.5	91.7	100.0[3]	30.5	9.5

[1] 60,959 net counts
[2] 906,824 net counts
[3] 36,095 net counts
* The expression increase is the ratio of mRNA expression of *S. clavuligerus arg*R::*aph* in relation to the wild type strain

present in the cells. The results of transcription initiation showed in Figure 2 for *argR* and comparison with the homologous *S. coelicolor argR* sequence suggest that the translation initiation of ArgR is the ATG present at nt 4,594 (Figure 2).

Transcription of Genes for the Late Steps of the Pathway
Upstream from the coding sequence of *argG* a putative promoter region formed by a -35 TTGACG sequence separated by 17 nt from the -10 CATACT box was predicted (Rodríguez-García *et al*., 1995). S1 analysis using the *argRG* probe showed a 263-264 nt protection band, that confirms that *argG* transcription starts at a A or a C located 29 or 30 nt upstream of the GTG translation initiation codon (Figure 2C) in agreement with the initially proposed promoter. Protection with the *argGH* probe showed a single hybridization band of 550 bp. The hybridization bands were stronger in the *argR*- disrupted mutant (Figure 3A, lanes 2, 4 and Figure 3C, lanes 2, 4) than in the wild type strain. Northern hybridization of total RNA from the wild type *S. clavuligerus* with a 0.83 kb *Kpn*I-*Nru*I DNA fragment (probe VI in Figure 2A) internal to *argG* showed a strong signal of about 1.5 kb which might correspond to a monocistronic *argG* transcript. Additionally, a degraded transcript of higher molecular weight was observed (Figure 4B). The large transcript was also observed when a 0.4 kb *Pvu*II probe internal to *argH* (probe VII) was used as probe, suggesting that both genes are co-transcribed.

Discussion

The presence of a cluster of seven contiguous genes for the biosynthesis of arginine in *S. clavuligerus* was unexpected since *arg* auxotrophic mutants map in four different loci in the *S. coelicolor* genome (Redenbach *et al*., 1996).

The hybridization studies of the *S. coelicolor* DNA with our probes suggest that there are at least two clusters of *arg* genes in this model actinomycete. One cluster containing *argC,J,D,R* and *B* while *argH* is located close but not linked to the first subcluster. These results agree with those of Hindle *et al*. (1994) who found an *argCJB* cluster in *S. coelicolor* and with recent data on the genetic map of *S. coelicolor* (Sanger Center Sequencing Group).

In addition, the unstable *argG* of *S. coelicolor* maps near one of the ends of *the S. coelicolor* linear chromosome (Redenbach *et al*., 1996) at difference of what occurs *in S. clavuligerus* where *argG* forms part of the arginine cluster as reported in this article. Unstable *argG* genes are also present in *Streptomyces lavendulae, Streptomyces lividans, Streptomyces cattleya, Streptomyces scabies* and *Streptomyces alboniger.* There are two types of *argG* genes in *Streptomyces* (Rodríguez-García *et al*., 1995) and it is likely that the stable *argG* genes of *Streptomyces* species might be located in a large *arg* cluster whereas the unstable *argG* gene is located in a separate location near the end of the linear genome. This would suggest that *argG* suffered a translocation in a line of *Streptomyces* ancestors but not in other group of this large genus.

The absence of an *argF* gene, encoding ornithine carbamoyltransferase (forming citrulline) in the *arg* cluster of *S. clavuligerus* is surprising; *argF* is not located in the vicinity, either upstream or downstream of the *S. clavuligerus arg* cluster although *argF* is in a central position in the *arg* cluster of mycobacteria. Attempts to clone this gene by complementation of *argF* mutants of *E. coli* were unsuccessful, perhaps due to the lack of expression of its promoter in *E. coli*.

Regulation of the expression of the arginine cluster in *S. clavuligerus* by the arginine level is exerted at the transcription level. Two canonical 18-nt Arg-boxes with 70 to 80% similarity to the consensus *Streptomyces* Arg-box (Rodríguez-García et al., 1997) occur upstream of the *argC* and *argG* genes but only one 18-nt sequence with 50% similarity to Arg-boxes, is present upstream of *argR* (nucleotides 4581-4598). Indeed formation of the monocistronic *argR* transcript appears to be weak and constitutive (Table 1). This is in agreement with the results of arginine regulation in *E. coli* in which the presence of two arginine-boxes exerts a cooperative effect resulting in a stronger binding of the regulatory ArgR protein. Binding of the *B. subtilis* AhrC arginine repressor to the *arg*-boxes of *argC* has been reported previously (Rodríguez-García *et al*., 1997). Expression of the arginine biosynthesis genes in the arginine R-disrupted mutant is in the order of 10 to 44-fold that of the wild type strain (depending on the medium) (Table I). Arg-boxes are also present upstream of *argC* (Hindle *et al*.,1994), *argG* (Rodríguez-García *et al*., 1995) and *argH* (Sanger Center Sequencing Group) of *S. coelicolor* suggesting a similar type of regulation.

Our results are compatible with the co-transcription of the *argG* and *argH* genes. However the differences in intensity between the 3.3 kb and the 1.5 kb band obtained by Northern hybridization with the *argG* probe suggest that

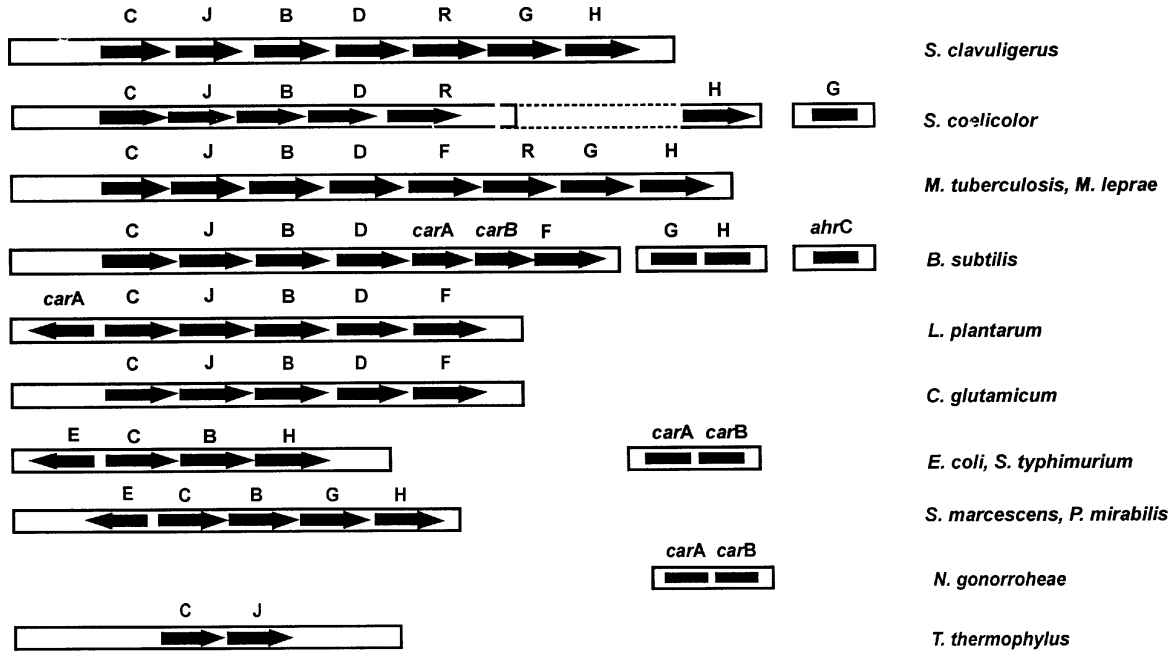

FIgure 5. Organization of arginine biosynthesis gene clusters in different Gram positive and Gram negative bacteria and Archaebacteria. Note the different arrangement of the *arg* genes in *S. clavuligerus* and *S. coelicolor* (*S. coelicolor* StL24 cosmid sequence at the Sanger Center).

the inverted repeat sequence existing 22 nt downstream of the *argG* TGA stop codon, which might form a stem and loop structure (Rodríguez-García *et al.*, 1995), acts as a terminator (attenuator) and therefore only a small percentage of the RNA polymerase-iniciated transcript are able to readthrough to the *argH* gene. This is an interesting example of differential expression of clustered genes in *Streptomyces*, that seems to occur also in long polycistronic transcript of antibiotic biosynthesis genes (Enguita *et al.,* 1998).

Experimental Procedures

Bacterial Strains and Plasmids

The wild type S**.** *clavuligerus* ATCC 27064 was used as source of DNA. *E. coli* auxotroph strains XSlD2 *(argE)* and XB25 *(argB)* were kindly provided by S. Baumberg (Leeds University, UK). Strains *E. coli* CGSC 5359 (*argH*) and *E. coli* CGSC 4538 *(argD)* were obtained from the *E. coli* Genetic Stock Center.

Plasmids pULML31, pULAR1 and pULAR11 were constructed by Ludovice *et al.* (1992) and Rodríguez-García *et al.* (1995). Plasmid pULD contains in pBSK(+) a 1.38 kb *Kpn*I DNA fragment carrying the *argD* gene. Plasmid pULB contains the *argB* gene obtained by PCR in the *Eco*RV site of pULMA (Rodríguez-García *et al.,* 1995), donwstream of the *argC* promoter. Plasmids pULJ1 and pULJ2 are clones used for the sequencing of the 6.1 kb *Sau*3A1 DNA fragment; they contain in pBSK(+) a 1.2 kb DNA insert containing the complete *argJ* gene.

Culture Conditions

E. coli auxotrophy complementation was tested in VB medium (Smith and Yanofsky, 1962). *Streptomyces* strains were grown from mycelium stock kept in glycerol (20%) at -75ºC. One milliliter of the frozen mycelium was used to seed a 100 ml TSB (3% trypticaseine-soy) culture. The TSB culture was grown for 36 h, the optical density was adjusted to O.D. 5.0 and 5 ml were centrifuged and used to inoculate 500 ml trippled baffled flasks containing 100 ml of either TSB or GSPG medium (containing glutamate, proline, glycerol and sucrose) (Romero *et al.*, 1984).

Enzyme Assays

N-acetylglutamate synthetase (EC 2.3.1.1.), ornithine N-acetyltransferase (EC 2.3.1.35), N-acetylornithine aminotransferase (EC 2.6.1.11) and N-acetylornithinase (EC 3.5.1.16) activities were assayed in different *E. coli* transformants according to the methods of Haas *et al.* (1972), Denés (1970), Vogel and Jones (1970) and Vogel and McLellan (1970) respectively. The enzyme activities encoded by genes expressed from the *lacZ* promoter were measured in *E. coli* transformants grown for 8 h in the presence of 0.1 mM IPTG.

DNA Manipulation and Transformation

Plasmid pHZ*argR*-K (a pHZ1351-derived vector) contains the *aphII* gene inserted as a blunt-ended fragment into the *Bst*XI site of the *argR* gene. *S. clavuligerus* was transformed with pHZ*arg*R-K to obtain the disrupted mutant *S. clavuligerus argR*::*aph*. *E. coli* strains were transformed according to Hanahan (1983). All DNA manipulations including PCR studies and Northern hybridization were made following standard procedures (Sambrook *et al.*,1989; Hopwood *et al.*, 1985). S1 mapping was performed essentially as described by Sambrook *et al.* (1989). Total RNA from *S. clavuligerus* ATCC 27064 (200 μg) or S. *clavuligerus argR::aph* (100 μg) was used and 100-200 μg tRNA from *Saccharomyces cerevisiae* (type X-SA, Sigma) was utilized as negative control. The hybridization reaction was done at 67ºC overnight. The S1 reaction sample (4 μl DNA sample plus 6 μl carrier buffer) was electrophoresed in 6% urea 7M. As size control the sequencing reactions from M13mp18 were used. Quantification of the hybridization signals was done by densitometry using an Electronic Autoradiography Instant Imager (Packard Ins, Meriden, USA).

RNA Isolation

RNA in the mycelium from 100 ml of a 24 h (TSB medium) or 48 h culture (GSPG medium) was extracted as described before (Pérez-Redondo *et al.*, 1998) except that prior to the RNA purification, most of the DNA was specifically precipitated from the aqueouse phase using isopropanol (v/v).

Probes Used in the S1 Analysis

Probe CJ: A 626 nucleotide fragment was amplified by PCR using as template the *Sma*I-*Sal*I DNA fragment corresponding to *S. clavuligerus argCJ* intergenic region and the oligonucleotides 1J**:** 5' -GGGCTGGC CGCTCTCCTT-3' and reverse M13mp18(-20) as primers. The 626 nt amplified region was treated with *Bst*XI to obtain a 464 nt fragment with a protuding 5' end, non-susceptible to labelling with polynucleotide kinase. This probe has 36 nt of heterologous DNA [from pBSKS (+)].

Probe JB: The oligonucleotides 1B: 5'-CTTGGGGAGGGCGTTGTTCTT-3' and reverse M13mp18 were used to amplify by PCR a DNA template (*Sal*I-*Nco*I 349 nt DNA fragment corresponding to the *S. clavuligerus argJB* intergenic region) subcloned in pBSKS(+). The 354 nt amplified region was digested with *Kpn*I giving a 290 nt probe that contains 21 nt of heterologous DNA [corresponding to pBSKS (+)].

Probe DR: An oligonucleotide 1R: 5´-ACGCTCAGTCCGTTGTCCG-3' corresponding to the non-coding strand of *argR* was purified by PAGE and 10 pmol were labelled at its 5' end with [γ-^{32}P] ATP (>185 TBq/mmol, Amersham Ltd., England) and 8 units of T4 polynucleotide kinase (MBI Fermentas, Lithuania). The labelled oligonucleotide was purified by precipitation with ammonium acetate using glycogen (20 μg) as carrier. Using this oligonucleotide and M13mp18 (-20) as primers a labelled DNA probe of 449 nt corresponding to *S. clavuligerus argDR* intergenic region was amplified by PCR. The 76 nt corresponding to the 3'-end of the non-labelled chain correspond to pBSSK(+).
Probe RG: The oligonucleotide 1R: 5' -*GTTGCTGAATTC*GCGGAGCTGC ACGACATC-3' , (in which the nt in italics correspond to pBSSK(+) non homologue DNA) and 2G: 5'-CTTGATCGCAGGGAGGCAGTA-3' were used to amplify from plasmid pULAR10 (Rodríguez-García *et al*., 1995) a 649 nt DNA fragment that was digested with *Eco*RI.
ProbeGH: A 561 nt DNA probe with 11 nt non-homologous nucleotides at the 5'-end of the non-labelled strand was amplify by PCR using the oligonucleotides 1G: 5'-GTTGCTGAATTCGGCTGGTCGGCATCAAGT-3' and 2H: 5'- CGAACCGT CCGCCCCAGAG-3'.

All the probes were purified from 1.5% agarose gels using the Quiaex Gel Extraction Kit (Quiagen, Germany) and denatured (except probe DR), by incubation for 5 min at 65ºC in the presence of NaOH (final concentration 0.1 M). The probes were precipitated with sodium acetate pH 5.2:ethanol in the presence of glucogen (10 μg) as carrier, suspended in reaction buffer (Sambrook *et al*., 1989) and labelled with [γ-^{32}P] ATP and 14 units of T4 polynucleotide kinase. After 50 min at 37ºC the enzyme was inactivated by heating and the probes were purified by precipitation with ammonium acetate:ethanol. Additionally, probes I to VII (Figure 2) were used for low resolution S1 mapping, Northern analysis or hybridization with *S. coelicolor* genome.

Acknowledgements

This work was supported by grants from the CICYT (Madrid) Bio96-0827 and by Antibióticos SA, León. We thank H. Kieser (Norwich, U.K.) for providing the *S. coelicolor* cosmid library. Álvaro de la Fuente and Rosario Pérez-Redondo received fellowships from the PFPI (Madrid) and the University of León, respectively.

References

Bringel, F., Frey, L., Boivin, S., and Hubert, J.C. 1997. Arginine biosynthesis and regulation in *Lactobacillus plantarum*: the *carAB* gene and the *argCJBDF* cluster are divergently transcribed. J. Bacteriol. 179: 2697-2706.

Crabeel, M., Charlier, D., Cunin, R., and Glansdorff, N. 1979. Cloning and endonuclease restriction analysis of *argF* and of the control region of the *argECBH* bipolar operon in *Escherichia coli*. Gene 5: 207-231.

Denés, G. 1970. Ornithine Acetyltransferase. Methods Enzymol. 17A: 273-277.

Eastcott, M., Griffin, A., Barton, B., and Baumberg, S. 1998. Studies on an *argJ* homologue located within the clavulanic acid gene cluster in *Streptomyces clavuligerus*. Abstract of the Genetics of Industrial Microoganisms (GIM 98) Congress, Jerusalem, Israel, p. 88

Enguita, F.J., Coque, J.J.R., Liras, P., and Martín, J.F. 1998. The nine genes of the *Nocardia lactamdurans* cephamycin cluster are transcribed into large mRNAs from three promoters, two of them located in a bidirectional promoter region. J. Bacteriol. 180: 5489-5494.

Haas, D., Holloway, B.W., Schamböck, A., and Leisinger, T. 1977. The genetic organization of arginine biosynthesis in *Pseudomonas aeruginosa.* Mol. Gen. Genet. 154: 7-22.

Hanahan, D. 1983. Studies of transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166: 557-580.

Hass, D., Kurer, V., and Leisinger, T. 1972. N-Acetylglutamate synthetase of *Pseudomonas aeruginosa.* An assay *in vitro* and feedback inhibition by arginine. Eur. J. Biochem. 31: 290-295.

Higgins, D.G., and Sharp, P.M. 1989. Fast and sensitive multiple sequence alignement on a microcomputer. Comp. Appl. Biosci. 5: 151-153.

Hindle, Z., Callis, R., Dowden, S., Rudd, B.A.M., and Baumberg, S. 1994. Cloning and expression in *Escherichia coli* of a *Streptomyces coelicolor* A3(2) *argCJB* gene cluster. Microbiology 140: 311-320.

Hodgson, J.E., Fosberry, A.P., Rawlinson, N.S., Ross, H.N.M., Neal, R.J., Arnell, J.C., Earl, A.J., and Lawlor, E.J. 1995. Clavulanic acid biosynthesis in *Streptomyces clavuligerus*: gene cloning and characterization. Gene 166: 49-55.

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate, D.J., Smith, C.P., Ward, J.M., and Schrempf**,** H. (1985) Genetic manipulation of *Streptomyces*: a laboratory manual. The John Innes Foundation. Norwich, U.K.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Danchin, A., *et al*. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature 390: 249-256.

Leskiw, B.K., Lawlor, E.J., Fernández-Ábalos, J.M. and Chater, K.F. 1991. TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proc. Natl. Acad. Sci. USA 88: 2461-2465.

Ludovice, M., Martín, J.F., Carrachás, P., and Liras, P. 1992. Characterization of the *Streptomyces clavuligerus argC* gene encoding N-acetyl glutamyl-phosphate reductase: Expression in *Streptomyces lividans* and effect on clavulanic acid production. J. Bacteriol. 174: 4606-4613.

Matsumoto, H., Hosogaya, S., Suzuki, K., and Tazaki, T. 1975. Arginine gene cluster of *Serratia marcescens*. Jpn. J. Microbiol. 19: 35-44.

Mountain, A., Mann, N.H., Munton, R.N., and Baumberg, S. 1984. Cloning of a *Bacillus subtilis* restriction fragment complementing auxotrophic mutants of eight *Escherichia coli* genes of arginine biosynthesis. Mol. Gen. Genet. 197: 82-89.

North, A.K., Smith M.C.M., and Baumberg, S. 1989. Nucleotide sequence of the *Bacillus subtilis* arginine regulatory gene and homology of its products to the *E. coli* arginine repressor. Gene 80: 29-38.

O'Reilly, M., and Devine, K.M. 1994. Sequence and analysis of the citrulline biosynthetic operon *argC-F* from *Bacillus subtilis*. Microbiology 140: 1023-1025.

Pérez-Redondo, R., Rodríguez-García, A., Martín J.F., and Liras, P. 1998. The *claR* gene of *Streptomyces clavuligerus*, encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase (*car*) gene. Gene 211: 311-321.

Picard, F.J., and Dillon, J.R. 1989. Cloning and organization of seven arginine biosynthetic genes from *Neisseria gonorrhoeae.* J. Bacteriol. 171: 1644-1651.

Piggot P.J., and Hoch, J.A. (1985) Revised genetic linkage map of *Bacillus subtilis*. Microbiol. Rev. 49: 158-179.

Prozesky, O.W. 1968. Transductional analysis of arginine less mutants in *Proteus mirabilis.* J. Gen. Microbiol. 54: 127-143.

Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H., and 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.

Rodríguez-García, A., Ludovice, M., Martín, J.F., and Liras, P. 1997. Arginine boxes and the *argR* gene of *Streptomyces clavuligerus*: evidence for a clear regulation of the arginine pathway. Mol. Microbiol. 25: 219-228.

Rodríguez-García, A., Martín, J.F., and Liras, P. 1995. The *argG* gene of *Streptomyces clavuligerus* has low homology to unstable *argG* from other actinomycetes: Effect of amplification on clavulanic acid biosynthesis. Gene 167: 9-15.

Romero, J., Liras, P., and Martín, J.F. 1984. Dissociation of cephamycin and clavulanic acid biosythesis in *Streptomyces clavuligerus.* Appl. Microbiol. Biotechnol. 20: 318-325.

Sakanyan, V., Petrosyan, P., Lecocq, M., Boyen, A., Legrain, C., Demarez, M., Hallet, J.N., and Glansdorff, N. 1996. Genes and enzymes of the acetyl cycle of arginine biosynthesis in *Corynebacterium glutamicum*: enzyme evolution in the early steps of the arginine pathway. Microbiology 142: 99-108.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sanderson, K.E. 1970. Current linkage map of *Salmonella typhimurium***.** Bacteriol. Rev. 34: 176-193.

Smith, O.H., and Yanofsky, C. 1962. Biosynthesis of tryptophan. Methods Enzymol. 5: 794-806.

Strohl, W.R. 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. Nucleic Acid Res. 20: 961-974.

Valentine, B.B., Bailey, C.R., Doherty, A., Morris, J., Elson, S.W., Baggaley, K.H. and Nicholson**,** N.H. 1993. Evidence that arginine is a later metabolic intermediate than ornithine in the biosynthesis of clavulanic acid by *Streptomyces clavuligerus.* J. Chem. Soc. Chem. Commun. 1993: 1210-1212.

Vogel, H.J., and Jones, E.E. 1970. Acetylornithine δ-aminotransferase (*Escherichia coli*). Methods Enzymol. 17A: 260-264.

Vogel, H.J., and McLellan, W.L. 1970. Acetylornithinase. Methods Enzymol. 17A: 265-269.

Woods, S.A., Miles J.S., and Guest, J.R. 1988. Sequence homologies between argininosuccinase, aspartase and fumarase: a family of structurally-related enzymes. FEMS Microbiol. Lett. 51: 181-186.

Yonaha, K., Nishie, M., and Aibara, S. 1992. The primary structure of α-amino acid:pyruvate aminotransferase. J. Biol. Chem. 267: 12506-12510.