

Métabolisme – Régulation

La régulation génétique dans *Lactococcus lactis* : le modèle de la biosynthèse du tryptophane

A Chopin, J Bardowski, R Raya, SD Ehrlich

Laboratoire de génétique microbienne, INRA, 78352 Jouy-en-Josas, France

Résumé — Pour contrôler la fermentation lactique, il est nécessaire de contrôler l'expression des gènes des bactéries lactiques. Pour cela, il faut mieux connaître les modes de régulation génétique de ces bactéries. La régulation de la biosynthèse du tryptophane dans *Lactococcus lactis* a été étudiée. Les 7 gènes nécessaires à la synthèse du tryptophane ont été clonés et caractérisés. Ils sont présents dans l'ordre *trpEGDCFBA*, occupent une région de 7 968 paires de bases et sont transcrits en un ARN messager unique d'environ 8 kb, indiquant une organisation en opéron. La synthèse de ce messager est modulée d'un facteur 150 par la présence de tryptophane dans le milieu de culture. Le promoteur de transcription a été identifié. De petits ARN messagers sont également produits à partir de ce promoteur. Cependant, leur synthèse n'est pas modulée par le Trp. Le gène *lacZ* de *Escherichia coli* qui code pour une β -galactosidase a été introduit dans le chromosome de *L. lactis* sous le contrôle des signaux d'expression de l'opéron *trp*. Dans cette souche, la synthèse de β -galactosidase est modulée par la présence de tryptophane.

***Lactococcus lactis* / tryptophane / biosynthèse / régulation / gène**

Summary — Gene regulation in *Lactococcus lactis*: the example of tryptophan biosynthesis. To obtain more insight into the mechanisms which control gene expression in *Lactococcus lactis*, we have undertaken studies on the regulation of the tryptophan (*Trp*) biosynthetic genes. This model was chosen because the *trp* genes and their products have been characterized in many prokaryotes and represent a paradigm for the study of gene organization and regulation. The chromosomal *trp* region from *L. lactis* subsp. *lactis* IL1403 was cloned and sequenced. The 7 structural genes are present in the *trpEGDCFBA* order and span a 7 968-bp segment. Each gene is preceded by a putative ribosome-binding site. Upstream of *trpE*, a 457-bp non-coding DNA segment contains several regions fitting the consensus for Gram-positive promoters and one region strongly resembling a transcription terminator. Northern hybridisation experiments revealed a 8-kb mRNA transcript, indicating that *trp* genes are organised in a single operon. The biosynthesis of this transcript was modulated hundred-fold by *Trp*. Transcription start site was determined by primer extension analysis. Our present data suggest that *Trp* biosynthesis in *L. lactis* is regulated by modulation of transcription termination. The *lacZ* gene was introduced into the *L. lactis* chromosome under the control of the *trp* expression signals. In this strain, the expression of the β -galactosidase is modulated by the *Trp* concentration.

***Lactococcus lactis* / tryptophan / biosynthesis / regulation / gene**

INTRODUCTION

Toutes les protéines ne sont pas également abondantes dans la cellule bactérienne. Dans une cellule d'*Escherichia coli* par exemple, le nombre de molécules de certaines protéines rares est de l'ordre de la dizaine alors que pour des protéines abondantes, ce nombre est de l'ordre de la centaine de milliers (Gouy et Gautier, 1982). Cette différence s'explique principalement par une différence dans la fréquence d'initiation de la transcription des gènes : tandis que certains gènes sont transcrits moins d'une fois par génération, d'autres sont transcrits jusqu'à une fois par seconde (Hoopes et McClure, 1987). En outre, l'expression de la plupart des protéines bactériennes est régulée en fonction des conditions de vie de la cellule. Par exemple, la concentration de β -galactosidase dans les cellules d'*E coli* est environ 1 000 fois plus élevée en présence qu'en absence de lactose (Beckwith, 1987). Ces phénomènes, qui existent dans toutes les bactéries étudiées, coordonnent l'ensemble de leur activité physiologique et il est important de les connaître pour pouvoir orienter et modifier la physiologie bactérienne.

Cette connaissance serait également importante dans le domaine des bactéries lactiques. Il est en effet crucial de pouvoir mieux contrôler les gènes industriels, de façon à obtenir leur expression au niveau optimum, au moment voulu du processus de fabrication. Cependant, les mécanismes qui contrôlent l'expression des gènes des bactéries lactiques n'ont été que peu étudiés, et cela nous a conduits à entreprendre une recherche approfondie sur la régulation génétique dans *Lactococcus lactis*.

Nous avons choisi comme modèle d'étude le système de biosynthèse du tryptophane (Trp) qui offre plusieurs avan-

tages. Les gènes de la voie de biosynthèse du Trp (fig 1), leurs produits et leur régulation ont déjà été caractérisés dans de nombreux organismes (Crawford, 1989) et cela devrait faciliter leur étude dans *L. lactis*. Par ailleurs, le Trp est l'acide aminé dont la biosynthèse est la plus coûteuse pour la cellule (78 mol d'ATP sont nécessaires pour la synthèse d'une mole de Trp) (Somerville, 1983). Comme les bactéries lactiques sont énergétiquement limitées, il est probable que l'expression de ces gènes de biosynthèse soit étroitement contrôlée par des mécanismes de régulation efficaces.

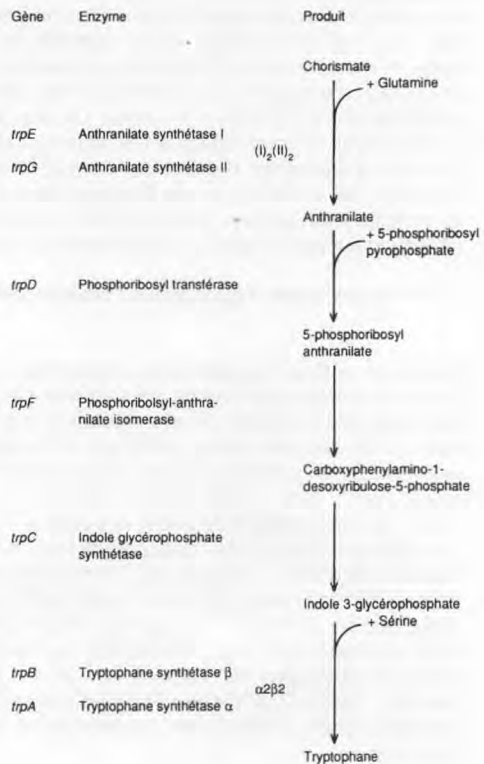


Fig 1. Gènes et enzymes spécifiques de la voie de biosynthèse du tryptophane (d'après Somerville, 1983).

LES GÈNES DE BIOSYNTÈSE DU TRYPTOPHANE DE *LACTOCOCCUS LACTIS* SUBSP *LACTIS*

Clonage de la région *trp*

La région du chromosome de *L. lactis* subsp *lactis* contenant les gènes *trp*, a été clonée dans un mutant Trp^- de *Bacillus subtilis*. Un transformant Trp^+ a été obtenu dans lequel avait été cloné un fragment d'ADN chromosomique de *L. lactis* subsp *lactis* IL1403 de 12,7 kilobases (kb) (Bardowski *et al*, 1992). L'analyse de ce fragment par complémentation de différentes mutations *trp* de *E. coli* a révélé qu'il complémentait les mutations *trpA*, *trpB*, *trpC* et *trpD*, mais non *trpE*. Le clonage de la région adjacente du chromosome de *L. lactis* a permis d'isoler un fragment d'ADN de 3,6 kb, chevauchant partiellement le précédent, et capable de complémenter la mutation *trpE* de *E. coli*.

Identification des séquences codantes et des enzymes de la voie de biosynthèse

Les gènes *trp* et leur organisation ont été caractérisés en détail par séquençage des fragments d'ADN clones. La séquence nucléotidique d'une région de 8,656 kb a été déterminée, dans laquelle 8 séquences codantes probables ont été identifiées (Bardowski *et al*, 1992). Ces séquences codantes sont précédées, 5 à 7 paires de bases en amont, par des séquences homologues à la séquence de fixation des ribosomes de *L. lactis* (de Vos, 1987). Les séquences codantes potentielles ont été traduites en protéines qui ont été comparées aux protéines de séquences connues. Cela a permis d'établir que 7 de ces 8 séquences codaient pour des protéines ho-

mologues aux protéines de biosynthèse du Trp connues chez d'autres organismes, et d'identifier les gènes *trp* de *L. lactis* subsp *lactis*. La protéine déduite de la huitième séquence codante potentielle ne présentait d'homologie significative avec aucune séquence protéique connue et a été désignée OrfX. La position des gènes *trp* de *L. lactis* subsp *lactis* dans la région étudiée est représentée sur la figure 2.

Les protéines Trp de *L. lactis* subsp *lactis* présentent de 20 à 61% d'homologie par rapport aux protéines Trp d'autres organismes (tableau I). La plus faible homologie a été observée dans le cas de TrpF (20 à 34%) et la plus forte dans le cas de TrpB (50 à 61%). L'homologie des protéines Trp n'est pas significativement supérieure entre bactéries phylogénétiquement proches (telles que *L. lactis* subsp *lactis* et *Bacillus subtilis* ou *Lactobacillus casei*) qu'entre organismes phylogénétiquement très éloignés (tels que *L. lactis* subsp *lactis* et *Saccharomyces cerevisiae* ou *Methanobacterium thermoautotrophicum*). Cela est en accord avec le modèle proposé par

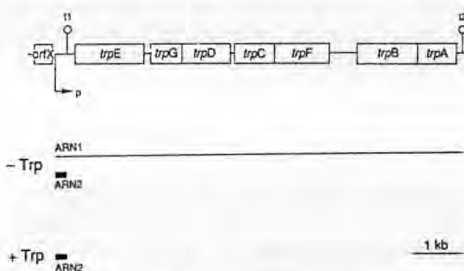


Fig 2. Structure et transcription de la région d'ADN de *L. lactis* subsp *lactis* IL1403 contenant les gènes *trp*. Les gènes sont représentés par des rectangles et les régions d'ADN non codantes sont représentées par une ligne. Dans la partie inférieure sont représentés les ARN messagers transcrits à partir de l'ADN en présence et en absence de tryptophane. L'épaisseur du trait indique une différence dans l'abondance relative des ARN messagers. p : promoteur de transcription; t1 et t2 : terminateurs de transcription.

Tableau I. Homologie de la séquence en acides aminés des enzymes de biosynthèse du tryptophane de *Lactococcus lactis* subsp *lactis* avec celles d'autres organismes.

Organismes comparés	Identité de séquence (%) ^a							Références
	TrpE	TrpG	TrpD	TrpC	TrpF	TrpB	TrpA	
<i>Lactobacillus casei</i>	– ^b	–	38	38	29	56	32	Natori <i>et al</i> (1990)
<i>Bacillus subtilis</i>	40	44	32	38	24	61	36	Band <i>et al</i> (1984); Henner <i>et al</i> (1984)
<i>Brevibacterium lactofermentum</i>	33	34	35	32	25	50	25	Matsui <i>et al</i> (1986)
<i>Escherichia coli</i>	36	36	35	33	26	54	28	Yanofsky <i>et al</i> (1981)
<i>Pseudomonas aeruginosa</i>	43	38	36	36	–	59	30	Essar <i>et al</i> (1990); Hadero <i>et Crawford</i> (1986)
<i>Methanobacterium thermoautotrophicum</i>	36	40	35	33	20	55	28	Meile <i>et al</i> (1991)
<i>Saccharomyces cerevisiae</i>	34	36	25	36	34	55	33	Furter <i>et al</i> (1986); Tschumper <i>et Carbon</i> (1980); Zalkin <i>et Yanofsky</i> (1982); Zalkin <i>et al</i> (1984)

^a Cette valeur a été calculée à partir des alignements multiples de séquence obtenus par la méthode de Corpet (1988), en divisant le nombre d'acides aminés conservés à la même position par le nombre d'acides aminés dans la séquence protéique la plus courte; ^b séquence non publiée.

Crawford (1989) pour l'évolution des protéines *Trp*, selon lequel les régions hautement conservées des protéines sont celles qui sont essentielles pour les fonctions enzymatiques et sont entourées de régions variables dans lesquelles les substitutions d'acides aminés sont plus fréquentes.

Structure et organisation des gènes *trp*

Les 7 gènes *trp* nécessaires à la biosynthèse du Trp sont regroupés sur le chromosome de *L. lactis* subsp *lactis*. C'est

également le cas dans d'autres eubactéries, comme *Escherichia coli* (Yanofsky *et al*, 1981), *Brevibacterium lactofermentum* (Matsui *et al*, 1986) ou, à un gène près, dans *Bacillus subtilis* (Henner *et al*, 1984; Slock *et al*, 1990) ainsi que dans l'archaéobactérie *Methanobacterium thermoautotrophicum* (Meile *et al*, 1991). En revanche, dans *Pseudomonas*, les gènes *trp* sont répartis dans 4 endroits différents du chromosome (Crawford *et Gunsalus*, 1966; Calhoun *et al*, 1973).

Dans *L. lactis* subsp *lactis*, les 7 gènes *trp* recouvrent une région de 7,968 kb et

sont présents dans l'ordre *trpEGDCFBA* qui est le plus fréquent chez les bactéries. Tous les gènes *trp* utilisent AUG comme codon d'initiation de traduction, à l'exception de *trpG*, qui commence par un codon UUG. *trpG*, *trpF* et *trpB* se terminent par le codon UGA et les 4 autres gènes *trp* se terminent par le codon UAA. Chaque gène est précédé d'une séquence riche en purines, contenant un site potentiel de fixation au ribosome. Trois paires de gènes, *trpG-trpD*, *trpC-trpF* et *trpB-trpA*, se chevauchent de 11, 11 et 4 paires de bases respectivement. Trois autres paires de gènes, *trpE-trpG*, *trpD-trpC* et *trpF-trpB*, sont séparées par des espaces intercistroniques de 124, 46 et 585 paires de bases respectivement. Il n'a pas été observé de fusion entre les gènes *trp* de *L. lactis* subsp. *lactis*.

RÉGULATION DE L'EXPRESSION DES GÈNES *trp*

L'expression des gènes trp est régulée par la présence de tryptophane

L'observation que les gènes *trp* sont rassemblés dans une région unique du chromosome de *L. lactis* subsp. *lactis* suggère une expression et une régulation coordonnées de ces gènes. Les ARN messagers transcrits à partir des gènes *trp* ont été analysés dans des expériences d'hybridation par Northern (Raya, en préparation). Les résultats sont schématisés sur la figure 2. En présence de Trp dans le milieu de culture, de petits ARN messagers sont abondamment transcrits. En revanche, en absence de Trp, on observe en plus la transcription d'un ARN messenger de 8 kb environ, représentant la totalité des gènes *trp*. Ces résultats indiquent que la transcription des gènes *trp* est contrôlée par la présence de Trp dans le milieu. Le point

de départ de la transcription a été déterminé par cartographie à la transcriptase inverse ce qui a permis d'identifier le promoteur de transcription (Raya, en préparation) qui présente les caractéristiques typiques des promoteurs de transcription des lactocoques (de Vos, 1987).

Les points d'arrêts de transcription observés en amont de *trpE* et en aval de *trpA* coïncident avec des séquences caractéristiques des terminateurs de transcription bactériens rho-indépendants (d'Aubenton Carafa *et al*, 1990). Ces terminateurs de transcription potentiels sont désignés *t1* et *t2* (fig 2).

L'ensemble de ces résultats indique que la biosynthèse du Trp dans *L. lactis* subsp. *lactis* est régulée, au moins en partie, au niveau de la transcription et que les gènes *trp* qui sont transcrits sous forme d'un ARN messenger unique constituent un opéron.

Construction et régulation d'une fusion trpE::lacZ

Pour mieux caractériser la régulation de l'expression des gènes *trp*, nous avons mesuré le niveau d'expression de l'opéron *trp* dans des conditions inductrices et non inductrices. Pour cela, un gène hybride formé de la fusion du début du gène *trpE* avec le gène *lacZ* de *E. coli* qui code pour la β -galactosidase, précédé en amont par la région leader (*reg*) de l'opéron *trp* a été intégré dans le chromosome de *L. lactis* subsp. *lactis*, en amont des gènes *trp*. La méthode utilisée pour cette construction a déjà été décrite (Chopin *et al*, 1989) et le détail de la construction (Bardowski, en préparation) est donné sur la figure 3.

La production de β -galactosidase par cette souche a été mesurée dans un milieu de culture synthétique contenant des concentrations croissantes en Trp. Les résultats obtenus montrent que l'expression

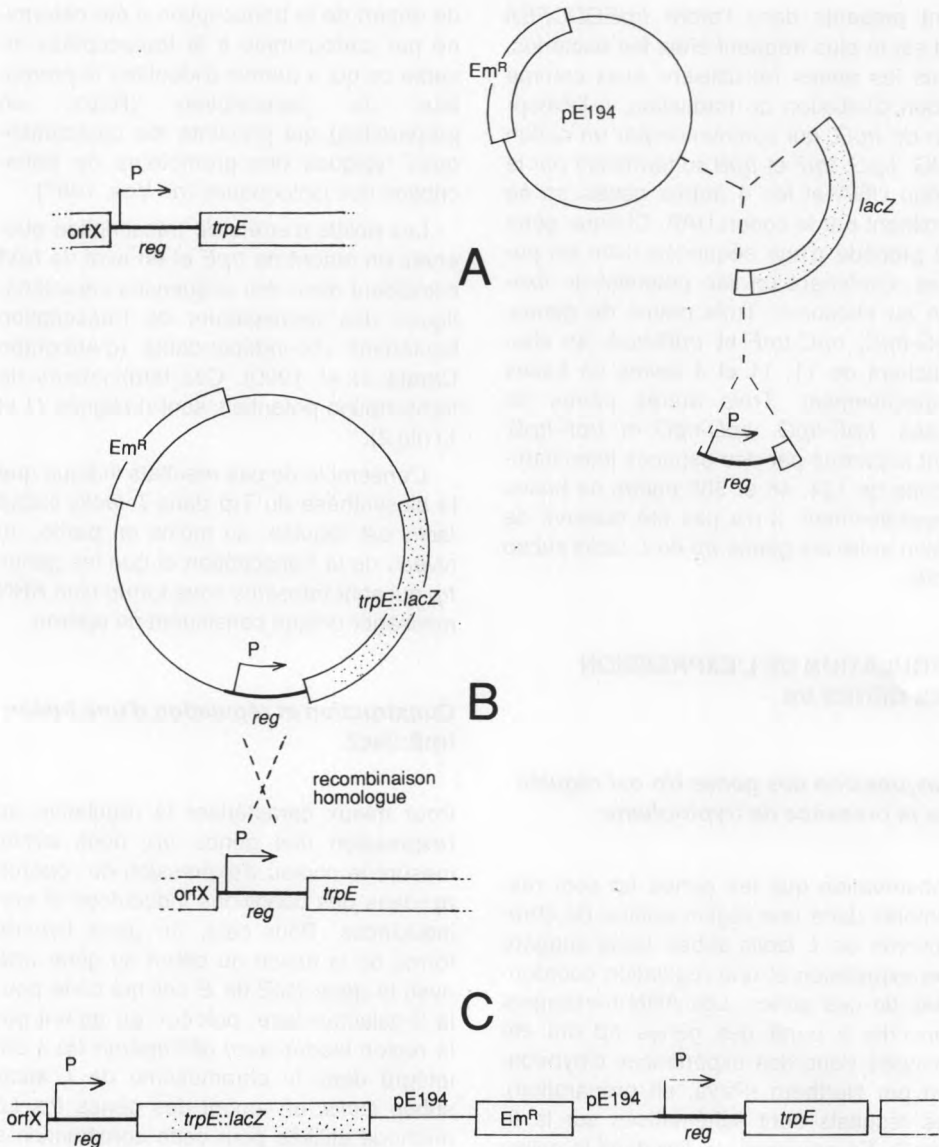


Fig 3. Intégration d'un gène hybride *trpE::lacZ* dans la chromosome de *L. lactis* subsp *lactis* IL1403. La méthode utilise la particularité du plasmide *pE194* à se répliquer dans *Bacillus subtilis*, mais non dans *Lactococcus lactis* (Chopin *et al*, 1989). **A.** Un fragment d'ADN contenant la région leader (*reg*) en amont de *trpE* et la partie 5' du gène est cloné en phase en amont du gène *lacZ* de *E. coli*, puis l'ensemble est inséré dans *pE194*. Le plasmide résultant est répliqué dans *B. subtilis*, puis utilisé pour transformer *L. lactis*. **B.** Comme le plasmide est incapable de se répliquer, le seul événement pouvant produire des transformants résistant à l'érythromycine (*Em^R*) est l'intégration du plasmide dans le chromosome par recombinaison homologue. La structure finalement obtenue dans le chromosome est indiquée en **C**.

de la fusion est contrôlée par la concentration en Trp. Quand la concentration en Trp est limitante, l'expression du gène hybride (et donc des gènes *trp*) est augmentée d'un facteur 150. Cela démontre qu'il est possible de contrôler efficacement l'expression d'un gène cloné sur le chromosome de *L. lactis*.

L'ensemble des résultats permet de proposer un modèle expliquant la régulation de la transcription des gènes *trp* en réponse à la concentration en Trp du milieu. Que le Trp soit présent ou non, une forte transcription est initiée au niveau du promoteur en amont de *trpE*. En l'absence de Trp, une fraction des transcrits n'est pas arrêtée par le terminateur de transcription *t1* et se poursuit à travers les gènes *trp* jusqu'au terminateur *t2*: les gènes *trp* sont exprimés. En revanche, en présence de Trp, on n'observe plus de transcrits des gènes *trp* et ceux-ci ne sont plus exprimés. Il semble donc que la modulation de la transcription se fasse par la modulation de l'efficacité du terminateur de transcription *t1*. Ce mécanisme semble différent des mécanismes de régulation déjà caractérisés en détail dans d'autres bactéries et nous poursuivons son étude au niveau moléculaire.

PERSPECTIVES

Les travaux que nous avons rapportés, ainsi que d'autres actuellement en cours dans notre laboratoire (Godon *et al*, 1992; Delorme *et al*, 1992; Godon *et al*, en préparation; Delorme et Renault, en préparation) ont permis d'identifier des mécanismes efficaces de contrôle de l'expression génétique dans *L. lactis* subsp *lactis*. Dans l'exemple de la biosynthèse du Trp, nous avons montré qu'il était possible d'utiliser l'un de ces mécanismes pour contrôler l'expression d'un gène étranger

en réponse à un paramètre du milieu de culture.

L'approche présentée dans cet article, le savoir-faire et les concepts acquis au cours de ce travail pourront prochainement être appliqués à l'étude de la régulation de gènes dont l'expression conditionne l'activité des lactocoques au cours de la transformation fromagère.

REMERCIEMENTS

Le travail décrit a été financé en partie par le Contrat Bridge-Biot-CT91-0263 du Biotechnology Action Programme de la Commission of the European Communities.

RÉFÉRENCES

- Band L, Shimotsu H, Henner DJ (1984) Nucleotide sequence of the *B. subtilis* *trpE* and *trpD* genes. *Gene* 27, 55-65
- Bardowski J, Erlich SD, Chopin A (1992) Tryptophan biosynthesis genes in *Lactococcus lactis* subsp *lactis*. *J Bacteriol* 174, 6563-6570
- Beckwith J (1987) The lactose operon. In: *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology* (Neidhardt FC, ed). Am Soc Microbiol, Washington DC
- Calhoun DH, Pierson DL, Jensen RA (1973) The regulation of tryptophan biosynthesis in *Pseudomonas aeruginosa*. *Mol Gen Genet* 121, 117-132
- Chopin MC, Chopin A, Rouault A, Galleron N (1989) Insertion and amplification of foreign genes in the *Lactococcus lactis* subsp *lactis* chromosome. *Appl Environ Microbiol* 55, 1769-17774
- Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16, 10881-10890
- Crawford IP (1989) Evolution of a biosynthetic pathway: the tryptophan paradigm. *Annu Rev Microbiol* 43, 567-600
- Crawford IP, Gunsalus IC (1966) Inducibility of tryptophan synthetase in *Pseudomonas putida*. *Proc Natl Acad Sci USA* 56, 717-724

- d'Aubenton Carafa Y, Brody E, Thermes C (1990) Prediction of rho-independent *Escherichia coli* transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* 216, 835-858
- Delorme C, Ehrlich SD, Renault P (1992) Histidine biosynthesis genes in *Lactococcus lactis* subsp *lactis*. *J Bacteriol* 174, 6571-6579
- de Vos WM (1987) Gene cloning and expression in lactic streptococci. *FEMS Microbiol Rev* 46, 281-295
- Essar DW, Eberly L, Han C, Crawford IP (1990) DNA sequences and characterization of four early genes of the tryptophan pathway in *Pseudomonas aeruginosa*. *J Bacteriol* 172, 853-866
- Furter R, Paravicini G, Aebi M, Braus G, Prantl F, Niederberger P, Hutter R (1986) The TRP4 gene of *Saccharomyces cerevisiae*: isolation and structural analysis. *Nucleic Acids Res* 14, 6357-6373
- Godon JJ, Chopin MC, Ehrlich SD (1992) Branched chain amino acids genes in *Lactococcus lactis* subsp *lactis*. *J Bacteriol* 174, 6580-6589
- Gouy M, Gautier C (1982) Codon usage in bacteria: correlation with gene expressivity. *Nucleic Acids Res* 10, 7055-7074
- Hadero A, Crawford IP (1986) Nucleotide sequence of the genes for tryptophan synthase in *Pseudomonas aeruginosa*. *Mol Biol Evol* 3, 191-204
- Henner DJ, Band L, Shimotsu H (1984) Nucleotide sequence of the *Bacillus subtilis* tryptophan operon. *Gene* 34, 169-177
- Hoopes BC, McClure WR (1987) Strategies in regulation of transcription initiation. In: *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology* (Neidhardt FC, ed), Am Soc Microbiol, Washington DC
- Matsui K, Sano K, Ohtsubo E (1986) Complete and deduced amino acid sequences of the *Brevibacterium lactofermentum* tryptophan operon. *Nucleic Acids Res* 14, 10113-10114
- Meile L, Stettler R, Banholzer R, Kotik M, Leisinger T (1991) Tryptophan gene cluster of *Methanobacterium thermoautotrophicum* Marburg: molecular cloning and nucleotide sequence of a putative trpEGCFBAD operon. *J Bacteriol* 173, 5017-5023
- Natori Y, Kanao Y, Imamoto F (1990) Nucleotide sequences and genomic constitution of five tryptophan genes of *Lactobacillus casei*. *J Biochem* 107, 248-255
- Slock J, Stahly DP, Han DY, Six EW, Crawford IP (1990) An apparent *Bacillus subtilis* folic acid biosynthetic operon containing *pab*, an amphibolic trpG gene, a third gene required for synthesis of para-aminobenzoic acid, and the dihydropteroate synthase gene. *J Bacteriol* 172, 7211-7226
- Somerville RL (1983) Tryptophan: biosynthesis, regulation, and large-scale production. In: *Amino acids: biosynthesis and genetic regulation* (Herrmann KM, Somerville RL, eds). Addison-Wesley Publishing Company. Reading, Massachusetts
- Tschumper G, Carbon J (1980) Sequence of a yeast DNA fragment containing a chromosomal replicator and the TRP1 gene. *Gene* 10, 157-166
- Yanofsky C, Platt T, Crawford IP, Nichols BP, Christie GE, Horowitz H, van Cleemput M, Wu AM (1981) The complete nucleotide sequence of the tryptophan operon of *Escherichia coli*. *Nucleic Acids Res* 9, 6647-6668
- Zalkin H, Yanofsky C (1982) Yeast gene TRP5: structure, function, regulation. *J Biol Chem* 257, 1491-1500
- Zalkin H, Paluh JL, van Cleemput M, Moye WS, Yanofsky C (1984) Nucleotide sequence of *Saccharomyces cerevisiae* genes TRP2 and TRP3 encoding bifunctional anthranilate synthase: indole-3-glycerol phosphate synthase. *J Biol Chem* 259, 3985-3992