KOR. JOUR. MICROBIOL April, 1994 p. 120~125 Vol. 32, No. 2

Cloning and Sequence Analysis of the *trpB*, *trpA* and 3' *trpC(F)* Genes of *Vibrio metschnikovii* Strain RH530

Kwon, Yong Tae, Jin Oh Kim, Young Dong Yoo, and Hyune Mo Rho*

Department of Molecular Biology and Research Center for Cell Differentiation, Seoul National University, Seoul 151-742, Korea

The genes, *trpB*, *trpA* and 3' *trpC(F)* of *Vibrio metschnikovii* strain RH530 were cloned and sequenced. The *trpB* and *trpA* genes had open reading frames of 1,173 bp and 804 bp encoding 391 and 268 amino acids, respectively. The *trpB* and *trpA* genes had conventional ribosome-binding sequences and overlapped with each other by one nucleotide, suggesting that these two genes are translationally coupled. 115 nucleotide upstream the *trpB* start codon, there was an incomplete open reading frame of the 3'-end of the *trpC(F)*. The amino acid sequences of *trpB*, *trpA* and *trpC(F)* of *V. metschnikovii* RH530 had identities of 64.2%, 82.4% and 73.7%, respectively, for those of *V. parahaemolyticus*; 58.7%, 72.3% and 54.9%, respectively, for *Salmonella typhimurium*; and 42.6 %, 54.1% and 12.5%, respectively, for *Brevibacterium lactofermentum*. The genetic organization of these genes, especially in the noncoding region between *trpC(F)* and *trpB*, was distinct from that of Enterobacteriaceae.

KEY WORDS □ ***Vibrio metschnikovii*, DNA sequencing, tryptophan synthase, intercistronic region, sequence homology**

The biosynthesis of tryptophan has been the object of intensive genetic and biochemical investigations in a variety of bacterial species. To date, genes of *trp* operons of *Escherichia coli* (21), *Bacillus subtilis* (8), *Pseudomonas putida* (9), *Brevibacterium lactofermentum* (14), and *Vibrio parahaemolyticus* (3) have been cloned, and the nucleotide sequence showed that their genetic organization and regulation of the pathway had considerable diversity. In most bacteria, the genes for tryptophan synthesis are linked each other, and thus constituting an operon (1, 15). On the other hand, in *P. putida* and *Acinetobacter*, the operons are scattered along the chromosomal DNA (1). Tryptophan synthase (TrpBA) (EC 4.2.1.20), the final enzyme in the tryptophan biosynthesis pathway, is composed of 2 α-subunits which cleave indolglycerol phosphate to indole and glyceraldehyde-3-phosphate and 2 β-subunits which synthesize L-tryptophan from indole and L-serine (16). In most organisms, the genes for TrpBA are preceded by the *trpC(F)* gene encoding indolglycerol phosphate synthase (EC 4.1.1.48). In the enteric bacteria *Brevibacterium* and some filamentous fungi, this gene is fused with *trpF*, the gene for N-phosphoribosylanthranilate isomerase (5).

In the family Vibrionaceae, no *trp* genes except those of *V. parahaemolyticus* (3) have been cloned. Therefore, relatively little is known about the *trp* genes of the family Vibrionaceae and their relationship with those of Enterobacteriaceae. In this paper, the *trpB*, *trpA* and part of *trpC(F)* genes of *Vibrio metschnikovii* strain RH530 (12) were cloned in *E. coli* and sequenced. This information was used for a comparative evaluation between the *trp* genes of *V. parahaemolyticus* (3), and the unrelated species *S. typhimurium* (4).

MATERIALS AND METHODS

Bacterial strains and growth media

Vibrio metschnikovii strain RH530 isolated and classified as *Vibrio* sp. by Kwon *et al.* (12) was recently reclassified as *V. metschnikovii* (Bae, K.S., personal communication). Cells were routinely cultivated on LSC medium (pH 10.5) which contains 1% bactotrypton, 0.5% yeast extract, 1% sodium chloride and 50 mM sodium carbonate buffer.

Screening of the *trp* genes

Chromosomal DNA of *V. metschnikovii* RH530 was isolated by the method of Marmur (13). The *trp* genes were screened by the colony hybridization (7). On the basis of the amino acid homology between TrpB of *E. coli* and *V. parahaemolyticus* (3), and the codon usage of *V. metschnikovii* strain RH530 (Kwon, Y.T., J.O. Kim,

S.Y. Moon and H.M. Rho, unpublished data), an oligonucleotide (5'-CAHGGMGGMGCGC-AHAANACMAAT-3', in which equimolar mixtures of the following nucleotides were incorporated at H, M and N: A and T for H; C and T for M; and A and G for N) corresponding to the region containing Lys-87 of *V. parahaemolyticus* TrpB which forms the Schiff base linkage with the coenzyme pyridoxal phosphate was synthesized. The radioactive probe was prepared by the 5'-end labelling of the 24-mer probe with T4 kinase and γ-^{32}P-ATP (3,000 Ci/mmole). The transferred colonies were lysed with 1% SDS solution, denatured with 0.5 N NaOH and renatured with 0.5 M Tris-HCl (pH 8.0). The replica was prehybridized in buffer P (6× SSC, 5× Denhardt's solution, 0.1% SDS, 100 μg/ml salmon sperm DNA) and hybridized in buffer H (6× SSC, 1× Denhardt's solution, 100 μg/ml yeast RNA, 0.05% sodium pyrophosphate) at 48℃ for 24 hr and then washed five times in 6× SSC/0.1% sodium pyrophosphate at 50℃.

Nucleotide sequencing

The 3.4-Kb *Eco*RI-*Hin*dIII-fragment of pVT1 was inserted into M13mp18 and M13mp19 and serially deleted by the exonuclease activity of T4 DNA polymerase as described by Dale *et al.* (6). The sequencing reaction was done by the dideoxy chain termination method (20) using Sequenase and dITP (US Biochemicals). The nucleotide sequence was analysed using PCGENE (Intelligenetics), and homology searches of GenBank and Swiss Protein. Database was carried through the EMBL computer network.

RESULTS AND DISCUSSION

The *Eco*RI-digested chromosomal DNA was ligated into the corresponding site of pUC19, and transformed into *E. coli* HB101. The colony hybridization (7) using the above-mentioned probe showed a single positive clone, which contained a recombinant plasmid (pVT1) with a 6.5 Kb-*Eco*RI-fragment (Fig. 1). The subcloning experiments showed that the 3.4 Kb-*Eco*RI-*Hin*dIII-fragment of the initial 6.5 Kb fragment reacted with the above-mentioned probe. The 3.4 Kb-*Eco*RI-*Hin*dIII-fragment was subcloned into both of M13mp18 and M13mp19, and the nucleotide sequence was determined using dideoxy-chain termination method (20) with [α-^{35}S]dATP and T7 Sequenase.

The nucleotide sequence of the *trpB*, *trpA* and 3' *trpC(F)* genes of *V. metschnikovii* RH530 as well as the deduced amino acid sequences of the three gene products are shown in Fig. 2. The *trpB* and *trpA* genes had open reading frames of 1,173 bp and 804 bp encoding 391 amino acids (42 kDa) and 268 amino acids (29 kDa), respectively, compared with *E. coli* (397 and 268 amino acids

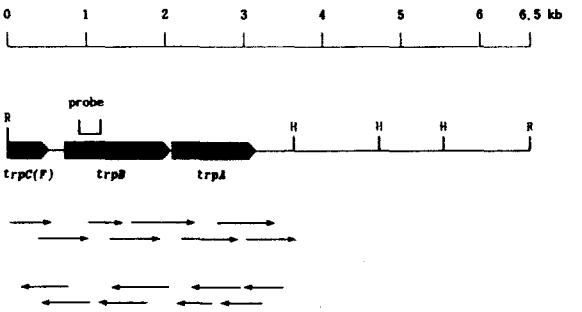

Fig. 1. *Restriction map and sequencing strategy of the 6.5 Kb-EcoRI-fragment of the plasmid pVT1 and arrangement of trpB, trpA and 3' trpC(F) of V. metschnikovii RH530.*
The solid box represents the *trp* genes. Clones used for sequencing were obtained by deletion with T4 DNA polymerase (6) as described in Materials and Methods. The arrows indicate the extent of sequence obtained from each deletion clone. The horizontal bar above the solid box represents the 24-mer oligonucleotide probe. H and R indicate *Hin*dIII and *Eco*RI site, respectively.

for *trpB* and *trpA*) and *V. parahaemolyticus* (396 and 268 amino acids for *trpB* and *trpA*). Both of the *trpB* and *trpA* genes had conventional ribosome-binding sequences 9 bp ahead of the ATG start codon. The TAA stop codon of *trpB* overlapped with the *trpA* start codon by one nucleotide, adenosine of the ATG start codon. This structure suggests that these two genes are translationally coupled as seen in *E. coli trpBA* genes (21). But, in the 3'-flanking region of the *trpA* gene, no stable stem-loop structure was found like a rho-independent terminator.

116 nucleotide upstream the *trpB* start codon, there was an incomplete open reading frame which seems to be the 3'-end of the *trpC(F)*, the fused form of *trpC* and *trpF*, on the basis of the comparison of the deduced amino acids with others (Fig. 3). In most enteric bacteria including *E. coli* and Gram-positive *B. lactofermentum*, *trpC* are fused with *trpF*, and *trpC(F)* is separated with trpBA by 8-14 nucleotides (21, 14). Recently, *trpC (F)* of *V. parahaemolyticus* belonging to the family Vibrionaceae has been shown to be separated with *trpBA* by 45 nucleotides (3). On the basis of these results, the length of the spacer between *trpC(F)* and *trpB* of the family Vibrionaceae may be longer and more variable than that of the family Enterobacteriaceae. However, since this noncoding region does not contain a promoter-like sequence or a transcriptional termination signal, its function remains unclear. The G+C contents of the coding regions of *V. metschnikovii*

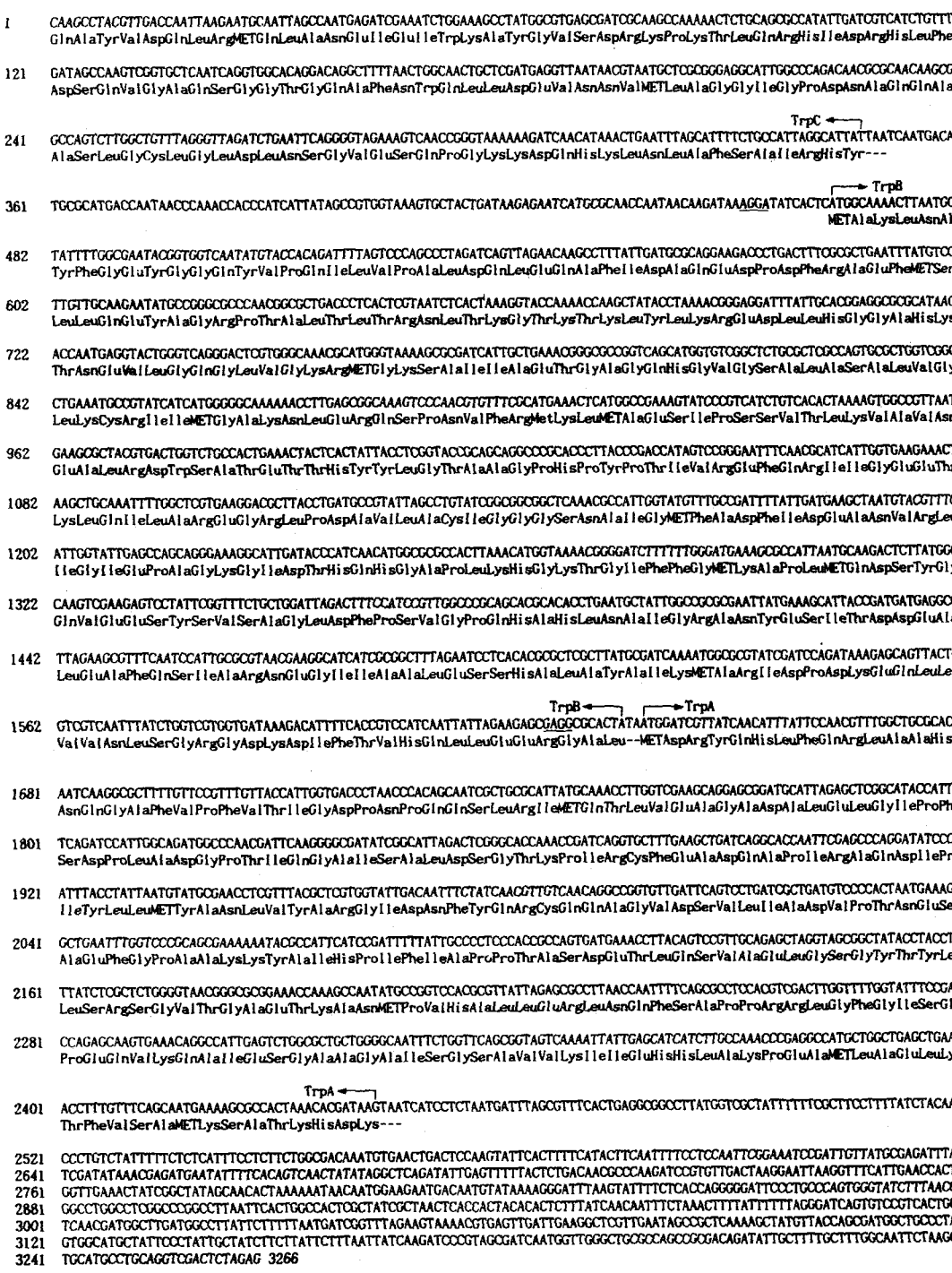

Fig. 2. *Nucleotide sequences and deduced amino acid sequences of trpB, trpA and 3' trpC(F) of V. metschnikovii RH530.*

The nucleotide sequence is numbered from 5'-end of the sequenced fragment. Probable Shine-Dalgarno sequences are underlined. The sequence data in this paper have been submitted to the EMBL Data Libraries under the accession number of Z19090.

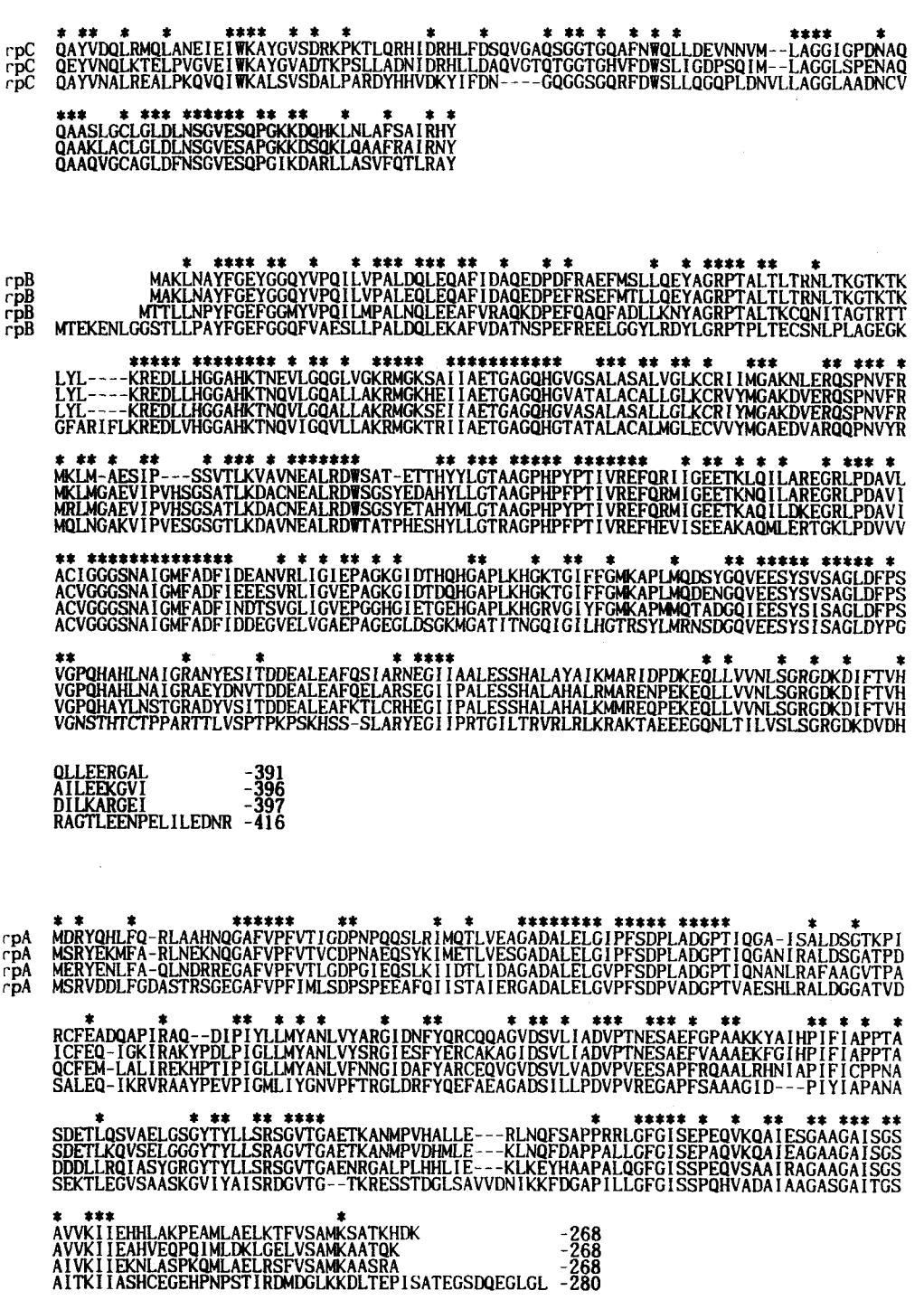

Fig. 3. *Alignment of amino acid sequences for TrpB, TrpA and the C-terminus of TrpC(F) of V. metschnikovii (VM), V. parahaemolyticus (VP), S. typhimurium (ST) and B. lactofermentum (BL).*
The amino acid sequence of *B. lactofermentum* TrpC(F) is excluded in the comparison of TrpC(F) due to its low homology to *V. metschnikovii* TrpC(F). Standard single-letter abbreviations are used. To maximize the similarity, a few gaps were introduced. Homologous residues for all given proteins are indicated as asterisks.

RH530 *trpB, trpA* and 3' *trpC(F)* are 49.3%, 49.0% and 45.9%, respectively, close to those of the genus *Vibrio*.

The amino acid sequences of TrpB, TrpA and the C-terminus of *trpC(F)* of *V. metschnikovii* RH 530 deduced from the nucleotide sequences were aligned with those of *V. parahaemolyticus* (3), *S. typhimurium* (4, 19, 10), and *B. lactofermentum* (14) (Fig. 3). The deduced amino acid sequences of *trpB, trpA* and 3' *trpC(F)* of *V. metschnikovii* RH530 had similarities of 84.1%, 73.1% and 61.4%, respectively, for those of *V. parahaemolyticus*; 75.2 %, 62.3% and 48.3%, respectively, for *S. typhimurium*; and 52.2%, 41.8% and 26.7%, respectively, for *B. lactofermentum*. This result shows that TrpB is more conservative than TrpA and that TrpC(F) is more variable than the other two polypeptides, agreeing with the previous report (2).

The amino acid residues important for catalysis have been assigned from the results of site-directed mutagenesis (17, 18, 22) and the three-dimensional structure (11) of the *S. typhimurium* tryptophan synthase. For the β subunit, Lys-87 (numbers for the *S. typhimurium* subunit) forms a Schiff base with a coenzyme, pyridoxal phosphate. Other conserved regions, including a Gly-Gly-Gly-Ser-Asn-Ala stretch at 232 to 237, form a pocket for the phosphate group of the coenzyme. Ser-377 and Gly-378 bind other parts of the coenzyme molecule. Glu-109 and His-115 are involved in binding the L-serine substrate. All the residues mentioned above are conserved in *V. metschnikovii* RH530. For the α subunit, Glu-49 and Asp-60 are essential for enzymatic activity. Phe-22, Leu-100, Tyr-102, Leu-127 and Ala-129, and Ile-153, Tyr-175, Gly-211, Phe-212, Gly-213, Gly-234 and Ser-235 are important in substrate binding. All the residues mentioned above are also conserved in *V. metschnikovii* RH530.

ACKNOWLEDGEMENT

This work was supported in part by research grants from the Ministry of Science and Technology and from KOSEF through the Research Center for Cell Differentiation (94-4-1).

REFERENCES

1. **Crawford, I.P.**, 1975. Gene rearrangements in the evolution of the tryptophan synthetic pathway. *Bacteriol. Rev.* **39**, 87-120.
2. **Crawford, I.P.**, 1987. Evolution of a biosynthetic pathway: The tryptophan paradigm. *Ann. Rev. Microbiol.* **43**, 576-600.
3. **Crawford, I.P., C.Y. Han, and M. Silverman**, 1991. Sequence and features of the tryptophan operon of *Vibrio parahaemolyticus*. *DNA Sequence* **1**, 189-196.
4. **Crawford, I.P., B.P. Nicholas, and C. Yanofski**, 1980. Nucleotide sequence of the *trpB* gene in *Escherichia coli* and *Salmonella typhimurium*. *J. Mol. Biol.* **142**, 480-502.
5. **Creighton, T.E.**, 1970. N-(5'phosphoribosyl) anthranilate isomerase-indole-3-glycerol phosphate synthetase of tryptophan biosynthesis: Relationship between the two activities of the enzyme from *Escherichia coli*. *Biochem. J.* **120**, 699-707.
6. **Dale, R.M.K., B.A. McClure, and J.P. Houchins**, 1985. A rapid single-stranded cloning strategy for production a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18S rDNA. *Plasmid* **13**, 31-40.
7. **Grunstein, M. and D.S. Hogness**, 1975. Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Aca. Sci. USA* **72**, 3961-3965.
8. **Henner, D.J., L. Band, and H. Shimotsu**, 1984. Nucleotide sequence of the *Bacillus subtilis* tryptophan operon. *Gene* **34**, 169-177.
9. **Holloway, B.W. and A.F. Morgan**, 1986. Genomic organization in *Pseudomonas*. *Ann. Rev. Microbiol.* **40**, 79-105.
10. **Horowitz, H., J.V. Arsdell, and T. Platt**, 1983. Nucleotide sequence of the *trpD* and *trpC* genes of *Samonella typhimurium*. *J. Mol. Biol.* **169**, 775-797.
11. **Hyde, C.C., S.A. Ahmed, E.A. Padlan, E,W. Miles and D.R. Davies**, 1988. Three-demensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *S. typhimurium*. *J. Biol. Chem.* **263**, 17857-17871.
12. **Kwon, Y.T., S.Y. Moon, J.O. Kim, Y.H. Kho, and H.M. Rho**, 1992. Characterization of extracellular proteases from alkalophilic *Vibrio* sp. strain RH 530. *Kor. J. Microbiol.* **30**, 501-506.
13. **Marmur, J.**, 1961. A procedure for the isolation of deoxyribonucleic acid from microorganism. *J. Mol. Biol.* **3**, 208-218.
14. **Matsui, K., K. Sano, and E. Ohtsubo**, 1986. Complete nucleotide sequence and deduced amino acid sequences of the *Brevibacterium lactofermentum* tryptophan operon. *Nucl. Acids Res.* **14**, 10113-10114.
15. **Matsui, K., K. Sano, and E. Ohtsubo**, 1987. Sequence analysis of the *Brevibacterium trp* operon. *Mol. Gen. Genet.* **209**, 299-305.
16. **Miles, E.W., R. Bauerle, and S.A. Ahmed**, 1987. Tryptophan synthase. *Methods Enzymol.* **142**, 398-414.
17. **Miles, E.W., H. Kawasaki, S.A. Ahmed, H. Morita, and S. Nagata**, 1989. The β-subunit of tryptophan synthase: Clarification of the roles of histidine 86, lysine 87, arginine 148, cysteine 170, and cysteine 230. *J. Biol. Chem.* **264**, 6280-6287.
18. **Nagata, S., C.C. Hyde, and E.W. Miles**, 1989. The α-subunit of tryptophan synthase: Evidence that aspartic acid as a catalytic residue and that the double alteration of residues 175 and 221 in a

second-site revertant restores the proper geometry of the substrate binding site. *J. Biol. Chem.* **264**, 6288-6296.
19. **Nicholas, B.P. and C. Yanofski,** 1979. Nucleotide sequences of *trpA* of *Salmonella typhimurium* and *Escherichia coli*: An evolutionary comparison. *Pro. Natl. Acad. Sci. USA* **76**, 5244-5248.
20. **Sanger, F., S. Nicklen, and A.R. Coulson,** 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5466.
21. **Yanofsky, C., T. Platt, I.P. Crawford, B.P. Nichols, G.E. Christie, H. Horowitz, N. VanCleemput, and A.N. Wu,** 1981. The complete nucleotide sequence of the tryptophan operon of *Escherichia coli*. *Nucl. Acids Res.* **9**, 6647-6668.
22. **Yutani, K., K. Ogasahara, T. Tsujita, K. Kanemoto, M. Matsumoto, S. Tanaka, T. Miyashita, A. Matsushiro, Y. Sugino, and E.M. Miles,** 1987. Tryptophan synthase α subunit glutamic acid 49 is essential for activity, studies with 19 mutants at position 49. *J. Biol. Chem.* **262**, 13429-13433.

(Received March 7, 1994)
(Accepted March 19, 1994)

초 록: *Vibrio metschnikovii* 균주 RH530의 *trpB, trpA* 그리고 3′ *trpC(F)* 유전자의 클로닝 및 염기서열 결정

권용태 · 김진오 · 유영동 · 노현모* (서울대학교 자연과학대학 분자생물학과 및 세포분화 연구센터)

Vibrio metschnikovii 균주 RH530의 *trpB, trpA* 및 3′ *trpC(F)* 유전자를 대장균에 클로닝하여 염기서열을 결정하였다. *trpB* 및 *trpA* 유전자는 각각 391 및 268 아미노산을 coding할 수 있는 1,173 bp 및 804 bp의 open reading frame을 가졌다. *trpB* 및 *trpA* 유전자는 일반적인 ribosome-binding sequence를 가졌으며 1개의 nucleotide만큼 중복되어 있었다. 이는 이들 두 유전자가 translation 단계에서 연결되어 있음을 의미한다. *trpB* 유전자의 개시코돈에서 115 nucleotide 위에 *trpC(F)*와 *trpF*의 융합체인 *trpC(F)*의 3′-말단부위의 불완전한 open reading frame이 존재하였다. *V. metschnikovii* RH530의 TrpB, TrpA 및 TrpC(F)의 아미노산 서열은 *V. parahaemolyticus*의 그것들과 각각 64.2%, 82.4%, 73.7%; *S. typhimurium*과 58.7%, 72.3%, 54.9%; *B. lactofermentum*과 42.6%, 54.1%, 12.5%의 유사성을 보였다. 이는 TrpB가 TrpA보다 서열이 잘 보존되어 있음을 나타내며 TrpC(F)는 다를 두 polypeptide에 비해서 보다 서열의 변이가 큼을 나타낸다. 이들 유전자들의 구조, 특히 *trpC(F)*와 *trpB* 사이의 noncoding 부위는 Enterobacteriaceae 비롯한 다를 개체들과는 다른 특징을 보였다.