Journal of General Microbiology (1993), **139**, 3275–3281. *Printed in Great Britain*

Cloning and characterization of a tryptophanase gene from *Enterobacter aerogenes* SM-18

KOSEI KAWASAKI,[1] ATSUSHI YOKOTA,[1]* SHIGERU OITA,[1]† CHISA KOBAYASHI,[1]‡ SHUJI YOSHIKAWA,[1]§ SHIN-ICHI KAWAMOTO,[2] SHOICHI TAKAO[1]‖ and FUSAO TOMITA[1]

[1]*Laboratory of Applied Microbiology and* [2]*Laboratory of Molecular Biology, Faculty of Agriculture, Hokkaido University, Kita-9, Nishi-9, Kita-ku, Sapporo 060, Japan*

(*Received 14 April 1993; revised 12 July 1993; accepted 21 July 1993*)

A tryptophanase gene from *Enterobacter aerogenes* SM-18 was cloned and sequenced. The structural gene for tryptophanase, *tnaA*, consisted of 1389 bp encoding 462 amino acid residues, and its nucleotide sequence and deduced amino acid sequence showed significant homology to those of *tnaA* from *Escherichia coli* K12. A short open reading frame consisting of 31 amino acid residues was found upstream of *tnaA*, and it showed some similarity to the *E. coli tnaC* gene known to be a *cis*-acting regulatory element for transcription. A partial open reading frame homologous to the 5′ end of *E. coli tnaB* was observed at the 3′-flanking region of *tnaA*. These genes may thus constitute an operon as in *E. coli*.

Introduction

Tryptophanase is a catabolic enzyme that catalyses the degradation of L-tryptophan to indole, pyruvic acid and ammonia by an α,β-elimination reaction, and requires pyridoxal 5′-phosphate as a co-factor (Wood *et al.*, 1947; Snell, 1975). This enzyme has been found mainly in enteric bacteria. The genes for tryptophanase from *Escherichia coli* K12 (Deeley & Yanofsky, 1981), *Proteus vulgaris* (Kamath & Yanofsky, 1992), *Symbiobacterium thermophilum* (Hirahara *et al.*, 1992) and *Alcaligenes faecalis* (Omori *et al.*, 1987) have been cloned, and the first three have been sequenced.

Since the reverse reaction of tryptophanase is possible at high concentrations of ammonia and pyruvic acid, this enzyme may be used as a catalyst for the production of L-tryptophan (Watanabe & Snell, 1972; Nakazawa *et al.*, 1972*a*, *b*). We have tried to improve the industrial feasibility of L-tryptophan production using tryptophanase, in which one of the substrates, pyruvic acid, is supplied economically by microbial fermentation (Takao *et al.*, 1984; Yokota & Takao, 1984, 1989; Yokota *et al.*, 1989; Oita *et al.*, 1990). A pyruvic-acid-producing mutant of *Enterobacter aerogenes*, strain LT-94 was constructed from a high-tryptophanase strain, *Ent. aerogenes* AHU1540. After pyruvic acid was produced from glucose by strain LT-94, L-tryptophan production was achieved enzymically by direct addition of the remaining substrates, i.e. indole and ammonia, to the fermentation broth (Oita *et al.*, 1990).

The amount of L-tryptophan production depended mainly on the tryptophanase activity of the producer, and thus it seemed necessary to increase the activity of tryptophanase in order to improve L-tryptophan production (Oita *et al.*, 1990). The induction of tryptophanase in this strain is severely repressed by glucose or pyruvic acid (Oita *et al.*, 1990). Thus, it is important to release repression in order to get higher enzyme activity. For these purposes, we attempted to clone and characterize the tryptophanase gene (*tna* gene) from *Ent. aerogenes* SM-18 as presented in this paper.

* Author for correspondence. Tel. +81 11 726 3523; fax +81 11 716 0879.

† Present address: National Food Research Institute, Ministry of Agriculture, Forestry and Fisheries, Tsukuba 305, Japan.

‡ Present address: Technology and Engineering Laboratories, Ajinomoto Co. Inc., Kawasaki 210, Japan.

§ Present address: Hokkaido Food Processing Research Centre, Ebetsu 069, Japan.

‖ Present address: Department of Food Science, Rakuno Gakuen University, Ebetsu 069, Japan.

Abbreviations: Sm, streptomycin; Km, kanamycin; Tc, tetracycline; Ap, ampicillin; NTG, *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine; CRP–cAMP, cyclic AMP receptor protein–cyclic AMP complex.

The DDBJ/EMBL/GenBank accession no. for the sequence reported in this paper is D14297.

Methods

Media. LB medium containing 10 g Bacto-Tryptone (Difco) l^{-1}, 5 g Bacto-Yeast Extract (Difco) l^{-1} and 5 g NaCl l^{-1} was used as the complete medium. LB-HEPES was used for the study of tryptophanase induction, and was prepared by adding 0·1 M-HEPES and 0·5 mM-pyridoxin.HCl to LB followed by adjustment of the pH to 7·0. The *tna*$^+$ selective medium contained M9 minimal salts (Miller, 1972), 4 g glycerol l^{-1}, 1 g Bacto-Vitamin Assay Casamino acids (Difco) l^{-1}, 10 mg indole l^{-1}, 50 mg 5-methyl-DL-tryptophan (5MT, Nacalai Tesque Inc., Kyoto, Japan) l^{-1}, 50 mg each of L-threonine and L-leucine l^{-1}, 200 μg thiamin.HCl l^{-1} and 1 μM-$FeCl_3$. The L-tryptophan analogue 5MT was added as an inducer for tryptophanase (Yudkin, 1976). These media were solidified by 20 g Bacto-Agar (Difco) $^{-1}$ when used in plates.

Bacterial strains, plasmids and phages. Bacterial strains, plasmids and phages used in this study are listed in Table 1.

Construction of E. coli strains for cloning the tna gene. Strain TR11M, which has both L-tryptophan auxotrophy and a tryptophanase defect, was used for detection of the *tna* gene by complementation on *tna*$^+$ selective medium in cloning of the *tna* gene, and was constructed from *E. coli* C600 r^-m^- as follows. Among the tryptophan auxotrophs derived by UV mutagenesis, a mutant (strain T4) unable to utilize indole for growth was selected as a *trpB*-defective strain. This strain was further mutagenized by treatment with NTG in order to obtain a tryptophanase-negative mutant. The selection of the tryptophanase-negative mutants was carried out by penicillin G enrichment in the *tna*$^+$ selective medium, followed by replica-plating using the *tna*$^+$ selective agar medium supplemented with 50 mg L-tryptophan l^{-1} as a master plate and one without L-tryptophan as the selective plate. Strain TR11 was obtained as a representative strain. The inability of this strain to synthesize tryptophanase protein was confirmed by SDS-PAGE of total cell proteins (data not shown). Strain TR11R was obtained by conjugation of strain TR11 with *E. coli* MCL31 (Lorence & Rupert, 1983), and is a recombination-defective strain. TR11M was derived by lysogenization of Mu-1 into TR11R by the method described by Boucher *et al.* (1977). This lysogenization stabilized the cloned tryptophanase gene on a mobile plasmid, RP4::Mu*cts*62, by repressing heat induction.

Induction and assay of tryptophanase. To study induction of tryptophanase, 3 ml of a mid-exponential phase culture in LB was inoculated into 50 ml of LB-HEPES supplemented with or without 1 g L-tryptophan l^{-1}, and cultured with reciprocal shaking at 37 °C for 12 h. For repressing conditions, LB-HEPES supplemented with both 1 g L-tryptophan l^{-1} and 10 g glucose l^{-1} was used, and the culture time was shortened to 6 h. Cells were collected by centrifugation, washed with 50 mM-potassium phosphate buffer (pH 7·2), and used for the assay. Tryptophanase activity was determined by measuring indole formation from L-tryptophan by the enzyme reaction as described by DeMoss & Moser (1969) except that the concentration of pyridoxal 5′-phosphate in the complete reaction mixture was increased to 130 mg l^{-1}. One unit of tryptophanase activity was defined as the amount of the enzyme that formed 1 μmol indole min^{-1} under these assay conditions. Activity was expressed as units per mg dry wt cells.

Detection of the Tna$^+$ *phenotype.* The Tna$^+$ phenotype of the transconjugants or transformants was checked by detecting indole liberated from L-tryptophan into the culture medium by the tryptophanase reaction. Cells were grown in LB and the culture broth supernatant was mixed with an equal volume of colour reagent (DeMoss & Moser, 1969).

DNA manipulations. Recombinant techniques were according to standard methods (Sambrook *et al.*, 1989). Extraction of the large-size plasmids pRMT33 and pRMT38 was done by the method of Shibano *et al.* (1985). Transformation of *E. coli* strains was carried out by the method of Hanahan (1985).

Cloning of the tna gene from Ent. aerogenes SM-18. Cloning of the *tna* gene was carried out basically by the method of Deeley & Yanofsky (1981) using a mobile plasmid, RP4::Mu*cts*62 (Murooka *et al.*, 1981). *Ent. aerogenes* SM-18RM was used as a donor strain, and a tryptophanase-defective strain, *E. coli* TR11M was used as recipient strain. Strain SM-18RM was constructed by conjugation of *Ent. aerogenes* SM-18 with *E. coli* K-12MuR(pGMI117) by the method of Murooka *et al.* (1981). The conjugants were selected on LB agar plates containing 150 μg Ap ml^{-1} and 100 μg Sm sulphate ml^{-1}. The doner, *Ent. aerogenes* SM-18RM was cultured in LB at 39 °C without shaking for 16 h to allow partial heat induction of temperature-sensitive Mu*cts*. The culture broth was mixed with a mid-exponential phase culture of *E. coli* TR11M in LB, and the subsequent protocols for conjugative transfer of RP4::Mu*cts*62 from *Ent. aerogenes* to *E. coli* were by the method of Murooka *et al.* (1981) with the following modifications: the conjugated cells on filter membrane were resuspended in saline, and washed with saline three times to remove L-tryptophan, and appropriate dilutions were plated onto the *tna*$^+$ selective medium supplemented with 150 μg Ap ml^{-1}, and 25 μg Tc ml^{-1}. After incubation at 37 °C, two isolates were obtained as the transconjugants harbouring RP4::Mu*cts*62::*tna*$^+$. The plasmids were extracted, and their abilities to retransform *E. coli* TR11M to Tna$^+$ were confirmed. They were designated pRMT33 and pRMT38, respectively. Plasmid pRMT33 was digested with *Hin*dIII, and the fragments generated were ligated into the *Hin*dIII site of pBR328. The ligation mixture was used to transform *E. coli* TR11R selecting for Tna$^+$. A plasmid containing a 30 kbp *Hin*dIII fragment was designated pBT1. This plasmid was digested with *Pst*I and DNA fragments were ligated into the *Pst*I site of pUC19 and the ligation mixture was used to transform TR11R to Tna$^+$. A clone that showed Tna$^+$ was selected, and its plasmid, which contained a 5·8 kbp *Pst*I fragment, was designated pKT403.

DNA sequencing. The target DNA fragments were subcloned into pUC118 and/or pUC119. A series of nested deletions of each plasmid was constructed by the method of Yanisch-Perron *et al.* (1985) with mung-bean nuclease. Single-stranded plasmid DNA for the sequencing template was prepared by infecting *E. coli* MV1184 harbouring the plasmid with M13KO7. Dideoxy sequencing was performed using the Sequenase version 2.0 DNA sequencing kit (United States Biochemical Corp.) and [α-^{32}P]dCTP (Amersham). Computer analyses of nucleotide sequence and resulting amino acid sequence were performed with the software package DNASIS (Hitachi Software Engineering Co., Japan).

SDS-PAGE. This was done by the method of Laemmli (1970) with an 8% (w/v) acrylamide separation gel. Gels were stained with Coomassie Brilliant Blue R-250. Densitometric analysis was done with a Shimadzu model CS-9000.

Results and Discussion

Subcloning and localization of the tna gene

A partial restriction map of pKT403 was determined (Fig. 1). Southern blot analysis revealed that the cloned fragment originated from the *Ent. aerogenes* chromosome (data not shown). To locate the *tna* gene, deletion derivatives (from pKT404 to pKT408, and pKT421) were constructed from pKT403 as described in Table 1, and their abilities to complement Tna$^-$ were examined in *E. coli* TR11R with L-tryptophan as the inducer. As

Table 1. *Bacterial strains, phages and plasmids*

	Genotype, phenotype or description	Source or reference
Enterobacter aerogenes		
SM-18	Sm^r derived from AHU1540	Yokota & Takao (1984)
SM-18RM	RP4::Mucts62 Km^r Tc^s Ap^r, conjugation of SM-18 × *E. coli* K12MuR(pGMI117)	This study
Escherichia coli		
K12MuR(pGMI117)	RP4::Mucts62 Km^r Tc^s Ap^r	J. Fourment
C600 r^-m^-	*thr-1 leuB6 thi-1 hsdS1 lacY1 tonA21* λ^- *supE44*	ATCC 33525
MCL31	HfrPO201 Δ(*gpt–lac*)*5 relA1 rpsE2123 thi-1 supE44 TP3* Δ(*srl–recA*)*306*::Tn*10*(Tc^r)	Lorence & Rupert (1983)
MV1184	*ara* Δ(*lac–proAB*) *rpsL thi*(*ϕ*80 *lacZ*ΔM15) Δ(*srl–recA*)*306*::Tn*10*(Tc^r) F′[*traD36 proAB*$^+$ *lacI*^{q} *lacZ*ΔM15]	Vieira & Messing (1987)
T4	Trp^- derivative from C600 r^-m^- induced by UV treatment, probably *trpB* defective	This study
TR11	Tna^- derivative from T4 induced by NTG treatment	This study
TR11R	TR11 Δ(*srl–recA*)*306*::Tn*10* (Tc^r), conjugation of TR11 × MCL31	This study
TR11M	TR11R Mu-1 lysogen	This study
Phages		
Mu-1	Mu c^+	ATCC 23724-B9
M13KO7	Km^r	Vieira & Messing (1987)
Plasmids		
pBR328	Cloning vector, Ap^r Tc^r Cm^r	Soberon *et al.* (1980)
pUC18 and pUC19	Cloning vector, Ap^r	Yanisch-Perron *et al.* (1985)
pUC118 and pUC119	Cloning vector, Ap^r	Vieira & Messing (1987)
pRMT33 and pRMT38	RP4::Mucts62 inserted with *tna*$^+$ fragment from *Ent. aerogenes* SM-18	This study
pBT1	30 kbp (approximate) *Hin*dIII fragment from pRMT33 cloned in pBR328	This study
pKT403	5·8 kbp *Pst*I fragment from pBT1 cloned in pUC19 at the same restriction site	This study
pKT404	pKT403 cut with *Kpn*I and recircularized	This study
pKT405	pKT403 cut with *Acc*I* and *Hin*cII, and recircularized	This study
pKT406	5·2 kbp *Dra*I–*Eco*RI fragment from pKT403 cloned in pUC19 cut with *Sma*I–*Eco*RI	This study
pKT407	pKT403 cut with *Sph*I and recircularized	This study
pKT408	3·5 kbp *Sac*I–*Eco*RI fragment from pKT403 cloned in pUC19 cut with the same enzymes	This study
pKT421	2·8 kbp *Sph*I–*Kpn*I fragment from pKT403 cloned in pUC19 cut with the same enzymes	This study
pKT321	2·8 kbp *Sph*I–*Kpn*I fragment from pKT403 cloned in pUC18 cut with the same enzymes	This study

*This *Acc*I terminus was blunted with bacteriophage T4 DNA polymerase before ligation.

shown in Fig. 1, the complementing activity was retained in the 2·8 kbp *Sph*I–*Kpn*I fragment. The expression and regulation of the *tna* gene coded in this *Sph*I–*Kpn*I segment were investigated (Table 2). In *E. coli* TR11R(pKT421), tryptophanase expression is inducible by L-tryptophan and is repressed severely by addition of glucose. Similar results were obtained with *Ent. aerogenes* SM-18 and *E. coli* TR11R(pKT403). This *Sph*I–*Kpn*I fragment was subcloned into pUC18 at the same restriction sites. The resulting plasmid, pKT321, which has the same fragment as pKT421, but with inverted orientation against the *lac* promoter of the plasmid vector, also displayed characteristics similar to pKT421. These results suggested that this segment contains promoter and other regulatory elements as well as the structural gene for tryptophanase.

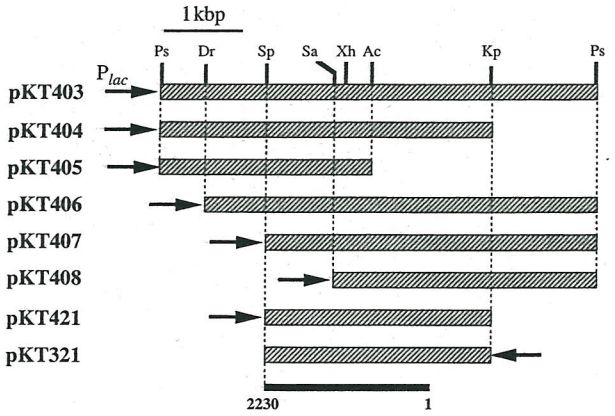

Fig. 1. Restriction map of pKT403 and complementation of Tna$^-$ by pKT403 and its deletion derivatives. Cloned DNA of each plasmid is represented by a hatched box. Arrows indicate the position and orientation of the *lac* promoter (P_{lac}). Each plasmid was introduced into tryptophanase-defective *E. coli* TR11R, and its Tna$^+$ was tested. The region sequenced and shown in Fig. 3 is indicated with coordinates. +, Complementation; –, no complementation. Abbreviations for restriction sites; Ps, *Pst*I; Dr, *Dra*I; Sp, *Sph*I; Sa, *Sac*I; Xh, *Xho*I; Ac, *Acc*I; Kp, *Kpn*I.

Table 2. *Tryptophanase induction and its repression by glucose in Ent. aerogenes SM-18 and E. coli TR11R transformants*

The strains were cultured in LB-HEPES, either unsupplemented (not induced), or supplemented with L-tryptophan (induced) or L-tryptophan plus glucose (repressed). Details are given in Methods.

Strain(plasmid)	Tryptophanase activity [units (mg dry wt cells)$^{-1}$] Not induced	Induced	Repressed
Ent. aerogenes SM-18	0·11	0·57	< 0·01
E. coli TR11R(pUC19)	< 0·01	< 0·01	< 0·01
E. coli TR11R(pKT403)	0·11	1·23	< 0·01
E. coli TR11R(pKT421)	0·25	1·27	< 0·01
E. coli TR11R(pKT321)	0·49	1·06	< 0·01

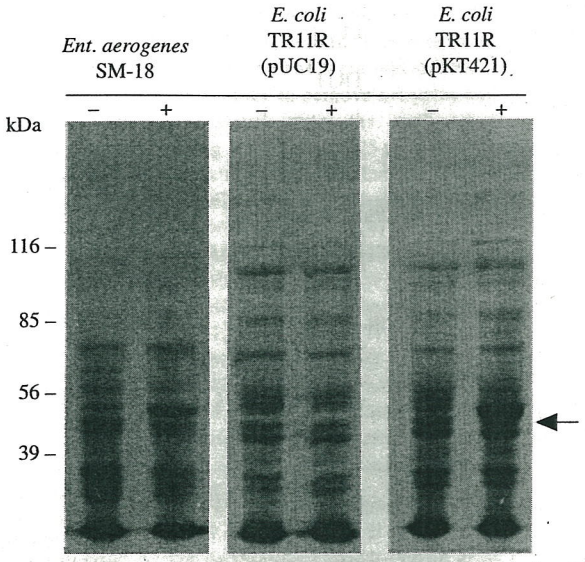

Fig. 2. SDS-PAGE analysis of proteins from whole cells. The tryptophanase was induced in *Ent. aerogenes* SM-18 and *E. coli* TR11R transformants by L-tryptophan as described in Methods. The position of the tryptophanase protein is indicated by an arrow. +, Inducing conditions; –, non-inducing conditions.

SDS-PAGE analysis of total cell proteins

This revealed that an extra protein was produced in *E. coli* TR11R(pKT421) grown under inducing conditions, which was not detected in *E. coli* TR11R(pUC19) (Fig. 2). The molecular mass of this protein was estimated to about 50 kDa. This 50 kDa protein was also observed in *Ent. aerogenes* SM-18 grown under inducing conditions. A densitometric analysis of the gel revealed that the ratio of 50 kDa protein to total cell protein was 26% for induced *E. coli* TR11R(pKT421) and 15% for induced *Ent. aerogenes* SM-18. This agreed with the results of experiments (Table 2) showing that the activity of tryptophanase in induced *E. coli* TR11R(pKT421) was twofold higher than that of induced *Ent. aerogenes* SM-18. Therefore, this 50 kDa protein was considered to be the tryptophanase protein.

Nucleotide sequence of the tna gene and deduced amino acid sequence

The nucleotide sequence of the *tna* gene region in the *Sph*I–*Kpn*I fragment was determined from both strands. Fig. 3 shows the 2230 bp sequence including the *Sph*I site at the 3′-terminus. Computer analysis of the sequence revealed one major open reading frame (ORF) comprising 1389 bp coding for 462 amino acids, and significant homology between the nucleotide sequences of this ORF and *tnaA* from *E. coli* K12 (Deeley & Yanofsky, 1981). Thus, it seems to be the structural gene for tryptophanase, and was designated *tnaA*. A Shine–Dalgarno (SD) sequence was found just before the ATG translation initiation codon. The molecular mass of the TnaA protein calculated from the predicted amino acid sequence is 51·8 kDa, which is in good agreement with the result of the SDS-PAGE (Fig. 2). The deduced amino acid sequence is 60% identical to that of *E. coli* TnaA (Fig. 4). A larger value, 70%, was observed between *Ent. aerogenes* TnaA and *P. vulgaris* TnaA (Kamath & Yanofsky, 1992). The lysine residue of *E. coli* TnaA (Lys-270 indicated by a triangle in Fig. 4) is known to

```
   1 CATGGTGCTGATCATCTCCGGCGGCATCATTATGAACGTCTTTTCCAGCATGTAAGTCACTCCTGCGCCT

  71 TCATTTATTGCGGTGAAGGCGCTAAATACCCCTACTAAATAAGCCGTCATTGGCACAATTAAAAAAACAA

                                                  AA TGTGA  T   TCACA TT
 141 AAAACCATCGCCTGCTAACCAAATAATCAATAACGCTAAAAATACGACGAGCGTCAAATTTTTATACTCC
                                                      CAP-cAMP

        -35                 -10
 211 ACTTATCAATAACTAATTTTTTGTGATTTTAATCGTGGTTATTTTCTCATATTACCTTATTTTTATGGCA

                               tnaC>
 281 TCCCTTCAACCACAGAGGACGATGTCATGACGCAAAAATCGATTTATCGCCTCATCGCCCAGGGCGGCCT
                    *****      M  T  Q  K  S  I  Y  R  L  I  A  Q  G  G  L

 351 GGACGACTGGTATAATTTAGATCACCGCATCTCCAACGACTTCCCTCGTTAATTCTCCGCTGCTGTTTAC
      D  D  W  Y  N  L  D  H  R  I  S  N  D  F  P  R  END

 421 TCCCTAAAAAAATTAATTCAGCACTGCGTTTAGTCGCTATTTGTTGCGCCCATTACTTGTCGCCAAATAG

                                         tnaA>
 491 CCAGACGCGGCACAGCATTCCCATGGAGAAAATAAATGAAACGTATCCCTGAGCCGTTCCGCATTAAAAT
                               ****      M  K  R  I  P  E  P  F  R  I  K  M

 561 GGTCGAAAATATTCGTATGACCACCTTTGACGATCGCGTCAAAGCGCTGAAAGAGGCGGGCTATAACCCG
      V  E  N  I  R  M  T  T  F  D  D  R  V  K  A  L  K  E  A  G  Y  N  P

 631 TTTTTGCTCAATAGCCAGGACGTGTATATCGACCTGCTGACCGACTCCGGCACCGGCGCGATGAGCGACC
     F  L  L  N  S  Q  D  V  Y  I  D  L  L  T  D  S  G  T  G  A  M  S  D  H

 701 ATCAATGGGCCGGGCTGATGATGGGCGACGAAGCCTACGCCGGGTCGCGTAACTATCAGCATCTGTGTGA
       Q  W  A  G  L  M  M  G  D  E  A  Y  A  G  S  R  N  Y  Q  H  L  C  E

                                 AccI
 771 AAAGGTCAAAGAGATTATCGGCTATCCGTATACCATTCCCACCCACCAGGGGCGCGGCGCGGAGCAGATT
      K  V  K  E  I  I  G  Y  P  Y  T  I  P  T  H  Q  G  R  G  A  E  Q  I

 841 CTGTTCCCAAGCCTGATTGCTCGCCGCAAGTCGGCGCATCCGGTGTTTATTTCCAACTTCCACTTCGATA
     L  F  P  S  L  I  A  R  R  K  S  A  H  P  V  F  I  S  N  F  H  F  D  T

 911 CTACCGCCGCTCACGTGGAGCTTAACGGCGCAAAAGCCATTAACGTGGTGACGCCGAAAGCCTTCGACAC
       T  A  A  H  V  E  L  N  G  A  K  A  I  N  V  V  T  P  K  A  F  D  T

 981 GACCTCGTGGTACGACTGGAAAGGCAACTTTGATATTGATTTACTGAAAGCGACCATCGCCGAGCACGGC
      T  S  W  Y  D  W  K  G  N  F  D  I  D  L  L  K  A  T  I  A  E  H  G

1051 GCGGAAAACGTGGCGGCGATTATTACTACCGTCACCTGTAACAGTTCCGGCGGCCAGCCGGTATCGCTGG
     A  E  N  V  A  A  I  I  T  T  V  T  C  N  S  S  G  G  Q  P  V  S  L  A

1121 CGAATATGCGCGAGGTGTATCAAATTGCGCGCCAGAACAATATTCCGGTGGTAATCGACTCCGCCCGCTT
       N  M  R  E  V  Y  Q  I  A  R  Q  N  N  I  P  V  V  I  D  S  A  R  F

1191 CTGCGAAAACGCCTGGTTTATTAAACAGCGCGAAGAGGGCTATGCCGATAAATCGGTGAAAGAGATTATT
      C  E  N  A  W  F  I  K  Q  R  E  E  G  Y  A  D  K  S  V  K  E  I  I

     XhoI
1261 CTCGAGATGTACCAGTATGGCGATATGCTGACCATGTCCGCCAAGAAAGACCCGATGGTGAATATTGGCG
     L  E  M  Y  Q  Y  G  D  M  L  T  M  S  A  K  K  D  P  M  V  N  I  G  G

                                 SacI
1331 GACTGTGCTGCTTCCGCGATGACGAAGAGCTCTTCAACGAAGTACGCATTCGCTGCGTGCCGATGGAAGG
       L  C  C  F  R  D  D  E  E  L  F  N  E  V  R  I  R  C  V  P  M  E  G

1401 CTTTGTCACCTACGGCGGCCTGGCGGGGCGCGATATGGAGGCGCTGGCTATCGGTCTGGAAGAAGGGACC
      F  V  T  Y  G  G  L  A  G  R  D  M  E  A  L  A  I  G  L  E  E  G  T

1471 AACGAGGATTACCTCGCTTATCGCATCAATCAGGTGGAGTACCTCGGTGAACGCCTGCGCGAAGGCGGCA
     N  E  D  Y  L  A  Y  R  I  N  Q  V  E  Y  L  G  E  R  L  R  E  G  G  I

1541 TTCCGATTCAGTATCCGACCGGCGGCCATGCGGTGTTCGTCGATGCTAAGCTGCTGCTGCCGCATATTCC
       P  I  Q  Y  P  T  G  G  H  A  V  F  V  D  A  K  L  L  L  P  H  I  P

1611 GCCGGAGCAGTTCCCGGCGCACGCGCTAAACAACGAGCTGTATCTGGAAGCGGGCATTCGCAGCGTGGAA
      P  E  Q  F  P  A  H  A  L  N  N  E  L  Y  L  E  A  G  I  R  S  V  E

1681 ATCGGCTCCCTGCTGCTGGGACGCGACCCGGAAACCGGTAAGCAGAAGGCGTCGCCGATGGAGCTGCTGC
     I  G  S  L  L  L  G  R  D  P  E  T  G  K  Q  K  A  S  P  M  E  L  L  R

1751 GCCTGACCATTCCGCGTCGGGTTTATACCAACGATCATATGGACTATATCGCTGATGCGCTGATCGCCGT
       L  T  I  P  R  R  V  Y  T  N  D  H  M  D  Y  I  A  D  A  L  I  A  V

1821 TAAGGCGCGCGCCGCCACGATTAAGGGCCTGACCTTTACCTATGAACCGCCGGTACTGCGCCACTTCGTC
      K  A  R  A  A  T  I  K  G  L  T  F  T  Y  E  P  P  V  L  R  H  F  V

1891 GCCAGACTGAAACCGGTGAAATAAACCCTATCAGGGCGCTCTCGCAATAAGAGCGCCATTAATTTAACTC
     A  R  L  K  P  V  K  END           >>>>>>>>         <<<<<<<<

1961 CATCTTTTTAAGCTCAATTATTCTGGTTGCGCCGGGTAGCCGTTTGCCCGCAGCCCGTTATTTACTATAT
                                       >>>>  >> >     < <<  <<<<

                               tnaB'>
2031 AGCCACGTGAGGAATGCACCTTATGGCTACATTATCAGTCTCAAAAGCGGTCTCAACTAATACACCATCG
           *****               M  A  T  L  S  V  S  K  A  V  S  T  N  T  P  S

2101 CTGATTGGCGGAGCGATGATTATCGGCGGCACGATTATTGGCGCAGGCATGTTTTCTTTACCGGTGGTGA
     L  I  G  G  A  M  I  I  G  G  T  I  I  G  A  G  M  F  S  L  P  V  V  M

                                                                   SphI
2171 TGTCCGGTGCCTGGTTTTTCTGGTCGCTGCTGGCGCTTATCTTTACCTGGTTTTGCATGC  3'
       S  G  A  W  F  F  W  S  L  L  A  L  I  F  T  W  F  C  M
```

Fig. 3. Nucleotide sequence of the *tnaA* region. The deduced amino acid sequence is shown. A potential binding site for CRP–cAMP is shown (underlined) with its consensus sequence. Putative −35 and −10 sequences are indicated by use of **bold** type. Asterisks indicate the SD sequences. The inverted repeat sequences at the 3′-flanking region of the *tnaA* are indicated by arrowheads.

```
E.AERO. (  1)  M---KRIPEPFRIKMVENIRMTTFDDRVKALKEAGYNPFLLNSQDVYIDLLTDSGTGAMS
E.COLI  (  1)  MENFKHLPEPFRIRVIEPVKRTTRAYREEAIIKSGMNPFLLDSEDVFIDLLTDSGTGAVT

E.AERO. ( 58)  DHQWAGLMMGDEAYAGSRNYQHLCEKVKEIIGYPYTIPTHQGRGAEQILFPSLIARR---
E.COLI  ( 61)  QSMQAAMMRGDEAYSGSRSYYALAESVKNIFGYQYTIPTHQGRGAEQIYIPVLIKKREQE

E.AERO. (115)  KSA-HP--VFISNFHFDTTAAHVELNGAKAINVVTPKAFDTTSWYDWKGNFDIDLLKATI
E.COLI  (121)  KGLDRSKMVAFSNYFFDTTQGHSQINGCTVRNVYIKEAFDTGVRYDFKGNFDLEGLERGI

E.AERO. (172)  AEHGAENVAAIITTVTCNSSGGQPVSLANMREVYQIARQNNIPVVIDSARFCENAWFIKQ
E.COLI  (181)  EEVGPNNVPYIVATITSNSAGGQPVSLANLKAMYSIAKKYDIPVVMDSARFAENAYFIKQ

E.AERO. (232)  REEGYADKSVKEIILEMYQYGDMLTMSAKKDPMVNIGGLCCFRDDE--ELFNEVRIRCVP
E.COLI  (241)  REAEYKDWTIEQITRETYKYADMLAMSAKKDAMVPMGGLLCMKDDSFFDVYTECRTLCVV
                                              ▲

E.AERO. (290)  MEGFVTYGGLAGRDMEALAIGLEEGTNEDYLAYRINQVEYLGERLREGGIPIQYPTGGHA
E.COLI  (301)  QEGFPTYGGLEGGAMERLAVGLYDGMNLDWLAYRIAQVQYLVDGLEEIGVVCQQ-AGGHA

E.AERO. (350)  VFVDAKLLLPHIPPEQFPAHALNNELYLEAGIRSVEIGSLLLGRDPETGKQKASPMELLR
E.COLI  (360)  AFVDAGKLLPHIPADQFPATGLACELYKVAGIRAVEIGSFLLGRDPKTGKQLPCPAELLR

E.AERO. (410)  LTIPRRVYTNDHMDYIADALIAVKARAATIKGLTFTYEPPVLRHFVARLKPVK
E.COLI  (420)  LTIPRATYTQTHMDFIIEAFKHVKENAANIKGLTFTYEPKVLRHFTAKLKEV
```

Fig. 4. Alignment of the deduced amino acid sequence of *Ent. aerogenes* SM-18 TnaA (E.AERO.) with *E. coli* K12 TnaA (E.COLI). Colons indicate identical residues. Closely related residues are indicated by dots. Gaps (dashes) were introduced into the sequences to improve the alignment. A triangle (▲) under the *E. coli* TnaA sequence indicates the lysine residue that forms a Schiff's base with pyridoxal 5′-phosphate.

form a Schiff's base with pyridoxal 5′-phosphate (Kagamiyama *et al.*, 1970). In the *Ent. aerogenes* sequence, a lysine residue corresponding to this Lys-270 was found.

In *E. coli*, *tnaA* is part of the tryptophanase operon. The promoter requires a CRP-cAMP for the initiation of transcription, and there is a CRP-cAMP binding site just upstream of the promoter sequence (Deeley & Yanofsky, 1982). It has been shown that for inducible expression of the operon, translation of the short ORF coding for 24 amino acids, *tnaC*, located between *tnaA* and the promoter region, is necessary (Stewart & Yanofsky, 1986). An antitermination site for the rho-dependent terminator, lying between *tnaC* and *tnaA*, plays an important role in gene expression (Stewart & Yanofsky, 1985; Stewart *et al.*, 1986). Furthermore, translation of a single Trp codon in *tnaC* is essential for inducible expression of this operon (Gollnick & Yanofsky, 1990). A structural gene for a tryptophan specific permease, *tnaB*, is at the 3′-flanking region of *tnaA* (Sarsero *et al.*, 1991). A similar gene organization has been reported in *P. vulgaris* (Kamath & Yanofsky, 1992).

The nucleotide sequences surrounding *Ent. aerogenes*

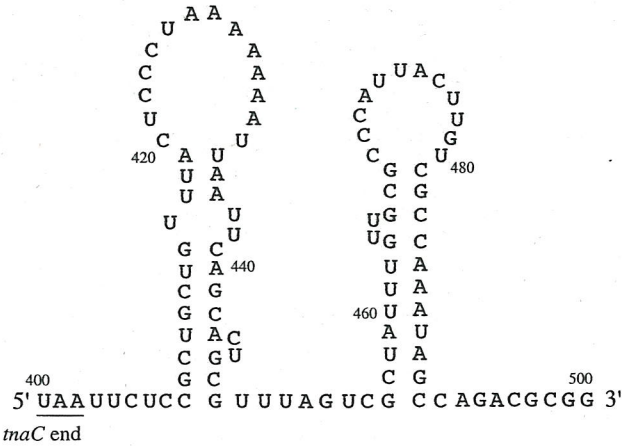

Fig. 5. Potential RNA secondary structure derived from the sequence located between *tnaC* and *tnaA*.

tnaA showed a similar organization. A symmetrical sequence homologous to the consensus sequence for the CRP–cAMP binding site of *E. coli* (De Crombrugghe *et al.*, 1984) was observed at about 335 bp upstream of the *tnaA* transcription initiation codon. Downstream of this region, a putative promoter (−35 and −10) was found. A small ORF coding for 31 amino acid residues preceded by a typical SD sequence was found to start 219 bp upstream of the *tnaA* translation initiation codon. Comparison of this ORF with *tnaC* of *E. coli* yielded no significant homology between them. However, there exists a single Trp codon as in *E. coli tnaC*, and the amino acid sequence following this Trp codon was Tyr-Asn-Leu-Glu, which is similar to that of *E. coli* (Phe-Asn-Ile-Glu; Gollnick & Yanofsky, 1990) (Fig. 3). When we consider the importance of the Trp codon in *tnaC* of *E. coli*, this similarity leads us to consider this ORF as an equivalent of *E. coli tnaC*. Therefore, we designated this ORF *tnaC*. However, the function of this ORF remains to be elucidated. In *E. coli*, a rho-dependent termination site was found between *tnaC* and *tnaA*. In *Ent. aerogenes*, a sequence which can form a sufficiently stable secondary structure was found between *tnaC* and *tnaA* (Fig. 5). Downstream of *tnaA*, a partial ORF that has homology to the 5′ end of *E. coli tnaB* was found (Sarsero *et al.*, 1991). As this partial ORF seemed to be a part of the structural gene for tryptophan permease, we designated it *tnaB′* (Fig. 3). Between *tnaB* and *tnaB′* were found two inverted repeats with stretches of T bases, which may serve as a rho-independent termination site (Fig. 3).

In conclusion, we have found that the organization of the *tna* gene from *Ent. aerogenes* SM-18 is similar to those from *E. coli* K12 and *P. vulgaris*. Therefore, we propose that the *tna* gene may constitute an operon structure in *Ent. aerogenes* SM-18.

We thank Dr J. Fourment for providing *E. coli* K12MuR(pGMI117), and Dr C. S. Rupert for providing *E. coli* MCL31. This work was supported in part by the Grants-in-Aid for Scientific Research (Nos 02760061 and 03760062) from the Ministry of Education, Science and Culture of Japan. This work was also supported in part by the Special Grant-in-Aid for Promotion of Education and Science in Hokkaido University provided by the same ministry.

References

Boucher, C., Bergeron, B., De Bertarmio, M. B. & Dénarié, J. (1977). Introduction of bacteriophage Mu into *Pseudomonas solanacearum* and *Rhizobium meliloti* using the R factor RP4. *Journal of General Microbiology* **98**, 253–263.

De Crombrugghe, B., Busby, S. & Buc, H. (1984). Cyclic AMP receptor protein: role in transcription activation. *Science* **224**, 831–838.

Deeley, M. C. & Yanofsky, C. (1981). Nucleotide sequence of the structural gene for tryptophanase of *Escherichia coli* K-12. *Journal of Bacteriology* **147**, 787–796.

Deeley, M. C. & Yanofsky, C. (1982). Transcription initiation at the tryptophanase promoter of *Escherichia coli* K-12. *Journal of Bacteriology* **151**, 942–951.

DeMoss, R. D. & Moser, K. (1969). Tryptophanase in diverse bacterial species. *Journal of Bacteriology* **98**, 167–171.

Gollnick, P. & Yanofsky, C. (1990). $tRNA^{Trp}$ translation of leader peptide codon 12 and other factors that regulate expression of the tryptophanase operon. *Journal of Bacteriology* **172**, 3100–3107.

Hanahan, D. (1985). Techniques for transformation of *E. coli*. In *DNA Cloning: a Practical Approach*, vol. 1, pp. 109–135. Edited by D. M. Glover. Oxford: IRL Press.

Hirahara, T., Suzuki, S., Horinouchi, S. & Beppu, T. (1992). Cloning, nucleotide sequences, and overexpression in *Escherichia coli* of tandem copies of a tryptophanase gene in an obligately symbiotic thermophile, *Symbiobacterium thermophilum*. *Applied and Environmental Microbiology* **58**, 2633–2642.

Kagamiyama, H., Morino, Y. & Snell, E. E. (1970). The chemical structure of tryptophanase from *Escherichia coli*. I. Isolation and structure of a pyridoxyl decapeptide from borohydride-reduced holotryptophanase. *Journal of Biological Chemistry* **245**, 2819–2824.

Kamath, A. V., Yanofsky, C. (1992). Characterization of the tryptophanase operon of *Proteus vulgaris*. Cloning, nucleotide sequence, amino acid homology, and *in vitro* synthesis of the leader peptide and regulatory analysis. *Journal of Biological Chemistry* **267**, 19978–19985.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.

Lorence, M. C. & Rupert, C. S. (1983). Convenient construction of *recA* deletion derivatives of *Escherichia coli*. *Journal of Bacteriology* **156**, 458–459.

Miller, J. H. (1972). *Experiments in Molecular Genetics*, pp. 431–432. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Murooka, Y., Takizawa, N. & Harada, T. (1981). Introduction of bacteriophage Mu into bacteria of various genera and intergeneric gene transfer by RP4::Mu. *Journal of Bacteriology* **145**, 358–368.

Nakazawa, H., Enei, H., Okumura, S., Yoshida, H. & Yamada, H. (1972*a*). Enzymatic preparation of L-tryptophan and 5-hydroxyl-L-tryptophan. *FEBS Letters* **25**, 43–45.

Nakazawa, H., Enei, H., Okumura, S. & Yamada, H. (1972*b*). Synthesis of L-tryptophan from pyruvate, ammonia and indole. *Agricultural and Biological Chemistry* **36**, 2523–2528.

Oita, S., Yokota, A. & Takao, S. (1990). Enzymatic production of tryptophan using a lipoic acid and thiamine double auxotroph of *Enterobacter aerogenes* having both pyruvic acid productivity and high tryptophanase activity. *Journal of Fermentation and Bioengineering* **69**, 256–258.

Omori, K., Shibatani, T. & Tosa, T. (1987). Cloning of tryptophanase gene of *Alcaligenes faecalis* for effective production of L-tryptophan. *Journal of Biotechnology* **5**, 17–28.

SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

SARSERO, J. P., WOOKEY, P. J., GOLLNICK, P., YANOFSKY, C. & PITTARD, A. J. (1991). A new family of integral membrane proteins involved in transport of aromatic amino acids in *Escherichia coli*. *Journal of Bacteriology* **173**, 3231–3234.

SHIBANO, Y., YAMAGATA, A., NAKAMURA, N., IIZUKA, T., SUGISAKI, H. & TAKANAMI, M. (1985). Nucleotide sequence coding for the insecticidal fragment of the *Bacillus thuringiensis* crystal protein. *Gene* **34**, 243–251.

SNELL, E. E. (1975). Tryptophanase: structure, catalytic activities, and mechanism of action. *Advances in Enzymology* **42**, 287–333.

SOBERON, X., COVARRUBIAS, L. & BOLIVAR, F. (1980). Construction and characterization of new cloning vehicles. IV. Deletion derivatives of pBR322 and pBR325. *Gene* **9**, 287–305.

STEWART, V. & YANOFSKY, C. (1985). Evidence for transcription antitermination control of tryptophanase operon expression in *Escherichia coli* K-12. *Journal of Bacteriology* **164**, 731–740.

STEWART, V. & YANOFSKY, C. (1986). Role of leader peptide synthesis in tryptophanase operon expression in *Escherichia coli* K-12. *Journal of Bacteriology* **167**, 383–386.

STEWART, V., LANDICK, R. & YANOFSKY, C. (1986). Rho-dependent transcription termination in the tryptophanase operon leader region of *Escherichia coli* K-12. *Journal of Bacteriology* **166**, 217–223.

TAKAO, S., YOKOTA, A. & TANIDA, M. (1984). Enzymatic production of tryptophan coupled to pyruvic acid fermentation. *Journal of Fermentation Technology* **62**, 329–334.

VIEIRA, J. & MESSING, J. (1987). Production of single-stranded plasmid DNA. *Methods in Enzymology* **153**, 3–11.

WATANABE, T. & SNELL, E. E. (1972). Reversibility of the tryptophanase reaction: Synthesis of tryptophan from indole, pyruvate, and ammonia. *Proceedings of the National Academy of Science of the United States of America* **69**, 1086–1090.

WOOD, W. A., GUNSALUS, I. C. & UMBREIT, W. W. (1947). Function of pyridoxal phosphate: resolution and purification of the tryptophanase enzyme of *Escherichia coli*. *Journal of Biological Chemistry* **170**, 313–321.

YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

YOKOTA, A. & TAKAO, S. (1984). Conversion of pyruvic acid fermentation to tryptophan production by the combination of pyruvic acid-producing microorganisms and *Enterobacter aerogenes* having high tryptophanase activity. *Agricultural and Biological Chemistry* **48**, 2663–2668.

YOKOTA, A. & TAKAO, S. (1989). Pyruvic acid production by lipoic acid auxotrophs of *Enterobacter aerogenes*. *Agricultural and Biological Chemistry* **53**, 705–711.

YOKOTA, A., OITA, S. & TAKAO, S. (1989). Tryptophan production by a lipoic acid auxotroph of *Enterobacter aerogenes* having both pyruvic acid productivity and high tryptophanase activity. *Agricultural and Biological Chemistry* **53**, 2037–2044.

YUDKIN, M. D. (1976). Mutations in *Escherichia coli* that relieve catabolite repression of tryptophanase synthesis. Mutations distant from the tryptophanase gene. *Journal of General Microbiology* **92**, 125–132.