

Cloning and Genetic Analysis of the Tryptophan Genes of *Azospirillum lipoferum*

S. RAMSCHÜTZ, K. KRAFT and W. KLINGMÜLLER

Department of Genetics, University of Bayreuth, W-8580 Bayreuth, Germany

Fax 49 (921) 552535

Received October 20, 1992; Accepted June 2, 1992

Abstract

The *trp*-genes of *A. lipoferum* were subcloned from a cosmid gene bank of *A. lipoferum* by complementation of an *E. coli trp*⁻ mutant which lacks the complete *trp* operon. On a fragment of about 20 kb in size the genes *trpE*, *trpD*, *trpC* and *trpA* were localized by complementation of different *E. coli trp*⁻ mutants with subcloned DNA fragments. At least three different *trp* gene clusters could be identified. The *trpE* gene was subcloned in pUC18 and sequenced. After hybridization of the *trpE* gene against total genomic DNA two signals were observed. It therefore could be possible that *A. lipoferum* possess two copies of the *trpE* gene.

Keywords: *Azospirillum*, tryptophan genes, *trpE* sequence

Abbreviations: B: *Bam*HI, E: *Eco*RI, H: *Hind*III, K: *Kpn*I, P: *Pst*I, S: *Sal*I, Sm: *Sma*I, X: *Xba*I

1. Introduction

The soil bacterium *Azospirillum* is of interest because of its ability to produce phytohormonal substances (Zimmer and Bothe, 1988, Ruckdäschel, 1987). Inoculation experiments indicated that the production of indole-acetic-acid (IAA) stimulates plant growth.

As a precursor of IAA the amino acid tryptophan is of interest. For *Azospirillum lipoferum* there was a lack of information about the genes involved in the *trp*-biosynthesis and about the organization of these genes. The results of our work provide some information about these aspects.

2. Materials and Methods

All DNA manipulations were done as described previously (Abdel-Salam and Klingmüller, 1987). For DNA sequencing a pUC18 sequencing kit was used (Boehringer Mannheim, Germany). Hybridization was done at high stringency conditions (20×SSC, 66°C) using a non-radioactive Digoxigenin detection kit (Boehringer Mannheim, Germany).

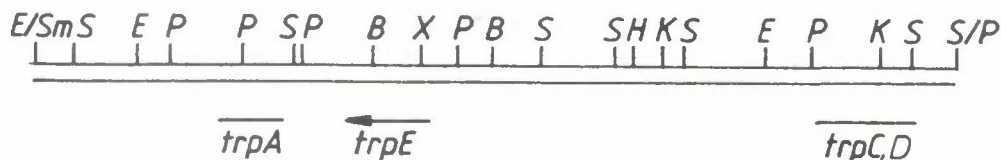
3. Results and Discussion

Cloning of the trp-genes

Total genomic DNA of *A. lipoferum* was partially digested with *Sau3A* and fragments with a size of 20–22 kb were cloned into the *Bam*HI-site of the cosmid vector pRK312 (Ruckdäschel, 1987). This gene bank was conjugated into an *E. coli* strain which lacks the complete *trp*-operon (Yanofsky et al., 1981). By using complementation analysis two cosmid-clones, named pK1 and pT, were isolated that allowed the mutant to grow on minimal medium without tryptophan. Both cosmids had a DNA-region of about 20 kb in common and must carry all genes necessary for *trp*-biosynthesis. Till now we subcloned the genes *trpE*, *trpD*, *trpC* and *trpA* of *A. lipoferum* (Fig. 1). *TrpB* and *trpF* could not yet be identified.

Sequencing of trpE gene

The *trpE* coding for anthranilate – synthase - α - subunit was sequenced (Fig. 2).



1.3kb

Figure 1. Restriction map of the *trp*-region of *A. lipoferum* and location of different *trp*-genes.

1 ACGGCGCTCGACCCGGGACCGCCCTCGACCCGGTGAICGACGGCGTGGACGCCCGCGG 60
 61 GCGTCTGCTGTCCAGCGGGGTGGAGCGCCGGCGCGCTACCGCGCTCACCGCTGGGCT 120
 121 TCACCGACCCCGGCTGGCGCTCACCGCGCGTGGCGGGACGCTCGGCATCGACCGGGCTG 180
 181 AACGGCGGGGGCAAGTGCCTGCCCGCGCTCGCCGAGGCCCTGCGTGGCCTGGGAGCG 240
 241 CCTGGCGGCTAGAGGAGGCGCGCTCGGGGTCACTGCCCTCGTCCGGAAGCCCAAGCA 300
 301 CCGCTTCCCGGAGGAGGAGCGGAGCCCGCAAGCCCTCCGTTTTCGGTCTCGCGGGCGT 360
 361 GCTGATCTGTTTGCAGCCCAACCGCTTGCCTGGCTTACGGCTTGGCTACACCTCG 420
 421 GCGTTCAGTTCGAGCGGATCGCGCAGCGGTGGAGCGGCCGACGAACCGCGGATCT 480
 481 GCTGCTCTACCTGCCGACCGGCTCGTCCGCTGGACCCCATCGAGGACTCGCCCGCT 540
 541 M S S S R R R R A A P R G W S A A G A
 CGTCCGCTATGAGTTCATCACGGCGCGGCGAGCACCGAGGGCTGGATGCGCGGGCGG 600
 601 T T P T V P T P T P R P A A T T R P V T
 CGACACCCCTACCGTCCGACACCAACCGGAGGCGGCTCGGACCAACCGCGCGTGA 660
 661 I S A S S R R A P R P P S A A A T C S R W
 CTATCAGCGGCTCGTCGAGAGCGCAAGCGCGCTTCCGCCGGCGACCTGTTGAGGT 720
 721 C P A R P S P S P G R R A F V G V P G L
 GGTGCCCGCGACAGCTTCGCGCAGCCCTGGCCGACCGCTTCTGCGGTGTTCCGGGC 780
 781 P A A N P A P Y E A F V N L G R G E F L
 TCCCGCGCCCAACCGCGGCTTACGAGGCTTCTGCAACCTCGGGCGGGCGAGTTC 840
 841 V A A S P E M Y V R V A G G R V E T C P
 TCGTCCCGCGACCGCGGAGATGTATGCGGGTGGCGGGCGGGTGAACCTCGC 900
 901 J S G T V A R G A D A L G R R P A G P G
 CGATCTCCGGCAGCTGGCGCGGGCGGCGGCTGGGGCAGCCCGCGAGTCTG 960
 961 A C L T S A K D A A E L T M C T D V D R
 GCGCTGCTGACCTCGCGCAAGGACCGGGGAGTGAACATGTGACCGAGCTGGACC 1020
 1021 N D K G A G V R A G I R P G D R A A D D
 GCAACGACAGGGCGCGGTGTGCGACCGGGATCCGTCGGGTGATCGGGCGGGATG 1080
 1081 R A V L P S D P H G G P C G G T A A V R
 ATCAGCTGTACTCCGCTCTGATCCACAGGTTGACCATGTGGAGGACCGCTCGCTCC 1140
 1141 N G R A G R L P H P Q L G G D G D R R A
 GGAATGGACCGCTGGACCGCTTCTCACCCACAGCTGGCGGTGACGGTGAACCGCGG 1200
 1201 Q A L G H A V P G G Y G A I A G A A G T
 CCAAGCGTGGGCEATGAGTCTCTCGAGGATACGGAGCAATCGCCGGCGCGCTGGTA 1260
 1261 A G P F F G R L G F D G G M D T G L T L R
 CGCGGGCGGCTTGGCGCGCTGGGCTTGCACCGGGGATGGACCGCGCTGACCTGC 1320
 1321 T I R M A E G V A Y V R A G A T L L S D
 GCACCATCCGATGCGCCAGCGGCTTCTACCTGCGGGCGGGGCGACGCTGCTGCCG 1380
 1381 S D P D A E D A E C R L K A A A F R D A
 ACAGCATCCGACCGGAGCGGAGTGGCGGCTGCGCGCTCAAGCGCGCGGCTTCCCGACG 1440
 1441 I R G T A A G A A P T L P A A P R G G E
 CCATCGCGGGACCGCGCGGTGGCGGCCACGCTGCGCGCGCTCCCGTGGCGGGG 1500
 1501 G R R V L L V D N D D S F V H T L A D Y
 AGGGCAGCGGGTGTCTGTTGGATCAGCAGCAGGCTTCTCCACACGCTGGCCGACT 1560
 1561 L G O T G A S V T T L R H S H A R A A L
 ATCTCGGCGCAGACGGGCGCTTCCGTGACGACGCTGCTCACAGTCAACGACGGCGGCG 1620
 1621 A D G R P D L V V L S P G P G P P G G F
 TGGCGGAGGGGAGCGGATGCTGCTGCTGTCGCCCGTCCGGGCGCGCGCGGAT 1680
 1681 R R C G H H R R C A G P R P A G V R R L
 TTCAGCTGGCGGACCATGACCGCGGCTGGCCCTCGGCTGCGGCTTCCGGCTC 1740
 1741 P G P A R D G G G L R R R A G R A A G A
 TGCTGGGCTGCAAGGATGGTGGAGGCTTCCGCGCGCGCTGGACGCTGCTGCCGGAG 1800
 1801 R P R Q G D E V R V L G G A L F A G L P
 CCGTCCACGGCAAGCGGACGAGGTCGGGTGCTGGCGCGCGCTGTTCGCGCGCTCG 1860
 1861 E R L T V E R Y H S L V A R R D R L P A
 CGGAGCGCTCAGGTCGGGCGCTACCACTCTCTGGTCCCGGCGGACCGGCTGCGCG 1920
 1921 D L T V T A E T A D G L V M A V E H R R
 CGGACCTCAGGTCAGCCGGAGACTCCCGAGGCTGCTGGTATGGCGGTGAGCACCGCG 1980
 1981 L P L A A V Q F H P E S I L S L D G G A
 GGTTCGCTGCGCGCGCTGCACTTCCACCCGAGTCACTGCTGCTGCGCTGACGGTGGG 2040
 2041 G L A L L G N V M D R L A A G A L T D A
 CCGGCTTCCGCTGCTGGGCAACGTGATGGACCGGCTGGCGCGCGCGGCTGACGGAG 2100
 2101 A A
 CTGGCGCTGATCGGGCGCGCTAACGGGAAGGAGTGGGGTGGTACTCCACCATCC 2160
 2161 CCAGCGCTTCTTGTAGAGTCCAGGATGGCTTCTGCTCTGGCGGCTGGCGCTGTCCA 2220
 2221 TTTGGCGAGCGGATGATCTGGCGGATGATCTTGGTGTGAAACCGGTGCCCTTGGCT 2280
 2281 CCGAGTAGACCTCTTGTATGCTCTCTCGAGGTCAGCTCT 2320

Figure 2. Nucleotide and amino acid sequence of the *trpE*-gene of *A. lipoferum*. Start and stop condons are underlined. Upstream of the start codon, at position 525, a possible ribosomal binding site is boxed (from Ramschütz, 1991).

```

  1   ...MSSRRRAAPRGWSAAGATTPTUPTPTPRPAATTRPUTISASSRAP
  ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
201  YA AKAWIDRYDFARENLS TE GKAADIAEFEPFRSUDSIPPHGDHRFGEYAE
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
  47  RPPSAAATCSR.....WCPARPSPPSGRRRAFUGVGPLPAAN
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
251  LVV KAKESFRRGDLFEVVPGQKYERCESRPSEISNR.....LKAIN
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

  83  PAPYEA FUNLGRGEFLVAASPEMYVRVAGGRVETCPISGTVARGADALGR
  |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
293  PSPYSFFINLGNQEYL VGASPEMFV RVSGRRIETCPISGTIKRGDOPIAD
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
133  RPAGPGA CLTS AKDAAELT MCTDVDRNDKG .AGVRAGIRPGDRAADD...
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
343  SEQILKL .LNSKKDESELT M CSDVDRNDK SRVCUPGSVKVIGRRQIEMYS
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

179  RA VLPSDPHGGPCGGTAAV R NGRAGRLPH PQLGGD GDRRAQALGHAVPGG
  |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
392  RL IHTVDHIEGRL RD DMDAFD GFLSHAWAVTUTGAPKLWAMRFIESHEKS
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

229  YG AIAGAAGTAGPFGRLGFDGGMDTGLTLRTIRMAEGVAYVRAGATLLSD
  |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
442  PRAWY GGA ..... IGMVGFNGDMNTGLTLRTIRIKDGI AEV RAGATLLYO
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

279  SDPDAEDAECRLKAAAF RDAIRGT .AAGAAPTLP AAPRG GEGRRULLVDH
  |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
487  SNPEEEEEAE TELKASAMIAAIRDAKSANS AKSAR DVAAVGAGVSILLVDH
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

328  DDSFVHTLADYLGQTGASV T TLRHSHARAALADGRPD LVVLSPGGPGPPGG
  .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
537  EDSFVHTLANYFRQ T GASVTTVRTPVAEEIFDRUKPD LVVLSPGGPGTPKD
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

378  FRRGGHRRGAGPRPA.....GURRLPGPAR DGGGLRRRAGRAAGARPR
  |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
587  FDCKATI K KARARDLPIFGVCLGL QALAEAY .GGDLRQLAIPMHGKPSR
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

422  QGDEV RVLG .GALFAGLPERLTVGRYHSLVARRDRLPADLTVTAETADGL
  .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
635  ....IRVLEPGIVF SGLGKEVTVGRYHSIFADPSNL PEFVITAESEDGT
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

471  UMAVEHRR LPLAAVQFH P ESILSLDGGAGLALLGNVMDRLAAGALTDAAA
  .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.
681  IMGIEH SK EPVA AVQFH P ESIMT LGGDAGMRMIENUVAHLAKRAKTKAA
    ..  .| .| .|. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |. |.

```

Figure 3. Amino acid sequence (as deduced from the nucleotide sequence) of the *trpE*-protein for *A. lipoferum* (upper line) and *R. meliloti* (lower line). Identities are indicated by vertical bars.

An ORF of 1560 bp encoding a 520 aa-polypeptide as identified. The G/C content was 75%. The *A. lipoferum* *trpE* gene showed a low amino acid-homology to the *E. coli* *trpE* gene (43%), and to 13 other *trpE*-genes (*Leptospira biflexa*, *Pseudomonas syringae*, *Pseudomonas putida*, *Corynebacterium glutamicum*, *Methanobacterium thermoautotrophicum*, *Spirochaeta aurantia*, *Thermus thermophilus*, *Vibrio parahaemolyticus*, *Bacillus lactofermentum*, *Lactobacillus biflexa*, *Acinetobacter calcoaceticus*, *Salmonella typhimurium* and *Clostridium thermocellum*). However, stronger homologies at the amino acid level to the *trpE* sequence of *Rhizobium meliloti* were found (Fig. 3; *Rhizobium* data from Bae et al., 1989). Neither leader/attenuator regions nor typical promoter regions could be identified in a sequenced region of *Azospirillum* of more than 500 bp upstream of the *trpE* startcodon (data not shown).

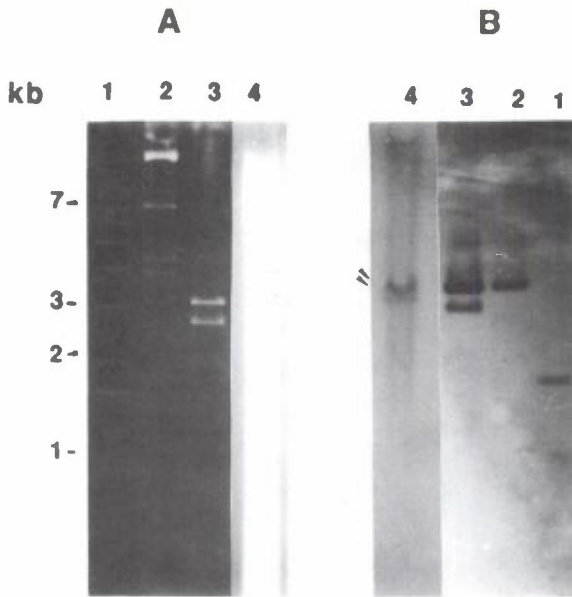


Figure 4. Hybridization with the *Bam*HI/ *Xba*I fragment of pT, carrying an internal part of the *trpE* gene of *A. lipoferum*. Hybridization was done in 20×SSC at 66°C using a non-radioactive Digoxigenin detection kit (Boehringer Mannheim, Germany). Target DNA were 25 µg of the genomic DNA of *A. lipoferum*, cut with *Pst*I. 1: KBL; 2: pT, cut with *Pst*I; 3: pUC18, carrying *trpE* gene of *A. lipoferum* on a *Pst*I fragment; 4: total genomic DNA of *A. lipoferum*, cut with *Pst*I.

Organization of the *trp*-genes

In contrast to *E. coli* (Crawford, 1989) the *trp*-genes of *A. lipoferum* are not organized in a single operon but presumably there are at least three different gene clusters. While the genes *trpD* and *trpC* are localized on one restriction fragment of about 3 kb in size, *trypE* is about 10 kb distant from *trypD/trypC* and *trypA* again is about 2 kb distant from *trypE* (Fig. 1).

Tn5-mutagenesis and hybridization experiments with the *trpE* gene

By using the plasmids pGS9 and pSUP2021 for Tn5-mutagenesis in *A. lipoferum*, it had not been possible to isolate a *trp*⁻ mutant (Abdel-Salam and Klingmüller, 1987). We therefore addressed the question whether there could exist several copies of the *trp* genes. A *Bam*HI/ *Xba*I - internal fragment of the *trpE* of *A. lipoferum* was labeled and used for hybridization against total DNA of *A. lipoferum* digested with *Pst*I.

Two signals were obtained - one with the expected size of 3 kb and one with a size of 2.8 kb (Fig. 4). Further experiments will be done to confirm the existence of two copies of the *trpE* gene in *A. lipoferum*.

Acknowledgements

Thanks are due to Dr. C. Yanofsky for *trp*⁻ *E. coli* strains, and to Dipl. ing. engr. C. Rappold for help with sorting the sequence data.

REFERENCES

- Abdel-Salam, S.M. and Klingmüller, W. 1987. Transposon Tn5 mutagenesis in *Azospirillum lipoferum*: Isolation of indole acetic acid mutants. *Mol. Gen. Genet.* **210**: 165-170.
- Bae, Y.M., Homgren, E., and Crawford, I.P. 1989. *Rhizobium meliloti* anthranilate synthase: Cloning, sequence and expression in *Escherichia coli*. *J. Bacteriol.* **171**: 3471-3478.
- Crawford, I.P. 1989. Evolution of a biosynthetic pathway: The Tryptophan Paradigm. *Annu. Rev. Microbiol.* **43**: 567-600.
- Ramschütz, S. 1991. Klonierung der Tryptophangene aus *Azospirillum lipoferum*. Untersuchungen zur Organisation, Subklonierung eines *TrpE*-Gens. Ph.D. thesis, University of Bayreuth, Germany.
- Ruckdäschel, E. 1987. Aminotransferasen für aromatische Aminosäuren aus *Azospirillum lipoferum*: Identifizierung der Enzyme, Beteiligung an der Biosynthese von Indolessigsäure und Nachweis zugehöriger Gene in Cosmidgenbanken. Diplom-Thesis, University of Bayreuth, Germany.

- Yanofsky, C., Platt, T., Crawford, I.P., Nicols, B.P., Christie, G.E., Horowitz, H., Van Cleemput, M., and Wu, A.M. 1981. The complete nucleotide sequence of the tryptophan operon of *E. coli*. *Nucleic Acids Res.* **9**: 6647-6663.
- Zimmer, W. and Bothe, H. 1988. The phytohormonal interactions between *Azospirillum* and wheat. *Plant Soil* **110**: 239-247.