Evolution of bacterial *trp* operons and their regulation

Enrique Merino[1], **Roy A Jensen**[2], and **Charles Yanofsky**[3]

1*Department of Molecular Microbiology, Instituto de Biotecnologia, Universidad Nacional Autonoma de Mexico, Cuernavaca, Morelos 62271, México (merino@ibt.unam.mx)*

2*Emerson Hall, PO Box 110700, University of Florida, Gainesville, FL 32611, USA (rjensen@ufl.edu)*

3*Department of Biological Sciences, Stanford University, Stanford, CA 94305-5020, USA (yanofsky@stanford.edu)*

Summary

Survival and replication of most bacteria require the ability to synthesize the amino acid L-tryptophan whenever it is not available from the environment. In this article we describe the genes, operons, proteins, and reactions involved in tryptophan biosynthesis in bacteria, and the mechanisms they use in regulating tryptophan formation. We show that although the reactions of tryptophan biosynthesis are essentially identical, gene organization varies among species - from whole-pathway operons to completely dispersed genes. We also show that the regulatory mechanisms used for these genes vary greatly. We address the question - what are some potential advantages of the gene organization and regulation variation associated with this conserved, important pathway?

Introduction

Extensive knowledge exists for the genes, enzymes, operons, and reactions of L-tryptophan (Trp) biosynthesis of bacteria, and on the mechanisms they use in regulating Trp formation [**1-5]. All organisms with this pathway use structurally similar enzymes, suggesting that the genes for this pathway evolved just once, probably late in the evolution of genes for amino acid biosynthesis. Trp is one of the rarest amino acids in most proteins, and it is the most costly to synthesize. It is generally encoded by a single codon, UGG, which may have served as a stop codon in early codon evolution. The 3D structures of all seven of the Trp biosynthetic enzymes are known, and most resemble structures of proteins catalyzing similar reactions in other pathways. The organization of *trp* genes within operons and the regulatory mechanisms used to control *trp* operon expression both vary greatly, undoubtedly reflecting organismal divergence in relationship to different metabolic contexts of Trp biosynthesis. Various *trans*-acting proteins and *cis*-acting sites and regions are used to regulate *trp* operon expression. The variety of mechanisms used suggests that groups of organisms experienced differing selective pressures in response to differing capabilities and needs. Intracellular concentrations of Trp and of charged and uncharged tRNA^{Trp} are the cues most often sensed in *trp* operon regulation. In this article we review and update knowledge on the distribution of the genes, enzymes, and regulatory mechanisms used by bacteria in Trp biosynthesis.

Phylogenetic distribution of *trp* gene organization and associated regulatory factors, elements, and mechanisms

Figure 1 portrays a 16S rRNA tree of selected organisms that are color-coded according to major phylogenetic groupings in the Bacteria. Although some studies on *trp* gene regulation in Archaea have appeared recently [6-9], these are outside the scope of this review. Shown to the right are the *trp*-gene clustering, *trp*-gene fusions, and regulatory features. *trp* genes that have specialized functional roles other than primary Trp biosynthesis and which therefore are not regulated by Trp are not shown. A few pathogens (*Haemophilus*, *Coxiella*, and *Propionibacterium*) that appear to be in the process of losing the Trp pathway [4] are also shown. The extensive variation in *trp* gene arrangements within operons and the repertoire of alternative mechanisms used in regulating transcription of these genes (Fig. 2), illustrate the plasticity that bacteria must have had for optimization of Trp biosynthesis.

Sensing Trp and tRNA^{Trp} in Gram-negative bacteria, Actinobacteridae, and Deinococci

Not surprisingly, particular gene arrangements and regulatory strategies are often present in related phylogenetic groupings. For example, in *Escherichia coli* and other lower-Gammaproteobacteria the seven *trp* biosynthetic genes are organized within a single transcriptional unit, the *trp* operon, which always has a *trpCF* gene fusion and which, in a smaller group, has a *trpGD* fusion (Fig. 1). In this group of bacteria two principal mechanisms are used for transcriptional regulation - one sensing Trp and a second sensing uncharged tRNA^{Trp}. Initiation of transcription of the *trp* operon is regulated by repression by the Trp-activated TrpR repressor protein (Figs. 2a-2b) [10,11].

Regulation by TrpR in *Psychrobacter* and *Coxiella* has been predicted by bioinformatic analysis, and regulation by TrpR has been demonstrated experimentally [12] in the phylogenetically distant Chlamydiales (which are not Gram-negative bacteria). Thus far, *trpR* generally appears to be autoregulated, and it exists predominantly in a monocistronic operon (Fig. 1). However, in some organisms *trpR* is co-transcribed with other *trp* genes, allowing coordinate expression, e.g., in the *trpREG*, *trpRBA*, and *trpRDCFBA* operons of *Psychrobacter cryohalolentis*, *Chlamydia trachomatis* and *Chlamydophila caviae*, respectively (Fig.1).

Uncharged tRNA^{Trp} is also sensed by transcription attenuation mechanisms as a regulatory signal (Figs. 2c-2d) [11,*13,14]. Interestingly, although *trp* operons containing all the *trp* genes are commonly subject to regulation by transcription attenuation, in some organisms this mechanism is used to regulate only a subset of the *trp* genes (Fig. 1). Thus, transcription attenuation is used to regulate the *trpE* and *trpGDC* operons of *Pseudomonas aeruginosa*, the *trpEG* operon of some Alphaproteobacteria (e.g., *Rhizobium etli*) and Actinobacteridae (e.g., *Streptomyces coelicolor*), as well as the monocistronic *trpE* operon of Deinococci (e.g., *Deinococcus radiodurans*) (Fig 1). In *P. aeruginosa* (and *P. entomophila*) transcription of the *trpBA* operon is activated by the TrpI protein in response to the accumulation of indoleglycerol phosphate (InGP), a Trp biosynthetic intermediate [15,16] (Figs. 2e-2f). InGP accumulates whenever the cellular Trp concentration is low. The *trpI* and *trpBA* operons of these organisms are divergently transcribed (Fig. 1) and their transcription is regulated by the binding of TrpI at their overlapping regulatory regions (Figs. 2e-2f). When the Trp concentration is high, the InGP concentration is low, and TrpI binds near its promoter and autoregulates its own synthesis. When the Trp concentration is low and InGP accumulates, two molecules of activated TrpI are bound near the *trpBA* promoter, activating *trpBA* transcription (Figs. 2e-2f) [15,16].

Sensing Trp and tRNA^{Trp} in Gram-positive bacteria

Although Gram-positive bacteria also sense both the intracellular level of free Trp and the availability of charged tRNA^{Trp} as regulatory signals, quite different strategies are used by these organisms in regulating the *trp* biosynthetic operon (Fig. 1). For example, in *B. subtilis* and some of its close relatives, coordinate expression of all seven *trp* genes is required for Trp biosynthesis. Six of these genes are clustered in the *trp* operon, *trpEDCFBA*, which is transcribed as part of a supraoperon also containing genes involved in the common aromatic pathway, and in phenylalanine, tyrosine, and histidine biosynthesis. The seventh *trp* gene, *trpG*, (also functioning as *pabA*) is located in the unlinked folate operon. Transcription of the *trp* operon of *B. subtilis* is regulated by attenuation by the Trp-activated *trp* RNA-binding attenuation protein, TRAP [17,18] (Figs. 2g-2h). In the presence of excess Trp, TRAP-mediated transcription termination predominates, but those few transcripts that escape termination are subject to translational regulation via formation of a secondary structure that blocks ribosome access to the *trpE* ribosome-binding site [19]. Expression of *trpG* is coordinately regulated with the other *trp* genes by Trp-activated TRAP, which binds at the *trpG* mRNA ribosome-binding site, inhibiting its translation [20]. TRAP regulation of the *trp* operon also occurs in some Clostridia, although here *trp* operon organization varies significantly from that of *B. subtilis*. For example, in *Syntrophomonas wolfei* and *Carboxydothermus hydrogenoformans* all the *trp* genes are transcribed as part of a common operon. In the former, *trpE* and *trpG* are fused and located at the end of the operon while in the latter, the operon begins with the *aroF* gene, followed by unfused *trpE* and *trpG* genes (Fig. 1).

In *B. subtilis* the availability of uncharged tRNA^{Trp} is also sensed, and accumulation of uncharged tRNA^{Trp} regulates synthesis of AT, an anti-TRAP protein. AT binds to Trp-activated TRAP, inhibiting its function [21,22]. (Figs. 2i-2j). The structural gene for AT, *rtpA*, is transcriptionally regulated by the T box mechanism (see below), in response to the accumulation of uncharged tRNA^{Trp} [22]. AT synthesis is also regulated translationally by uncharged tRNA^{Trp} [21]. AT appears to have evolved very recently since it has only been identified in *B. subtilis* and the closely related *B. licheniformis* (Fig. 1). In the other TRAP-containing organisms it is not known whether some other regulatory mechanism is used to sense uncharged tRNA^{Trp} accumulation.

In the vast majority of Gram-positive bacteria, TRAP is not present. Transcription of the *trp* operon is regulated in response to uncharged tRNA^{Trp} accumulation, by the T box mechanism [23] (Figs. 2k-2l). This mechanism involves leader RNA recognition of uncharged tRNA^{Trp}-targeted with great specificity and affinity in the absence of any protein factors. The T box was originally identified as the regulatory element used for many aminoacyl-tRNA synthetase operons in Gram-positive bacteria [23]. It was later found to be associated with many other operons of amino acid metabolism, such as those containing transporter genes and biosynthesis genes. The T box elements regulating *trp* biosynthetic genes are often tandemly arranged in the upstream transcribed region of the *trp* operon. However, in some species only a single T box element is present [**1,**24] Fig. 1). In organisms that use the T box mechanism to sense uncharged tRNA^{Trp} in regulating *trp* operon expression, the existence of a free-Trp sensing mechanism is unknown. In organisms with tandem T box elements, the quantitative advantages derived from employing tandem T boxes might compensate for the lack of a separate regulatory mechanism for sensing Trp [**1,**24]. Regulation of the *trp* operon by the T box mechanism is almost exclusively confined to the Firmicutes.

Possible reasons for differences in operon organization and regulatory mechanisms

The Trp pathway requires the products of four other pathways for completion of Trp biosynthesis. Chorismate is the common precursor of all three aromatic amino acids, as well as a precursor of other essential metabolites, e.g., folic acid. Glutamine provides an essential amino group during anthranilate formation, PRPP is the source of several carbons of the indole ring of Trp, and L-serine provides the side-chain of Trp. PRPP, an important high-energy metabolite, must be used efficiently. If exogenous anthranilate were to enter the cell and generate unneeded Trp, this could decrease the availability of PRPP for L-histidine synthesis (and for other functions). This likely explains why anthranilate phosphoribosyl transferase is a second target in addition to anthranilate synthase (AS) for Trp feedback inhibition in enteric bacteria [25]. *Bacillus subtilis* employs an alternative - and indirect - mechanism whereby its AS is activated by histidine [26].

Beyond the fate of Trp as a substrate for protein synthesis, some intermediates in Trp biosynthesis – and Trp itself – can serve other functions, and these functions are highly individualistic in nature. In view of this, it is perhaps not surprising that *trp* operon organization varies appreciably in different bacterial species, and that different mechanisms are used in regulating its expression. **(i)** Organisms such as *P. aeruginosa* use anthranilate for several purposes. (See the later discussion of quinolone biosynthesis in the LGT section). Thus, the overall need for AS here may not be coordinated with the need for synthesis of other Trp pathway enzymes. These relationships may partially explain the split operon (Fig. 1) [4] and the differential regulation of Trp enzymes seen in this species and its close relatives. **(ii)** In plant-associated organisms such as *Azospirillum brasilense* [27] Trp assists plant growth by serving as precursor of indole-3-acetic acid. This organism possesses two AS enzymes, one sensitive to feedback inhibition by Trp and controlled by a TrpL leader peptide region. The second AS is neither sensitive to feedback inhibition by Trp, nor known to be subject to repression by Trp, suggestive of its specialized role in auxin synthesis [28]. **(iii)** *Chromobacterium violaceum* produces large quantities of the pigment, violacein, which is derived from two molecules of Trp [29]. The *trp* genes of this organism (not shown in Fig. 1) are mostly scattered and present as single copies. It would be interesting to know what control strategies dictate the two fates of Trp in *C. violaceum*. **(iv)** *Streptomyces coelicolor* possesses a large gene cluster associated with formation of a calcium-dependent antibiotic containing Trp. The cluster includes copies of four of the *trp* genes (not shown in Fig. 1), implying their specialized participation in the antibiotic pathway (4). Hence, the remaining three *trp* genes which are present as single copies are uniquely used for two different pathway functions. **(v)** species of *Cytophaga, Polaribacter, Xanthomonas*, and *Gemmata* (not shown in Fig. 1) synthesize quinolinate from Trp and use it as a precursor of niacin [30]. In some of these organisms only single *trp* genes are present, implying that regulatory mechanisms exist that do not depend upon differential regulation of multiple copies of *trp* genes.

An interesting alteration of metabolic context for the Trp pathway is exemplified by the mosaic *trp* operon of *Chlamydophila caviae* (Fig. 1), an intracellular pathogen which recycles Trp-derived kynurenine of the mammalian host back to pathogen-available Trp [2]. The *trp* operon of this organism lacks genes encoding AS, but possesses the remaining *trp*-pathway genes. Its mosaic *trp* operon also contains genes encoding PRPP synthase (not otherwise present in Chlamydia) and kynureninase (required for conversion of host kynurinine to pathogen anthranilate). The repressor gene, *trpR*, is also within the operon, indicating that it is auto-regulated (Fig. 1).

Enzymatic diversity

The reactions of Trp biosynthesis are invariant, but the enzymes that catalyze them often exhibit distinct characteristics. Many different gene-fusion combinations exist in different lineages [4], and such fusions are reliable markers of phylogenetic relatedness.

Anthranilate synthase (AS)

The TrpE polypeptide, in complex with the TrpG polypeptide, catalyzes anthranilate formation from chorismate and glutamine [31]. *p*-aminobenzoate (PABA) synthase, a close homolog of AS, performs a similar reaction to generate PABA for folate synthesis [32]. Both enzymes use an amidotransferase subunit to deliver glutamine-derived ammonia to the active site. AS, but not PABA synthase, is feedback inhibited by Trp. Since PABA synthase nevertheless has a Trp-binding pocket, this is consistent with an evolutionary scenario whereby the PABA synthase gene evolved from *trpE* by gene duplication, followed by loss of Trp feedback inhibition. PABA is not derived from chorismate in some organisms [*33].

Anthranilate phosphoribosyl transferase

TrpD is a highly conserved enzyme, which together with nucleoside phosphorylase and thymidine phosphorylase represent one of three types of phosphotransferase folds [34]. Interestingly, under certain conditions, TrpD supplies phosphoribosyl amine for thiamine biosynthesis [35].

The TrpF, TrpC, and TrpA polypeptides

These three enzymes are of similar size and all have αβ barrel structures. All three may have originated via gene duplication from an ancestral gene encoding an enzyme that may have been capable of catalyzing all three reactions. In some Actinomycetes the *trpF*-encoded isomerase polypeptide is missing, and its function is provided by the similar histidine-pathway isomerase, denoted PriA [36]. The group of organisms lacking *trpF* include all of the *Actinobacteridae* subclass, except a small group within the Family *Corynebacteriaceae* (Fig. 1) in which the entire *trp* operon (except for a *trpC* remnant) has been replaced via LGT (see later section).

TrpB

Some organisms have two TrpB proteins, one of which complexes with TrpA and functions in Trp formation, while the other (a clearly distinct subhomolog) may act as a serine deaminase [*37,38]. It has been proposed that TrpB and the enzyme catalyzing the final step in threonine biosynthesis have a common origin [39].

Lateral gene transfer

The transfer of whole-pathway and partial-pathway *trp* operons to other organisms via LGT has been extensively reviewed recently [3,40]. Even where LGT origin of a given operon is clearly evident, it is rare that both the donor and the recipient have been studied experimentally in order to deduce the evolution of altered functional roles and altered regulation. It is exciting that recent publications now allow a comprehensive understanding of *trp*-operon LGT in the following two examples.

Whole-pathway LGT in *Corynebacterium*

A common ancestor of some of the species within the genus *Corynebacterium* received a complete *trp* operon from a member of the enteric lineage [3,4]. Genes having the functional role of primary Trp biosynthesis in the recipient have been replaced by genes having the same functional role in the donor. The recipient genome can be pinpointed to an ancestor that existed

quite recently at a time **after** the divergence of *C. jeikeium* (which does not have the LGT operon; Fig. 1) but **before** the divergence of C. *glutamicum*, *C. efficiens*, and *C. diptheriae* (which all have the LGT operon). It is quite interesting that acquisition of the LGT operon has provided the isomerase encoded by *trpF*, a gene that is otherwise missing from a large group of actinomycete bacteria, as previously discussed. Thus, like most actinomycetes, *C. jeikeium* relies on the histidine-pathway isomerase (PriA) for dual function in two pathways, whereas the three sister species possess specialized isomerases (one for the Trp pathway and one for the histidine pathway). The transferred operon has the *trpC/trpF* fusion that is typical of the enteric lineage. *cis* elements of regulation that were transferred include a leader peptide associated with an attenuator region, as well as an internal promoter. The structural gene for the *trpR* repressor was not transferred, as might be expected of a regulatory gene unlinked to the biosynthetic structural genes. However, a recent publication shows that this has been compensated for by conscription of LtbR, an IclR-type transcriptional repressor [**41] that has a highly conserved role in regulation of leucine biosynthesis in actinomycete bacteria. Thus, the Trp pathway of *C. glutamicum* can be appreciated as a novel evolutionary jump that resulted from a combination of LGT from a distant donor and conscription of a native regulatory gene.

Partial-pathway LGT in *Pseudomonas aeruginosa*

P. aeruginosa is a very recent recipient of a *trpEG* operon from a donor within the enteric lineage, i.e., the proximal fragment of the seven-gene *trp* operon. As shown by a recent publication [**42], this exemplifies a case where the genes, which in the donor were dedicated to primary biosynthesis, were recruited for a new specialized function in the recipient (and hence not shown in Fig. 1 where only *trp* genes participating in primary tryptophan biosynthesis are illustrated). The control elements ahead of AS were either not transferred or were discarded. Sensitivity to feedback inhibition by Trp was also lost. The *trpEG* operon was placed adjacent to the *pqsABCDE* operon. Together these seven genes encode the gene products which produce 2-heptyl-3-hydroxy-4-quinolone (PQS), a quorum-sensing signaling compound and virulence factor. The two operons are controlled positively by PqsR, and the gene products are primarily expressed during stationary-phase metabolism. Hence, the alien *trp* genes were assimilated into a different functional context and placed under a completely different control regimen. This represents an extremely recent innovation at the species level since the LGT operon is not present in very close relatives of *P. aeruginosa*. The LGT event is sufficiently recent that one could deduce that the *trpEG* donor species might perhaps have been from *Vibrio*.

Conclusions

The genes and operons of the Trp biosynthetic pathway are organized differently in various bacterial species. These differences reflect evolutionary divergence as well as adjustment to unique metabolic capabilities and environmental interactions. Associated with these differences in operon organization are the use of a variety of *cis*-acting sites and *trans*-acting proteins in regulating operon transcription. In addition, some organisms may have acquired, in one step, via LGT, an operon that has experienced evolutionary modifications that allow it to provide a new function and/or new regulation. Thus LGT must have also contributed to the operon diversity we now observe.

References and recommended reading

Papers of particular interest, published recently, have been highlighted as:

* of special interest

** of outstanding interest

**1. Gutierrez-Preciado A, Jensen RA, Yanofsky C, Merino E. New insights into regulation of the tryptophan biosynthetic operon in Gram-positive bacteria. Trends Genet 2005;21:432–436. [PubMed: 15953653]The extensive studies of Trp-pathway regulation in Gram-positive bacteria within the *Bacillales* Order provide an excellent comparative complement to what is known of Gram-negative bacteria related to *Escherichia coli*.

2. Xie G, Bonner CA, Jensen RA. Dynamic diversity of the tryptophan pathway in chlamydiae: reductive evolution and a novel operon for tryptophan recapture. Genome Biol 2002;3:research0051. [PubMed: 12225590]

3. Xie G, Bonner CA, Song J, Keyhani NO, Jensen RA. Inter-genomic displacement via lateral gene transfer of bacterial *trp* operons in an overall context of vertical genealogy. BMC Biol 2004;2:15. [PubMed: 15214963]

4. Xie G, Keyhani NO, Bonner CA, Jensen RA. Ancient origin of the tryptophan operon and the dynamics of evolutionary change. Microbiol Mol Biol Rev 2003;67:303–342. [PubMed: 12966138]

5. Yanofsky C. Using studies on tryptophan metabolism to answer basic biological questions. J Biol Chem 2003;278:10859–10878. [PubMed: 12556463]

6. Cubonova L, Sandman K, Karr EA, Cochran AJ, Reeve JN. Spontaneous *trpY* mutants and mutational analysis of the TrpY archaeal transcription regulator. J Bacteriol 2007;189:4338–4342. [PubMed: 17400746]

7. Gast DA, Wasserfallen A, Pfister P, Ragettli S, Leisinger T. Characterization of *Methanobacterium thermoautotrophicum* Marburg mutants defective in regulation of L-tryptophan biosynthesis. J Bacteriol 1997;179:3664–3669. [PubMed: 9171414]

8. Tang X, Ezaki S, Fujiwara S, Takagi M, Atomi H, Imanaka T. The tryptophan biosynthesis gene cluster trpCDEGFBA from *Pyrococcus kodakaraensis* KOD1 is regulated at the transcriptional level and expressed as a single mRNA. Mol Gen Genet 1999;262:815–821. [PubMed: 10628865]

9. Xie Y, Reeve JN. Regulation of tryptophan operon expression in the archaeon *Methanothermobacter thermautotrophicus*. J Bacteriol 2005;187:6419–6429. [PubMed: 16159776]

10. Rose JK, Squires CL, Yanofsky C, Yang HL, Zubay G. Regulation of *in vitro* transcription of the tryptophan operon by purified RNA polymerase in the presence of partially purified repressor and tryptophan. Nat New Biol 1973;245:133–137. [PubMed: 4582892]

11. Yanofsky C. The different roles of tryptophan transfer RNA in regulating *trp* operon expression in *E. coli* versus *B. subtilis*. Trends Genet 2004;20:367–374. [PubMed: 15262409]

12. Akers JC, Tan M. Molecular mechanism of tryptophan-dependent transcriptional regulation in *Chlamydia trachomatis*. J Bacteriol 2006;188:4236–4243. [PubMed: 16740930]

*13. Merino E, Yanofsky C. Transcription attenuation: a highly conserved regulatory strategy used by bacteria. Trends Genet 2005;21:260–264. [PubMed: 15851059]Evidence is given that transcription attenuation is a frequently used mechanism of regulation.

14. Yanofsky C. Transcription attenuation: once viewed as a novel regulatory strategy. J Bacteriol 2000;182:1–8. [PubMed: 10613855]

15. Chang M, Crawford IP. The roles of indoleglycerol phosphate and the TrpI protein in the expression of *trpBA* from *Pseudomonas aeruginosa*. Nucleic Acids Res 1990;18:979–988. [PubMed: 2107533]

16. Olekhnovich I, Gussin GN. Effects of mutations in the *Pseudomonas putida miaA* gene: regulation of the *trpE* and *trpGDC* operons in *P. putida* by attenuation. J Bacteriol 2001;183:3256–3260. [PubMed: 11325956]

17. Babitzke P, Gollnick P, Yanofsky C. The *mtrAB* operon of *Bacillus subtilis* encodes GTP cyclohydrolase I (MtrA), an enzyme involved in folic acid biosynthesis, and MtrB, a regulator of tryptophan biosynthesis. J Bacteriol 1992;174:2059–2064. [PubMed: 1551827]

18. Shimotsu H, Kuroda MI, Yanofsky C, Henner DJ. Novel form of transcription attenuation regulates expression the *Bacillus subtilis* tryptophan operon. J Bacteriol 1986;166:461–471. [PubMed: 2422155]

19. Merino E, Babitzke P, Yanofsky C. *trp* RNA-binding attenuation protein (TRAP)-*trp* leader RNA interactions mediate translational as well as transcriptional regulation of the *Bacillus subtilis trp* operon. J Bacteriol 1995;177:6362–6370. [PubMed: 7592410]

20. Babitzke P, Stults JT, Shire SJ, Yanofsky C. TRAP, the *trp* RNA-binding attenuation protein of *Bacillus subtilis*, is a multisubunit complex that appears to recognize G/UAG repeats in the *trpEDCFBA* and *trpG* transcripts. J Biol Chem 1994;269:16597–16604. [PubMed: 7515880]

21. Chen G, Yanofsky C. Features of a leader peptide coding region that regulate translation initiation for the anti-TRAP protein of *B. subtilis*. Mol Cell 2004;13:703–711. [PubMed: 15023340]

22. Sarsero JP, Merino E, Yanofsky C. A *Bacillus subtilis* operon containing genes of unknown function senses tRNA^{Trp} charging and regulates expression of the genes of tryptophan biosynthesis. Proc Natl Acad Sci U S A 2000;97:2656–2661. [PubMed: 10706627]

23. Grundy FJ, Henkin TM. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. Cell 1993;74:475–482. [PubMed: 8348614]

**24. Gutierrez-Preciado A, Yanofsky C, Merino E. Comparison of tryptophan biosynthetic operon regulation in different Gram-positive bacterial species. Trends Genet 2007;23:422–426. [PubMed: 17555843]This expands and updates the comparative analysis of *trp-operon* regulation in Gram-positive organisms

25. Henderson EJ, Zalkin H, Hwang LH. The anthranilate synthetase-anthranilate 5-phosphoribosylpyrophosphate phosphoribosyltransferase aggregate. Catalytic and regulatory properties of aggregated and unaggregated forms of anthranilate 5-phosphoribosylpyrophosphate phosphoribosyltransferase. J Biol Chem 1970;245:1424–1431. [PubMed: 4315259]

26. Kane JF, Jensen RA. Metabolic interlock. The influence of histidine on tryptophan biosynthesis in *Bacillus subtilis*. J Biol Chem 1970;245:2384–2390. [PubMed: 4315152]

27. De Troch P, Dosselaere F, Keijers V, de Wilde P, Vanderleyden J. Isolation and characterization of the *Azospirillum brasilense trpE(G)* gene, encoding anthranilate synthase. Curr Microbiol 1997;34:27–32. [PubMed: 8939798]

28. Ge SM, Xie BE, Chen SF. Characterization of two *trpE* genes encoding anthranilate synthase alpha-subunit in *Azospirillum brasilense*. Biochem Biophys Res Commun 2006;341:494–499. [PubMed: 16430864]

29. Antonio RV, Creczynski-Pasa TB. Genetic analysis of violacein biosynthesis by *Chromobacterium violaceum*. Genet Mol Res 2004;3:85–91. [PubMed: 15100990]

30. Kurnasov O, Goral V, Colabroy K, Gerdes S, Anantha S, Osterman A, Begley TP. NAD biosynthesis: identification of the tryptophan to quinolinate pathway in bacteria. Chem Biol 2003;10:1195–1204. [PubMed: 14700627]

31. Morollo AA, Bauerle R. Characterization of composite aminodeoxyisochorismate synthase and aminodeoxyisochorismate lyase activities of anthranilate synthase. Proc Natl Acad Sci U S A 1993;90:9983–9987. [PubMed: 8234345]

32. de Crecy-Lagard V, El Yacoubi B, de la Garza RD, Noiriel A, Hanson AD. Comparative genomics of bacterial and plant folate synthesis and salvage: predictions and validations. BMC Genomics 2007;8:245. [PubMed: 17645794]

*33. Porat I, Sieprawska-Lupa M, Teng Q, Bohanon FJ, White RH, Whitman WB. Biochemical and genetic characterization of an early step in a novel pathway for the biosynthesis of aromatic amino acids and *p*-aminobenzoic acid in the archaeon *Methanococcus maripaludis*. Mol Microbiol 2006;62:1117–1131. [PubMed: 17010158]A nice example that new genes and new biochemical pathways remain to be discovered.

34. Kim C, Xuong NH, Edwards S, Madhusudan, Yee MC, Spraggon G, Mills SE. The crystal structure of anthranilate phosphoribosyltransferase from the enterobacterium *Pectobacterium carotovorum*. FEBS Lett 2002;523:239–246. [PubMed: 12123839]

35. Ramos I, Downs DM. Anthranilate synthase can generate sufficient phosphoribosyl amine for thiamine synthesis in *Salmonella enterica*. J Bacteriol 2003;185:5125–5132. [PubMed: 12923085]

36. Barona-Gomez F, Hodgson DA. Occurrence of a putative ancient-like isomerase involved in histidine and tryptophan biosynthesis. EMBO Rep 2003;4:296–300. [PubMed: 12634849]

*37. Merkl R. Modelling the evolution of the archeal tryptophan synthase. BMC Evol Biol 2007;7:59. [PubMed: 17425797]A detailed bioinformatic logic is presented for proposed evolutionary and functional relationships between the tryptophan synthase subunits of different archeal species.

38. Xie G, Forst C, Bonner C, Jensen RA. Significance of two distinct types of tryptophan synthase beta chain in Bacteria, Archaea and higher plants. Genome Biol 2002;3:RESEARCH0004. [PubMed: 11806827]

39. Parsot C. A common origin for enzymes involved in the terminal step of the threonine and tryptophan biosynthetic pathways. Proc Natl Acad Sci U S A 1987;84:5207–5210. [PubMed: 3110785]

40. Xie G, Bonner CA, Brettin T, Gottardo R, Keyhani NO, Jensen RA. Lateral gene transfer and ancient paralogy of operons containing redundant copies of tryptophan-pathway genes in *Xylella* species and in heterocystous cyanobacteria. Genome Biol 2003;4:R14. [PubMed: 12620124]

**41. Brune I, Jochmann N, Brinkrolf K, Huser AT, Gerstmeir R, Eikmanns BJ, Kalinowski J, Puhler A, Tauch A. The IclR-type transcriptional repressor LtbR regulates the expression of leucine and tryptophan biosynthesis genes in the amino acid producer *Corynebacterium glutamicum*. J Bacteriol 2007;189:2720–2733. [PubMed: 17259312]The whole-pathway *trp* operon of *C. glutamicum* originated via lateral gene transfer, and the key information in this paper completes one of the two best examples that exist for any experimental system where functional roles and regulation are known in both the donor and recipient organisms.

**42. Farrow JM 3rd, Pesci EC. Two distinct pathways supply anthranilate as a precursor of the *Pseudomonas* quinolone signal. J Bacteriol 2007;189:3425–3433. [PubMed: 17337571]The *trpEG* operon of *P. aeruginosa* originated via lateral gene transfer, and the key information in this paper completes one of the two best examples that exist for any experimental system where functional roles and regulation are known in both the donor and recipient organisms.

43. Pineiro S, Olekhnovich I, Gussin GN. DNA bending by the TrpI protein of *Pseudomonas aeruginosa*. J Bacteriol 1997;179:5407–5413. [PubMed: 9286994]

44. Valbuzzi A, Yanofsky C. Inhibition of the *B. subtilis* regulatory protein TRAP by the TRAP-inhibitory protein, AT. Science 2001;293:2057–2059. [PubMed: 11557884]
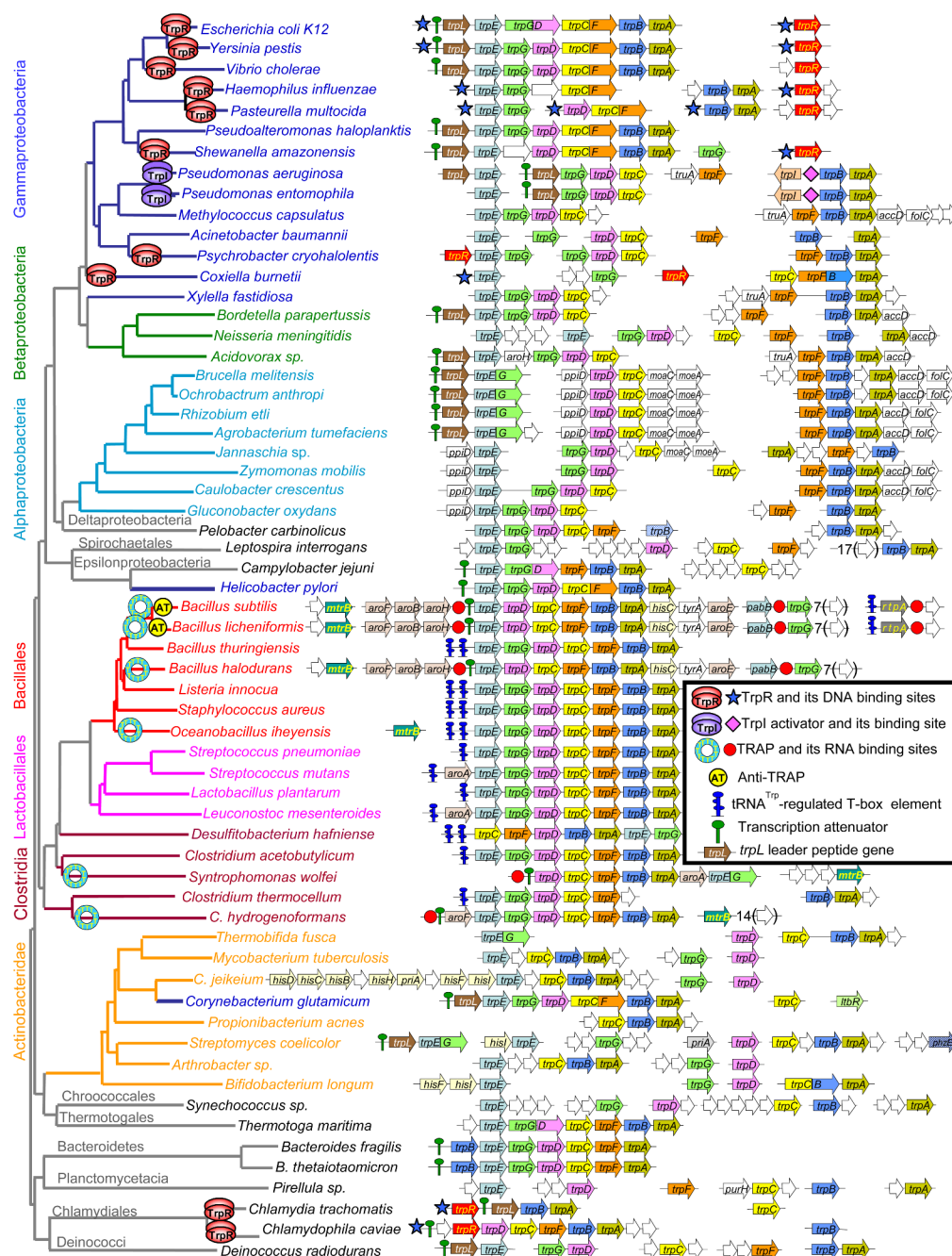
Figure 1. Operon organization and transcription regulation of *trp* genes

trp operon organization, and associated regulatory factors and elements, are "painted" on a phylogenetic tree of organisms. Only *trp* genes that function for primary tryptophan biosynthesis, not *trp* genes that participate in specialized secondary pathways, are shown. The branches for *Helicobacter pylori* and *Corynebacterium glutamicum* are color coded to indicate the LGT origins of their whole-pathway *trp* operons. A multiple alignment of the 16S rRNA sequences of representative genomes was used to estimate "genetic similarity distances" using the PROTDIST program of the phylogeny inference package program PHYLIP. Based on these estimated distances, successive clustering of lineages was performed using the neighbor-joining algorithm as implemented in the NEIGHBOR program. The regulatory DNA/RNA

binding sites, transcription attenuators, and leader peptides indicated have been identified experimentally or predicted from computer analysis of genome sequences. For large operons with more than 5 consecutive non-*trp* genes (white arrows), the number of these genes is indicated. Organisms with a tRNA[Trp]-regulated T box element controlling transcription of their *trp* operon generally also contain a tRNA[Trp]-regulated T box element controlling transcription of *trpS*, the structural gene for tryptophanyl-tRNA synthetase. *mtrB, rtpA*, and *ltbR*, are, respectively, the structural genes for the TRAP, AT, and LtrB polypeptides.
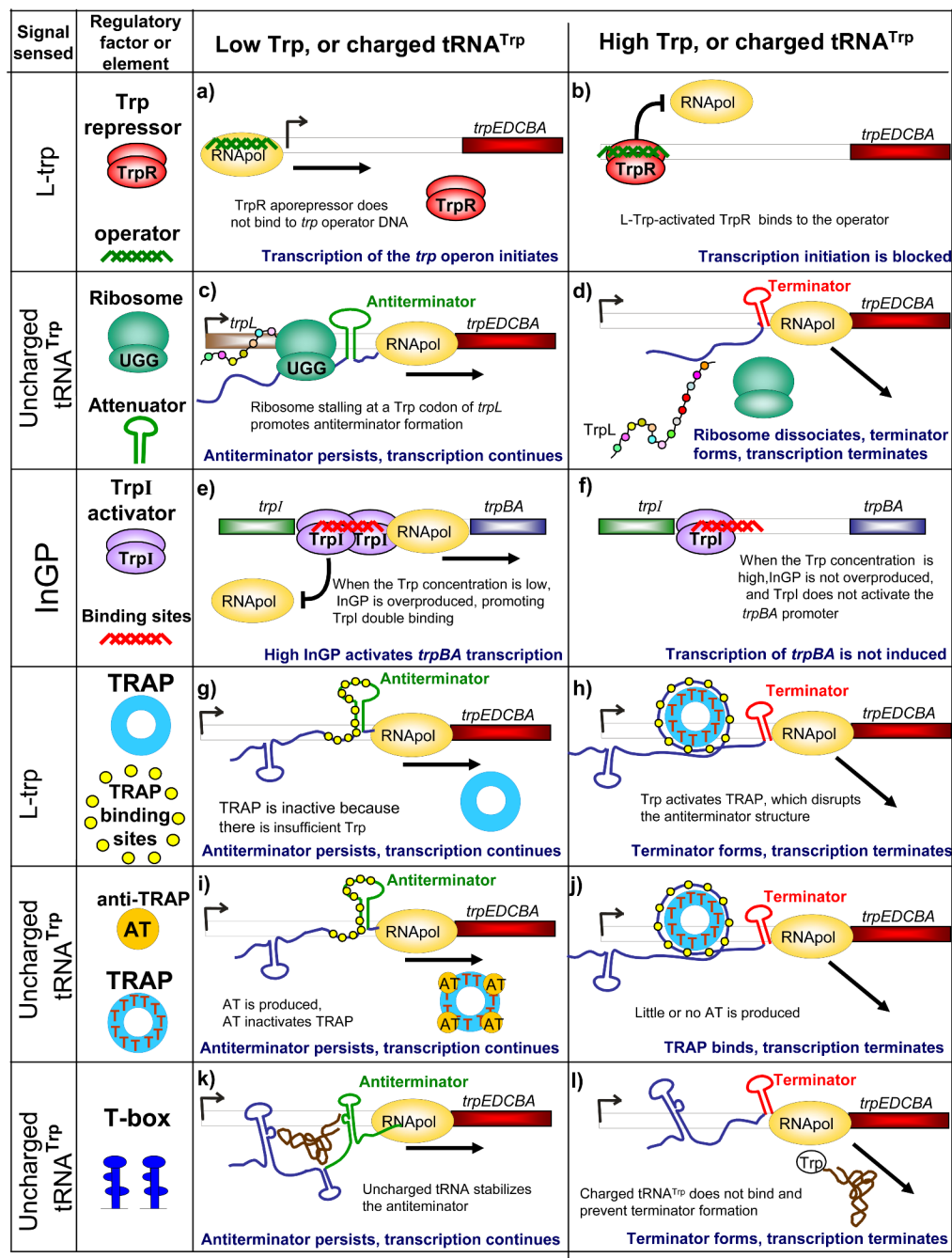
Figure 2. Molecular mechanisms used in regulating transcription of genes of the *trp* biosynthetic operon in bacteria

The variety of factors and elements participating in regulation of transcription of *trp* biosynthetic operons are shown. Some block the action of RNA polymerase, while others influence premature termination of transcription. The signals that are most often sensed are the intracellular concentration of free Trp, and the level of uncharged tRNA^{Trp}. These activate or regulate formation of different molecular components, including DNA and RNA binding proteins, RNAs, protein binding proteins, and translating ribosomes. In some instances where regulation by transcription termination has been predicted, the associated signal has not been identified. **Transcription regulation based on Trp activation of the TrpR repressor (a,**

b). The TrpR protein is found mainly in Gammaproteobacteria and in some Chlamydiales (Fig. 1). **(a)** When intracellular levels of Trp are low, TrpR exists in its inactive aporepressor conformation that cannot bind to *trp* operator DNA. Consequently, RNA polymerase can initiate *trp* operon transcription. **(b)** As the intracellular level of Trp rises, it binds to the TrpR aporepressor, altering its conformation. The repressor then binds at its cognate *trp* operators, overlapping the *trp* promoter, preventing initiation of *trp* operon transcription [10]. **Ribosome-mediated transcription attenuation (c, d).** In some organisms, the intracellular level of charged $tRNA^{Trp}$ determines whether a translating ribosome will stall at one of two adjacent *trpL* Trp codons while attempting synthesis of the TrpL leader peptide. **(c)** Under growth conditions where the charged $tRNA^{Trp}$ level is low, the translating ribosome stalls at one of the *trpL* Trp codons. Ribosome stalling favors formation of the antiterminator structure, rather than the terminator structure, hence transcription continues into the operon. **(d)** When the level of charged $tRNA^{Trp}$ is high, translation of *trpL* mRNA is completed and the translating ribosome dissociates. This allows formation of a Rho-independent transcription terminator in the transcript that terminates transcription in the leader region of the operon. **Regulation based on the activation of the TrpI protein (e, f).** The TrpI protein belongs to the LysR-family of prokaryotic regulatory proteins. Its structural gene, *trpI*, is adjacent to and transcribed divergently from the *trpBA* operon that encodes the subunits of the tryptophan synthase enzyme complex. **(e)** When the intracellular level of Trp is low, InGP, a biosynthetic intermediate and a tryptophan synthase substrate, is overproduced, favoring double binding of TrpI at operator sites in the *trpI-trpBA* intergenic region. TrpI binding activates transcription of the *trpBA* operon, either by direct interaction with RNA polymerase or by inducing DNA bending [43]. **(f)** Under growth condition where there is excess tryptophan, the InGP concentration is not sufficient to induce recognition of TrpI for its low affinity binding site, thus transcription of *trpBA* is not activated. Regulation of *trp* gene expression by TrpI appears to be restricted to *Pseudomonas aeruginosa*, *P. putida*, and *P. syringae* (Fig. 1). **Regulation based on the action of the TRAP protein (g, h).** TRAP is an RNA-binding protein found in some Bacillales, including *Bacillus subtilis*, and in a few Clostridia (Fig. 1). **(g)** Whenever Trp is growth limiting, TRAP is inactive. When TRAP cannot bind to RNA, the leader RNA of the *trp* operon folds to form the energetically favored secondary structure, the antiterminator. This structure prevents formation of a Rho-independent transcription terminator, consequently transcription proceeds into the operon's structural genes. **(h)** When there is an excess of tryptophan, TRAP is activated, preparing it to bind to the leader RNA segment of the *trp* operon transcript. This binding disrupts the antiterminator structure, favoring formation of the terminator structure, leading to premature termination of transcription in the leader region of the operon [17]. **Regulation based on the action of the anti-TRAP protein AT (i, j).** AT is produced in response to the accumulation of uncharged tRNA. **(i)** Whenever the uncharged $tRNA^{Trp}$ level is elevated, AT is synthesized and it binds to Trp-activated TRAP, inhibiting TRAP's activity [44]. **(j)** When the charged $tRNA^{Trp}$ level is high, AT synthesis is prevented, eliminating its effect on TRAP function [21,22]. **Regulation based on a T box element (k, l). (k)** Whenever the charged $tRNA^{Trp}$ level is low, uncharged $tRNA^{Trp}$ stabilizes the T box antiterminator structure, preventing formation of the terminator structure. This allows transcription to continue into the remainder of the operon. **(l)** When the charged $tRNA^{Trp}$ level is high, this charged tRNA is incapable of stabilizing the antiterminator structure. The terminator therefore forms, terminating transcription [23].