COMPARATIVE STUDIES ON THE REGULATION OF TRYPTOPHAN SYNTHESIS

Author: **Irving P. Crawford**
Department of Microbiology
University of Iowa
Iowa City, Iowa

Referee: Charles Yanofsky
Department of Biology
Stanford University
Stanford, California

INTRODUCTION

The past decade has witnessed a truly unprecedented number of fundamental discoveries in molecular biology and genetics. This has brought new insight into mechanisms of gene action and control, and this information must markedly affect the way biochemists approach the study of internal metabolism and its regulation. This review focuses on the use of recently developed genetic and biochemical methods ("recombinant DNA techniques") in analyzing the gene-enzyme relationships and regulation of a single biosynthetic pathway in microorganisms. The pathway of tryptophan synthesis is taken as the exemplar. After describing recent studies of this pathway in bacteria, the review will close with a critique of the use of these methods in evolutionary studies.

Until quite recently genetic diversity during evolution was assumed to have arisen chiefly as a result of an accumulation of point mutations, i.e., single base pair changes that, when they occur in structural genes, result in amino acid substitutions, premature termination of transcription, or frameshift events. Sporadic tandem duplications were evoked to give rise to "free" segments of raw material for such piecemeal evolution. The general similarity in the amino acid sequence of homologous proteins from related species (sometimes even distantly related ones) seemed in agreement with this mechanism, although the rather common occurrence of gaps or insertions of a few amino acids between blocks of homologous sequence[1] was somewhat disconcerting.

Recently microbial geneticists have become increasingly preoccupied with illegitimate recombinational events. By this is meant *recA*-independent recombination not requiring extensive regions of DNA homology and often associated with insertion sequences and transposons,[2] as well as in vitro recombination of unrelated DNA segments through recombinant DNA technology. These events give rise to deletions, insertions, inversions, and transpositions, the kinds of chromosomal rearrangements that have been studied in eucaryotes for a long time. It is now apparent that, at least under some conditions, the frequency of these events in procaryotes rivals the frequency of point mutations.[3] This review will examine some implications of this knowledge for the interpretation of established facts and the design of new experiments. Most of the conclusions are based on comparative analyses of the genes and enzymes of the tryptophan pathway.

There have been numerous reviews of the gene-enzyme relationships,[4,5] enzymology,[6-9] and regulation[10-12] of the tryptophan pathway in *Escherichia coli*. A review of comparative studies[5] pointed out that the structural genes for the tryptophan-specific enzymes of the pathway (whose reactions are diagrammed in Figure 1) assume a variety of chromosomal arrangements in the several major taxa of bacteria and fungi examined. Several examples of gene fusion (or fission) have been noted. Each chromosomal

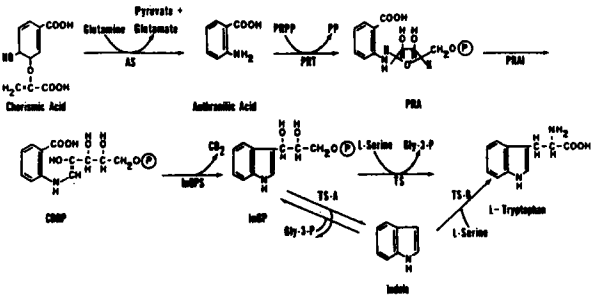

FIGURE 1. Biosynthetic pathway for tryptophan. Chorismic acid is the last intermediate common to the synthesis of aromatic amino acids and vitamins. Abbreviations: AS, anthranilate synthase (the physiological glutamine amidotransferase reaction is shown[4]); PRPP, 5-phosphoribosyl pyrophosphate; PP, pyrophosphate; PRT, anthranilate phosphoribosyltransferase; PRA, *N*-phosphoribosylanthranilate; PRAI, phosphoribosylanthranilate isomerase; CDRP, 1-(*o*-carboxyphenylamino)-1-deoxyribulose phosphate; InGP, indoleglycerol phosphate; InGPS, indoleglycerol phosphate synthaste; Gly-3-P, D-glyceraldehyde-3-phosphate; TS, tryptophan synthase; TS-A and TS-B, the A and B half-reactions of tryptophan synthase.

arrangement seems to be associated with a particular regulatory scheme, one bearing little resemblance to the other. Examples of end-product repression, substrate induction, and fixed levels of enzyme production are all found in one or another portion of the tryptophan pathway, sometimes all in the same organism. As these regulatory variations are observed in the presence of demonstrable amino acid sequence similarity,[4,5,8] it appears that a considerable remodeling of regulatory elements has occurred during evolution. The rearrangements in chromosomal location and the variations in regulatory pattern imply the frequent intervention of mechanisms that can mobilize and reinsert gene-sized or larger DNA segments.

Although in its comparative aspects the tryptophan pathway probably received earlier[13] and more extensive investigation than other metabolic pathways, only a few of the available bacterial taxa have been thoroughly investigated, and the coverage of eucaryotes is even more sparse. This review will concentrate on procaryotes; Figure 2 diagrams the chromosomal gene distributions found in several of the best studied major procaryotic taxa.

The enteric bacteria contain a fusion of two *trp* genes that are separate in all other bacterial groups investigated, *trpC* (for InGPS, the fourth enzyme), and *trpF* (for PRAI, the third enzyme). A subgroup of the enteric bacteria termed Class I, containing the genera *Escherichia, Shigella, Salmonella,* and *Enterobacter,* also possesses a second fusion involving the gene for the small, glutamine-binding subunit of the first pathway enzyme (*trpG*) and that for the second enzyme (*trpD*). The result is that the best studied tryptophan operons, those of *Escherichia coli* and *Salmonella typhimurium,* have the anomalously low number of five discrete genes for the pathway. No other well-studied bacteria have so few, although present evidence suggests that the eucaryote *Euglena gracilis* may use only two gene products to catalyze the same series of reactions.[14] All other enteric bacteria belong to Class II with respect to their tryptophan gene organization; in their case, the *trpC-trpF* gene fusion is present, but *trpG* and *trpD* are distinct. Class II includes a great variety of organisms, from *Serratia* and *Aeromonas* having GC ratios near 60%, to *Proteus* with some species having GC ratios near 40%,

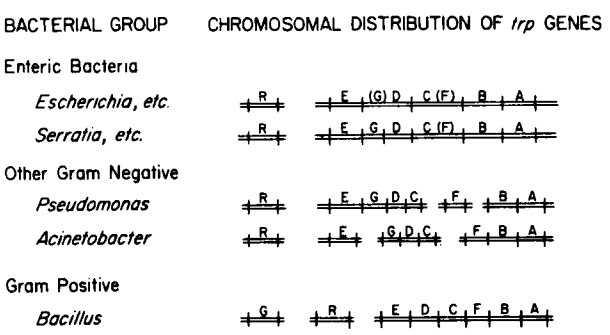

FIGURE 2. Tryptophan gene distributions in bacteria. Analogous genes are given the same letter designation throughout, but the *trp* prefix for all gene symbols has been omitted for simplicity. Gene-enzyme key (see legend to Figure 1 for enzyme abbreviations): *trpA*, TS α-chain; *trpB*, TS β-chain; *trpC*, InGPS; *trpD*, PRT; *trpE*, AS large subunit; *trpG*, AS glutamine amidotransferase subunit; *trpR*, tryptophan repressor. The gene fusions in the enteric bacteria are shown with functional components in the correct order; the component in parentheses is not ordinarily used to designate the gene. In *Acinetobacter*, the gene order in the three-gene clusters is not certain; the most probable order is shown.

and even the marine luminous species *Beneckea harveyi*.[15] A reasonable view is that Class I is a small subset of enteric organisms descended from a progenitor in which the *trpG-trpD* fusion occurred. Class II then includes all the rest of the enteric family, defined loosely enough to include *Aeromonas* and even certain marine species. All enteric bacteria seem to have the five or six structural genes for the tryptophan pathway clustered in one operon under the control of a single regulatory region.

The aerobic pseudomonads are a large and diverse group of Gram-negative bacteria. As with the enteric group, it would be surprising if all of them were to utilize a single pattern of chromosomal organization of the *trp* genes. Two members of the fluorescent pseudomonad subclass that have been studied in detail, *Pseudomonas putida* and *P. aeruginosa*, are very similar. Four genes responsible for the first, second, and fourth enzymes of the pathway are located near each other on the chromosome and are regulated by tryptophan repression. The gene for the third enzyme is separate and constitutively expressed. The two genes for the last enzyme, tryptophan synthase, are at a third location and are regulated by substrate induction.[16] There is genetic evidence that this inducibility involves autogenous regulation.[17]

Recently it has become apparent that this pattern of *trp* gene distribution and control is not the only one possessed by pseudomonads. *P. acidovorans*, investigated by Buvinger and Heath[18] and Morriss and Heath,[19] resembles more the next group of bacteria to be discussed, typified by *Acinetobacter calcoaceticus*. It is still too early to assess the taxonomic significance of this striking division in organisms presently sheltered under the *Pseudomonas* umbrella, but it may well reflect a major chromosomal and organizational difference.

The third major taxonomic cluster of Gram-negative bacteria includes *Acinetobacter, Moraxella*, and *Neisseria*;[20, 21] perhaps *Rhizobium* and *Agrobacterium* belong with this group rather than with their present companions in *Bergey's Manual*, the pseudomonads. The single organism from this group whose *trp* genes have been well studied is *Acinetobacter calcoaceticus*. This organism has its seven *trp* genes at three chromosomal locations but with a distribution different from that of the fluorescent pseudo-

monads (Figure 2). One interesting feature is that the gene for the glutamine amidotransferase subunit of anthranilate synthase (*trpG*), which also provides the same function for *p*-aminobenzoate synthase, is associated with genes for the second and fourth enxymes (*trpD* and *trpC*) and is separate from *trpE*, the gene for the large subunit of the enzyme.[22,23] These four gene products are regulated in concert[24] but not over as wide a range as those of the enteric group and the pseudomonads. The gene for the third enzyme (*trpF*) is linked to those for tryptophan synthase (*trpA* and *trpB*) but seems not to be regulated coordinately with them; the regulatory range for this cluster is even more restricted than for the other four genes, amounting to less than a sixfold difference at the extremes.

Gram-positive bacteria have not been genetically or enzymologically investigated as thoroughly as Gram-negative ones, but it is clear that they too have several different chromosomal and regulatory patterns for the *trp* genes.[5] Only the *Bacillus subtilis* pattern has been thoroughly characterized (Figure 2). In this case, all the *trp* genes but one are present in an operon regulated coordinately over as wide a range as the enteric bacterial *trp* operon is. The exception is *trpG*, the gene for the small glutamine amidotransferase subunit of anthranilate synthase.[23] This is located next to the gene for the large subunit of *p*-aminobenzoate synthase, leading to the suspicion that in this organism tryptophan synthesis has become dependent on a folate pathway gene, just as the reverse may have happened for *Acinetobacter*.[5,25]

All the aforementioned results were obtained using conventional techniques such as the selection, characterization, and mapping of auxotrophic and regulatory mutations and the assay of enzyme activity levels under various growth conditions. For the enzymes having dissimilar subunits, anthranilate synthase and tryptophan synthase, some intergroup complementation studies have been done to look for cooperative subunit interactions; these studies have yielded sporadic positive results[5] which do not seem to correlate closely with evolutionary distance as estimated in other ways. A number of amio acid sequence comparisons and immunological cross-reactivity measurements have also been performed on favorable *trp* gene products from an array of bacterial sources.[5,26,27] As far as they have been pursued these results are internally consistent, agree with generally accepted evolutionary distance estimates, and indicate the direct descent of the genes encoding enzymes of the tryptophan pathway, even those that have been involved in fusion-fission events or that have become attached to markedly different regulatory elements.

One recent study of the tryptophan synthetic enzymes in a Gram-positive organism merits special comment. *Brevibacterium flavum* is a coryniform organism used commercially to produce amino acids. Although genetic analysis of this organism is in its infancy, tryptophan auxotrophs as well as tryptophan overproducing mutants have been selected.[28] All the pathway enzymes show repression of synthesis in the presence of tryptophan, but this repression is not coordinate, except for the third and fourth enzymes.[29] When examined in extracts, the third and fourth enzyme activities were found to be on separate subunits of a complex enzyme,[30] a situation that is intermediate between the fused polypeptides seen in the enteric bacteria and the totally independent polypeptides noted in other Gram-negative bacteria and *Bacillus subtilis*. Genetic analysis of the *trp* genes in *Brevibacterium* is awaited with interest, and by virtue of the commercial importance of the organism, more information should soon be forthcoming, perhaps through application of some of the principles to be described in the following sections.

NEW FINDINGS WITH THE *ESCHERICHIA COLI TRP* REGION

The five-gene *E. coli trp* operon, with its specific control region (promoter-operator, leader, and attenuator) preceding *trpE* and a short "tail" extending beyond *trpA* to the transcription termination site,[31] comprises approximately 7000 nucleotide pairs. The acquisition of extensive knowledge of the mRNA transcript of this operon over the past decade has relied upon the existence of a group of *trp*-transducing phages derived from the lambdoid phage ϕ80 and its hybrids.[11] These phage-bacterial combinations are the result of illegitimate recombinational events. By good fortune the ϕ80 prophage attachment site (*att80*) lies quite close to the *trp* operon. None of the intervening genes is essential for growth, but one of them, *tonB*, encodes a cell surface protein used for attachment by phages T1 and ϕ80 and several colicins; this permits the application of a powerful selection for spontaneous *tonB* mutants, some of which are deletions. These circumstances have proved very useful in creating an array of ϕ80 and ϕ80-λ hybrid phages carrying segments of the *trp* operon suitable for hybridization with specific *trp* mRNA molecules. Such studies were invaluable in establishing the functional liason of the five structural genes as well as the basis for the phenomenon of polarity distal to a nonsense mutation.[11,12] These important studies, along with biochemical analysis of the bifunctional polypeptides and enzyme complexes of the pathway, make the *trp* operon one of the best known gene-enzyme systems in this well-studied organism. Even so, the successful cloning of the *trp* operon in an amplifiable plasmid vehicle in 1974[32] initiated an avalanche of new information concerning its regulation and mode of functioning. Clearly this has resulted from the availability of large quantities of pure, cleanly demarcated segments of the operon utilizable for both functional and DNA sequence analysis. The latter in particular, performed on segments from mutant and wild type cells primarily by Charles Yanofsky and his associates, has yielded a rich and unexpected harvest. Much of this information, though quite new, has been recently reviewed,[12] so it will be presented only in outline here.

The promoter-operator sequence was unambiguously located, its sequence determined, and competition for binding between active repressor and RNA polymerase at this site was clearly demonstrated.[33] The rather generous mRNA leader segment of 162 nucleotides preceding the *trpE* structural gene was sequenced and found unexpectedly to contain a strong ribosome binding and translation initiation site, followed by codons for a 14-residue peptide having tryptophans for its ninth and tenth residues.[34] This "mini-gene" precedes a conditional transcription termination sequence termed the attenuator, which functions as a second regulatory device in tandem with the operator. The attenuator is capable of aborting mRNA synthesis after the addition of 142 bases.[35] This new regulatory device, like the operator-repressor system, responds to the requirements of the cell for tryptophan, but in this case it monitors the charged tryptophanyl-tRNA level rather than the intracellular concentration of the free amino acid.[36-38] The mechanism for modulating optional transcription termination at the attenuator is intricate, almost certainly involves ribosomal movement along the 14-residue leader peptide gene, and is related to the choice of two alternative secondary structures assumed by the freshly synthesized leader mRNA segment.[39] Similar leader peptides with even more striking numbers of clustered codons for the appropriate amino acid have been found upon sequencing analogous regions preceding the *his* operon in *E. coli*[40] and *Salmonella typhimurium*[41] and the *pheA* gene in *E. coli*.[42]

The *trp* promoter-operator and leader sequences of *S. typhimurium*[34] and *Shigella dysenteriae*[43] are very similar to those of *E. coli*, although the latter bears a base substitution in the promoter making it a very inefficient transcription initiation site and contributing to the tryptophan auxotrophy of the strain.[44] *Serratia marcescens*, occupying a more distant position among the enteric bacterial relatives of *E. coli*, shows more divergence in the leader region but still retains a very similar promoter-operator sequence.[45] The *trp* operon of *E. coli* functions normally in the cytoplasm of *S. marcescens*.[46]

In *E. coli*, the entire complement of ribosome-binding, translation initiation sites (six, if that for the leader peptide is included) and the untranslated intercistronic regions are now known.[12,34,47-50] In instances where the products of adjacent genes form tightly bound enzyme complexes (*trpE-trpD* for *E. coli*, *Shigella dysenteriae*, and *Salmonella typhimurium*; *trpE-trpG* for *Serratia marcescens*, and *trpB-trpA* for *E. coli*), there is a one base overlap between the termination codon and the initiation codon, viz., UGĀUG. This doubly constrains the base sequence coding for the last few residues of the first gene product, for it must allow formation of a complete ribosome binding site for the following gene.[51] In instances where the two gene products do not form a complex, the intercistronic region is nine or more nucleotide pairs long. It remains to be seen whether this result will hold for the tryptophan gene clusters of organisms other than those clearly related to *E. coli* or for multicistronic operons other than *trp*. It can be seen from Figure 2 that *A. calcoaceticus* and *Bacillus subtilis* have genes for the two subunits of anthranilate synthase at separate chromosomal locations, suggesting that the enteric bacterial situation may represent a convenience rather than an imperative for the bacterium.

Although the best studied enteric bacteria, *E. coli*, *Salmonella typhimurium*, *Shigella dysenteriae*, and *Enterobacter aerogenes*, all have Class I *trp* operon organization, one case of an operon with Class II organization has been compared with them at the region of greatest interest, the *trpG-trpD* junction. It was postulated rather early (reviewed in References 5 and 6) that a mutation obliterating a stop codon at the end of *trpG*, thereby allowing continued reading through the *trpD* gene, could convert a Class II organism to a Class I organism. Recent investigation of the *trpG-trpD* intergenic region in *Serratia marcescens* offers circumstantial support for this mechanism of gene fusion, for the nucleotide sequence is capable of being converted to that of the fused genes by a few base substitutions and a single base pair deletion[53] (Figure 3).

As mentioned earlier, the *Escherichia coli* operon has been transported on a plasmid vehicle into *S. marcescens* where it is regulated normally. In the same paper[46] a similar result, effective repression by the product of the host *trpR* gene, was found for *Salmonella typhimurium*, *Enterobacter aerogenes*, and *Proteus mirabilis*. Most taxonomists place *Proteus* at an even greater distance from *Escherichia coli* than *Serratia*, a conclusion strengthened by the recent evidence that the overall chromosomal map of *P. mirabilis* differs markedly from that of other enteric bacteria.[52a,52b] From these results it may be surmised that the repressor-binding sequence of the *trp* operon is strongly conserved among the enteric bacteria. The only convincing test of repressor-operator recognition over greater evolutionary distance to date is the experiment of Nagahari et al.,[54] who constructed in vitro a plasmid carrying the *E. coli trp* operon and transferred it into *Pseudomonas aeruginosa*. *E. coli* tryptophan synthase was constitutively expressed under these conditions, showing no evidence that the *E. coli* operator responds to the *Pseudomonas trp* repressor.

There is an appreciable tail of mRNA, about 36 bases in length, following the last gene of the *E. coli* operon, *trpA*. At present, this region has no known function other

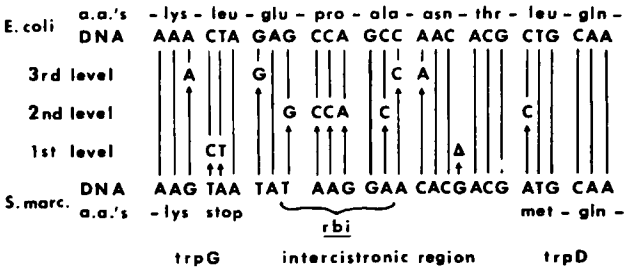

FIGURE 3. Hypothetical scheme for establishment of the Class I enteric bacterial *trpG-trpD* gene fusion. The *E. coli* and *S. marcescens* amino acid sequences and DNA coding strand sequences are shown. Changes at level 1 (involving one or the other of the two base pair substitutions in the ochre codon and a single base pair deletion) allow continuous translation through the intercistronic region and the *trpD* gene. Changes at level 2 will abolish any internal initiation of the newly fused gene. Changes at level 3 produce less drastic or inconsequential effects. The portion of the intercistronic region designated *rbi* is the ribosome binding and initiation region. (Modified from Miozzari, G. and Yanofsky, C., *Nature (London)*, 277, 486, 1979.

than transcription termination.[12,31,48] The extent of the tail does facilitate the attachment of other genes such as *lacZ* to the end of the *trp* operon, however, where they participate fully in the regulatory responses of the five *trp* genes.[55] Although these *trp-lac* fusions are artifically created and would doubtless soon be eliminated in the natural state, they neatly illustrate the point that illegitimate recombinational events can easily expand the number of individual genetic elements sharing a particular regulatory region. Clearly this fact must be taken into account when attempting to evaluate the significance of chromosomal and pathway regulatory variations in evolutionary terms.

OTHER BACTERIA STUDIED BY RECOMBINANT DNA METHODS

In Vivo Constructions

Recently it has become apparent that plasmids other than the F factor of *E. coli* are capable of both mobilizing chromosomal segments for transfer during conjugation and incorporating segments of the chromosome within themselves. When the plasmid is an antibiotic resistance transfer factor, the latter process is termed R-prime formation, in analogy with F-prime[56] formation by the classical F factor. If, as is sometimes the case, the R factor has a broad host range, the occurrence of an R-prime places a powerful tool in the hands of the geneticist. With it the chromosomal genes of one organism can be transported to the cytoplasm of very different organisms where they can replicate and function (if conditions permit) without the necessity of recombining with the chromosome of the host. Several examples of such R-primes allowing the transfer of genes from one Gram-negative organism to another have involved resistance transfer factors of the P1 class, such as RP4, R68.44, and R68.45. For example, the expression of the *nif* genes of *Klebsiella pneumoniae* has been studied in several bacterial genera.[57,58] Other examples that involve *trp* genes are the transfer of the *trpAB* gene pair in one case[59a] and the *trpEG* pair in another[59b] from *P. aeruginosa* into *E. coli*, and the transfer of *trpE, trpDC,* and *trpFBA* fragments from *Rhizobium meliloti* to *R. leguminosarum*[60] and from there to *P. aeruginosa* and *E. coli*.[61a] The first case is a reversal of the experiment[54] where the *trp* operon of *E. coli* was inserted into a class

P1 plasmid in vitro and moved into *P. aeruginosa*, where it functioned constitutively and at a high level. In the case of the *P. aeruginosa* genes in *E. coli*, however, both gene pairs examined produced little enzymic activity at first, although they could be improved by mutations that proved to be plasmid-borne.[59] In the last case, the *Rhizobium* genes functioned well in the second *Rhizobium* host, poorly in *P. aeruginosa*, and not at all in *E. coli*.[56] With the rhizobial *trpFBA* R-prime, however, a mutation located on the plasmid evoked some activity in *E. coli* and also improved the activity in *P. aeruginosa* without noticeably decreasing their function when returned to *Rhizobium*. Regulation of the *E. coli trp* operon in *Pseudomonas* cytoplasm was nil,[54] in sharp contrast to the situation when it was transferred to other members of the enteric bacterial group.

The *Pseudomonas trpAB* pair, on the other hand, was regulated by indoleglycerol phosphate induction in *E. coli* just as in its home cytoplasm.[59a] As the chromosomal segment in this R-prime is rather large (75 Mdaltons), the significance of this observation with respect to contribution of host factors to the regulatory apparatus, as well as the putative autogenous nature of the regulation, remains to be determined. No regulatory information is available concerning the rhizobial gene clusters in foreign cytoplasm.

The lesson to be drawn from these early experiments is that, even when the initial hurdle of the infrequent generation of these R-primes by illegitimate recombination is overcome, there sometimes remain severe problems of expression of foreign genetic material in the cytoplasm of distantly related hosts. The nature of these barriers to expression is not at all clear and may, in fact, not always be the same.[61b] Not all promoters active in one species may function well in a distantly related host. Once initiated, RNA synthesis may terminate prematurely, or the mRNA molecule may bind foreign ribosomes inefficiently for initiation of protein synthesis. The use of specific codons among several devoted to a single amino acid could vary between different organisms, resulting in translational problems. Specific nucleic acid or protein modification introduced by the new host may impair function. Moreover, there may be great differences in the stability of mRNAs and proteins designed by evolution for one organism and then transposed into the cytoplasm of an evolutionarily distant one. With all the potential pitfalls, however, the technique does provide an approach to previously unapproachable problems that should be exploited and can be used hand-in-hand with in vitro synthetic methods.

In Vitro Constructions

Most of the work described in the earlier section on enteric bacteria involved hybrid plasmids constructed from restriction enzyme-cleaved and ligated components. Experiments of this kind require organisms which have well-established vectors and a transformation system to recover the hybrids constructed. The experiments described in the preceding section where the *E. coli trp* operon was inserted into the plasmid RP4[54] showed that in vitro constructions can be used with wide host range vehicles as well. In this case the choice of vector was especially apt, for the same hybrid plasmid showing constitutive expression of the *trp* genes in *P. aeruginosa* could be transferred by conjugation back to *E. coli*, where the operon exhibited normal regulation.[54] Clearly this approach is capable of wide application and considerable refinement. Ordinarily a defined DNA segment can easily be recovered from the hybrid vector; it can then be shortened successively by in vitro methods until a particular function is lost or a regulatory region removed. Once crucial areas of the cloned segment are identified, their sequence can be determined. This combination of functional and structural analysis can reveal evolutionary divergence in detail not attainable heretofore.

The technique of inserting foreign DNA segments into suitable cloning vehicles

through in vitro manipulation is not restricted to conjugative bacteria. Most bacteria have the potential to become lysogenized with specific temperate phages. When investigated, many of these phages should prove to have dispensable segments of their genome that can be replaced by exogenous DNA. Moreover, a majority of the bacterial species studied have been found to harbor plasmids. Though many of these are cryptic, carrying no known selectable markers, techniques to make use of them as vehicles may not be long in appearing. Introduction of transposable antibiotic resistance markers would be an especially convenient means to this end. As an example of the opportunities such approaches can provide, some recent studies with the *trp* genes of Gram-positive, spore-forming bacilli will be described.

Although there exist temperate phages for *Bacillus subtilis* that are potentially utilizable as cloning vehicles,[62] and certain experiments have been done with enteric bacterial plamids into which *B. subtilis* phage genes have been inserted,[63] the most encouraging *Bacillus* gene cloning results to date have been obtained with a series of nonconjugative plasmids. So far the plasmids harbored by *B. subtilis* and its close relatives are either cryptic or have no selectable intrinsic markers.[64] Recently, however, Ehrlich[65] discovered that antibiotic resistance factors from *Staphylococcus aureus* can be introduced by transformation into *B. subtilis* and stably maintained there. This has provided new and convenient vehicles for recombinant DNA research. Subsequent work in several laboratories[66,67] showed that such plasmids make suitable cloning vehicles, that antibiotic resistance markers active in *S. aureus* and *B. subtilis* are also functional in *E. coli*[67] but that similar markers from *E. coli* R factors do not necessarily confer antibiotic resistance on *B. subtilis*, and that pUB110, a 3.0-Mdalton *S. aureus* plasmid conferring kanamycin/neomycin resistance, is a multicopy, amplifiable vehicle in *B. subtilis* having single cleavage sites for the restriction endonucleases *Eco*RI, *Xba* I, *Bam*HI, and *Bgl* II.[66] This work also showed that chimeras constructed with such a vehicle often sustain some subtractive shortening, presumably during their introduction or propagation in vivo.

Keggins and co-workers[68,69] showed that "shotgun" cloning of *Eco*RI digests of *B. pumilis*, *B. licheniformis*, and *B. subtilis* chromosomal DNA using pUB110 as a vehicle can provide hybrid plasmids able to complement the *B. subtilis trpC2* mutation. None of the original group contained the entire *trp* operon, as evidenced by their inability to complement certain *trp*$^-$ mutations outside *trpC*. One contained the genes *trpCFBA* in a functional state, another the *trpEDCF* group, and others still smaller segments of the operon. These do not appear to be reassemblies of an operon containing two or more *Eco*RI cleavage fragments, but rather represent the end products of the deletion and shortening events mentioned above, for in two cases only a single *Eco*RI cleavage site was found in the constructed plasmid although the chimeras were larger (by about 1.7 and 2.6 Mdaltons) than the original vector.[68] None of the plasmids complemented *aroB* or *tyrA* mutations; these are genes known to flank the *trp* operon closely. Hybrid plasmid DNA was shown to be able to transform a *trp*$^-$*recE4* strain; *recE4* strains take up chromosomal DNA normally during transformation but cannot form stable recombinants. It appears that homologous DNA segments, e.g., *B. subtilis trp* genes, on the vector and chromosome can complement one another but cannot recombine in this strain. This conclusion was confirmed in an experiment[69] where a segment of the *B. pumilis trp* operon bearing a particular *trpC* mutation was combined with the replicatory region of an indigenous, cryptic *B. subtilis* plasmid. This plasmid is compatible with pUB110. When it was present in *rec*$^+$ cells harboring a pUB110-*trp* hybrid with a different *trpC* mutation, recombination between the two plasmids readily generated *trp*$^+$ cells. When the host for the plasmid pair contained *recE4*, however, no recombination ensued. Thus, it seems that reassortment of mutational lesions can be avoided

in either of two ways, by cloning homologous but nonpairing chromosomal segments from different species or by using a *recE*⁻ recipient.

When cells containing these hybrid, nonconjugative plasmids were examined for *trp* enzyme activities complementary to chromosomal *trp*⁻ mutations, modest, fixed levels of activity were found.[70] The majority of these plasmids lack the normal transcription initiation site at the head of the operon, and perhaps all do, so the means of expression of these genes remains to be shown. When this system has been developed further it should become apparent whether the *Bacillus trp* operon is a unit of function in the same sense the enteric bacterial one is. The regulatory gene *mtr* (methytryptophan resistance) in *B. subtilis* has been compared with *trpR* in *E. coli*. Proof of this and an elucidation of the mechanism of regulation of the *trp* operon, whether by repression, attenuation, or both, should be forthcoming.

The unusual circumstance of a single gene, *trpG*, being linked to *pabA* rather than the *trp* operon (Figure 2) invites speculation that *B. subtilis* has lost an ordinary *trpG* gene by deletion or inactivation and that normal functioning of the first pathway enzyme has become dependent on the very similar glutamine amidotransferase subunit of p-aminobenzoate synthase. If this scenario is correct, there should be evidence of a functional *trpG* gene in the expected location, between *trpE* and *trpD*, in relatives of *B. subtilis*, as well as some trace of its former presence in that location in *B. subtilis*. In addition, some effort to examine a wider selection of *Bacillus* species should indicate whether or not this large group of Gram-positive spore-formers represents a homogeneous natural taxon like the enteric bacteria. As *S. aureus* plasmids form the backbone of this *Bacillus* effort, it is possible that the *trp* genes of this organism, of *Micrococcus*, and even *Streptococcus* species could become objects of study. The actual range of these nonconjugative staphylococcal plasmids remains speculative, but certainly the future will bring as many transformation experiments in Gram-positive species as prudence allows.

CRITIQUE

From the foregoing it should be clear that recombinant DNA methods can be fruitfully applied to the analysis of the effect of evolution on the genetic disposition and control of pathways of intermediary metabolism. These techniques are proving useful at all stages of investigation of the tryptophan pathway, from the first analysis of the chromosomal location of the genes of the pathway (as in *Rhizobium*) to the most detailed investigations of the nucleotide sequence of structural genes and regulatory regions as well as the elucidation of gene fusion mechanisms (exemplified by studies of the enteric bacterial operon). In fact, it is difficult to imagine that studies such as the reconstruction of the *trpG-trpD* fusion event or the analysis of functional capabilities of particular genes in foreign cytoplasm could be approached in any other way. It is probably useful to stress this point as another "benefit" of the recombinant DNA methodology, to be balanced against its still, fortunately, hypothetical "risks". It is not unreasonable to suppose that as methods improve and information accumulates we may be able to reconstruct the steps in the evolution of the pathway in a present-day organism by furnishing it with genetic raw material resembling that of an ancestral form and examining the progressive changes evoked by selective forces under presumed "normal" environmental conditions for that organism.

But rather than heap praise on a technique still in its infancy and undergoing vigorous development, it may be more useful here to present a list of caveats for investigators encouraged to embark on studies of this type. Some of these pitfalls can be avoided by a wise choice of existing methods, some may be solved by the development

of new methods as work progresses, and some may be intrinsic to the process and, hence, a perennial concern.

Investigators must be alert to the possibility that perfectly normal chromosomal segments from one organism may appear to contain nonfunctional genes even when properly cloned in a suitable vehicle in a different organism.[61b] As there are now several instances of eucaryotic genes functioning coherently in procaryotes, this is not a very predictable event, but it would seem reasonable at this time to assume that the likelihood of successful functioning decreases with greater evolutionary distance between donor and host. At times this functional disability can be overcome,[61a,61b] and perhaps further study will allow a rational approach to performing the experiments in ways that will avoid it. If it results from a failure of the promoter sequence of the donor to be recognized by the RNA polymerase of the host, it may be possible to insert the gene in such a way that a surrogate promoter on the plasmid can initiate transcription. Other possibilities exist, but for now it is well to remember that mutations in or near the cloned segment may restore some level of function, and if severe selection has been employed in the recognition and propagation of the host-vehicle combination, such a mutation may have already occurred, making the cloned segment abnormal. Whenever possible it would be wise to compare the sequence of the cloned segment with the original or to reintroduce it into its home cytoplasm to observe its ability to function there successfully.

In examining the level of expression of genes on plasmids that are not stringently controlled it is well to remember that the copy number may vary with cultural conditions and with the nature of the inserted DNA segment. Where possible it would be desirable to establish regulatory relationships in systems where the foreign DNA is inserted into a stringently controlled plasmid, or still better, a temperate phage vehicle where the copy number is assured to be one per chromosome. When selection is relaxed, as in the imposition of repression by supply of the end product, chimeric plasmids can delete their inserts or even be lost from the cell, since they are often of negative selective value.[71]

Another possibility for being misled in measurements of the level of expression involves interaction of the foreign gene product with host proteins. The interaction could be either an enhancement or a diminution of activity, and the interacting component may be specific or nonspecific, like a protease. It should be kept in mind that the foreign gene product evolved in a cell with a very different internal milieu, and unexpected interactions can take place. On the other hand, the foreign gene product might be influenced by the enzymes of the same pathway in its new host. Interactions between subunits of complex enzymes across rather large evolutionary gaps have been reported.[72,73] An interaction with the inactive product of the homologous gene of the host can be imagined. In any case where apparent regulation has been observed, it should be established that the real substance varying in amount is not a product of the host genome.

When working with DNA segments from a closely related organism, the possibility of recombination with the host genome exists. Recombination between *E. coli* and *Salmonella typhimurium* is infrequent but well documented.[74] Such an event could have the effect of bringing the foreign gene under control of the regulatory apparatus of the host. It should be avoided by use of *rec*$^-$ recipients whenever this is reasonable and prudent. Less closely related sequences may recombine by illegitimate means, and such events are not precluded in *rec*$^-$ hosts. After all, this is merely the reversal of the mechanism of formation of an R-prime in vivo. Although not apparently a highly frequent event, if drastic selection has been required to reveal the clone being examined, it will always be necessary to show that the chimeric vehicle it contains remains

unchanged. Inserting a foreign DNA segment into the chromosome of the host may have unpredictable effects on the functioning of genes in that segment, but it will have predictable effects on the genetic behavior of the host, making future recombination with homologous segments of foreign DNA possible.[63]

Finally, the uncontrolled and unexpected tendency of certain vehicles containing foreign DNA to shorten themselves spontaneously must be emphasized. This also seems to be part of the two-edged sword of illegitimate recombination. One cannot usually maintain selection for all the functions on the inserted segment of DNA. Deletions of various sizes can place the genes of interest under the control of plasmid promoters and modify the copy number of the vehicle. If the gene of interest interacts in any way with neighboring genes, deletion of these genes and loss of their products may have drastic consequences.

Many of these forewarnings may turn out to be gratuitous, but each of them has some basis in fact or experience. As the study of cloned exogenous chromosomal DNA segments becomes more quantitative and precise, the pitfalls posed become more serious and likely. Most of them are avoidable by good methodology and adequate controls, however, and it is not likely that they will impede for long the acquisition of new and unsuspected facets in our knowledge of the evolution and control of metabolic pathways in microorganisms.

REFERENCES

1. **Dayhoff, M. O.,** *Atlas of Protein Sequence and Structure,* Vol. 5, National Biomedical Research Foundation, Silver Spring, Maryland, 1972.
2. **Franklin, N.,** Illegitimate recombination, in *The Bacteriophage Lambda,* Hershey, A. D., Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1971, 175.
3. **Starlinger, P.,** DNA rearrangements in procaryotes, *Annu. Rev. Genet.,* 11, 103, 1977.
4. **Yanofsky, C. and Crawford, I. P.,** Tryptophan synthetase, in *The Enzymes,* Vol. 7, 3rd ed., Boyer, P. D., Ed., Academic Press, New York, 1972, 1.
5. **Crawford, I. P.,** Gene rearrangements in the evolution of the tryptophan synthetic pathway, *Bacteriol. Rev.,* 39, 87, 1975.
6. **Zalkin, H.,** Anthranilate synthetase, in *Advances in Enzymology,* Vol. 38, Meister, A., Ed., John Wiley & Sons, New York, 1973, 1.
7. **Kischner, K., Weishet, W., and Wiskocil, R. L.,** Ligand binding to enzyme complexes, in *Protein-Ligand Interaction,* Jaenicke, R. and Helmreich, E., Eds., Walter de Gruyter and Co., Berlin, 1975, 27.
8. **Crawford, I. P.,** Tryptophan synthetase, in *Subunit Enzymes, Biochemistry and Functions,* Vol. 2, Ebner, K. E., Ed., Marcel Dekker, New York, 1975, 223.
9. **Miles, E. W.,** Tryptophan synthase: structure, function, and subunit interaction, in *Advances in Enzymology,* Vol. 39, Meister, A., Ed., John Wiley & Sons, New York, in press, 1978, 127.
10. **Margolin, P.,** Regulation of tryptophan synthesis, in *Metabolic Pathways,* Vol. 5, Vogel, H. J., Ed., Academic Press, New York, 1971, 389.
11. **Imamoto, F.,** Translation and transcription of the tryptophan operon, in *Progress in Nucleic Acid Research,* Vol. 13, Davidson, J. N. and Cohn, W. E., Eds., Academic Press, New York, 1973, 339.
12. **Platt, T.,** Regulation and gene expression in the tryptophan operon of *Escherichia coli,* in *Molecular Aspects of Operon Expression,* Miller, J. H. and Reznikoff, W. S., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1978, 263.
13. **Bonner, D. M., DeMoss, J., and Mills, S. E.,** The evolution of an enzyme, in *Evolving Genes and Proteins,* Bryson, V. and Vogel, H. J., Eds., Academic Press, New York, 1965, 305.
14. **Hankins, C. N. and Mills, S. E.,** A dimer of a single polypeptide chain catalyzes the terminal four reactions of the L-tryptophan pathway in *Euglena gracilis, J. Biol. Chem.,* 252, 235, 1977.

43. Miozzari, G. and Yanofsky, C., A naturally occurring promoter down mutation: nucleotide sequence of the *trp* promoter/operator/leader region of *Shigella dysenteriae* 16, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 5580, 1978.
44. Manson, M. and Yanofsky, C., Naturally occurring sites within the *Shigella dysenteriae* tryptophan operon severely limit tryptophan biosynthesis, *J. Bacteriol.*, 126, 668, 1976.
45. Miozzari, G. and Yanofsky, C., personal communication.
46. Manson, M. D. and Yanofsky, C., Tryptophan operon regulation in interspecific hybrids of enteric bacteria, *J. Bacteriol.*, 126, 679, 1976.
47. Platt, T. and Yanofsky, C., An intercistronic region and ribosome-binding site in bacterial messenger RNA, *Proc. Natl. Acad. Sci. U.S.A.*, 72, 2399, 1975.
48. Platt, T., personal communication.
49. Selker, E. and Yanofsky, C., Nucleotide sequence of the *trpC-trpB* intercistronic region from *Salmonella typhimurium*, *J. Mol. Biol.*, 130, 135, 1979.
50. Nichols, B. P., Miozzari, G., van Cleemput, M., Bennett, G., and Yanofsky, C., *J. Mol. Biol.*, submited, 1980.
51. Steitz, J. A., Genetic signals and nucleotide sequences in messenger RNA, in *Biological Regulation and Control*, Goldberger, R., Ed., Plenum Press, New York, 1978.
52a. Coazee, J. N., Extension of a chromosome linkage group of *Proteus mirabilis*, *J. Gen. Microbiol.*, 107, 155, 1978.
52b. Coazee, J. N., Genetic circularity of the *Proteus mirabilis* linkage map, *Mol. Gen. Genet.*, in press, 1978.
53. Miozzari, G. and Yanofsky, C., Gene fusion during the evolution of the tryptophan operon in Enterobacteriaceae, *Nature (London)*, 277, 486, 1979.
54. Nagahari, K., Sano, Y., and Sakaguchi, K., Derepression of *E. coli trp* operon on interfamilial transfer, *Nature (London)*, 266, 745, 1977.
55. Guarente, L., Mitchell, D., and Beckwith, J., Transcription termination at the end of the *trp* operon of *E. coli*, *J. Mol. Biol.*, 112, 423, 1977.
56. Low, K. B., *Escherichia coli* K-12 F-prime factors, old and new, *Bacteriol. Rev.*, 6, 587, 1972.
57. Dixon, R., Cannon, F., and Kondorosi, A., Construction of a P plasmid containing nitrogen fixation genes from *Klebsiella pneumoniae*, *Nature (London)*, 260, 268, 1976.
58. Cannon, F. C. and Postgate, J. R., Expression of *Klebsiella* nitrogen fixation genes (*nif*) in *Azotobacter*, *Nature (London)*, 260, 271, 1976.
59a. Hedges, R. W., Jacob, A. E., and Crawford, I. P., Wide ranging plasmid bearing the *Pseudomonas aeruginosa* tryptophan synthase genes, *Nature (London)*, 267, 283, 1977.
59b. Hedges, R. W. and Crawford, I. P., unpublished results.
60. Johnston, A. W. B., Setchell, S. M., and Beringer, J. E., Interspecific crosses between *Rhizobium leguminosarum* and *R. melilote:* formation of haploid recombinants and R-primes, *J. Gen. Microbiol.*, 104, 209, 1978.
61a. Johnston, A. W. B., Bibb, M. J., and Beringer, J. E., Tryptophan genes in *Rhizobium* — their organization and their transfer to other bacterial general, *Mol. Gen. Genet.*, 165, 323, 1978.
61b. Ehrlich, S. D. and Sgaramella, V., Barriers to the heterospecific gene expression among procaryotes, *Trends Biochem. Sci.*, 3, 259, 1978.
62. Dean, D. H., Perkins, J. B., and Zarley, C. D., A potential temperate phage molecular vehicle for *Bacillus subtilis* 168, in *Spores VII*, Chambliss, G. C. and Vary, J., Eds., American Society for Microbiology, Washington, D.C., 1977.
63. Young, F. E., Duncan, C. H., and Wilson, G. A., Development of *Bacillus subtilis* model system for recombinant molecule technology, in *Recombinant Molecules: Impact on Science and Society*, Beers, R. E. and Bassett, E. G., Eds., Raven Press, New York, 1977, 33.
64. Lovett, P., Duvall, E. J. and Keggins, K. M., *Bacillus punilis* pPL10: properties and insertion into *Bacillus subtilis* 168 by transformation, *J. Bacteriol.*, 127, 817, 1976.
65. Ehrlich, S. D., Replication and expression of plasmids from *Staphyloccus aureus* in *Bacillus subtilis*, *Proc. Natl. Acad. Sci. U.S.A.*, 74, 1680, 1977.
66. Gryczan, T. J. and Dubnau, D., Construction and properties of chimeric plasmids in *B. subtilis*, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 1428, 1978.
67. Ehrlich, S. D., DNA cloning in *Bacillus subtilis*, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 1433, 1978.
68. Keggins, K. M., Lovett, P. S., and Duvall, E. J., Molecular cloning of genetically active fragments of *Bacillus* DNA in *Bacillus subtilis* and properties of the vector plasmid pUB110, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 1423, 1978.
69. Keggins, K. M., Duvall, E. J., and Lovett, P. S., Recombination between compatible plasmids containing homologous segments requires the *Bacillus subtilis recE* gene product, *J. Bacteriol.*, 134, 514, 1978.
70. Keggins, K. M., Lovett, P. S., Marrero, R., and Hoch, S. O., Insertional inactivation of *trpC* in cloned *Bacillus trp* segments: evidence for a polar effect on *trpF*, *J. Bacteriol.*, 139, 1001, 1979.

15. **Bieger, C. D. and Crawford, I. P.,** Tryptophan biosynthesis in the marine luminous bacterium *Beneckea harveyi,* manuscript in preparation.
16. **Crawford, I. P. and Gunsalus, I. C.,** Inducibility of tryptophan synthetase in *Pseudomonas putida, Proc. Natl. Acad. Sci. U.S.A.*, 56, 717, 1966.
17. **Proctor, A. R. and Crawford, I. P.,** Autogenous regulation of the inducible tryptophan synthase of *Pseudomonas putida, Proc. Natl. Acad. Sci. U.S.A.*, 72, 1249, 1975.
18. **Buvinger, W. E. and Heath, H. E.,** Organization of the *trp* genes in *Pseudomonas acidovorans, J. Bacteriol.*, submitted.
19. **Morriss, J. E., III and Heath, H. E.,** Structure of the *trp* gene clusters in *Pseudomonas acidovorans,* in preparation.
20. **Henriksen, S. D.,** *Moraxella, Acinetobacter* and the *Mimae, Bacteriol. Rev.*, 37, 522, 1973.
21. **Buchanan, R. E. and Gibbons, N. E., Eds.,** *Bergey's Manual of Determinative Bacteriology,* 8th ed., Williams & Wilkins, Baltimore, 1974.
22. **Sawula, R. V. and Crawford, I. P.,** Mapping of the tryptophan genes of *Acinetobacter caloaceticus* by transformation, *J. Bacteriol.*, 112, 797, 1972.
23. **Sawula, R. V. and Crawford, I. P.,** Anthranilate synthetase of *Acinetobacter calcoaceticus:* separation and partial characterization of subunits, *J. Biol. Chem.*, 248, 3573, 1973.
24. **Cohn, W. and Crawford, I. P.,** Regulation of enzyme synthesis in the tryptophan pathway of *Acinetobacter calcoaceticus, J. Bacteriol.*, 127, 367, 1976.
25. **Kane, J. F., Holmes, W. M., and Jensen, R. A.,** Metabolic interlock: the dual function of a folate pathway gene as an extra-operonic gene of tryptophan biosynthesis, *J. Biol. Chem.*, 247, 1587, 1972.
26. **Li, S.-L. and Hoch, S. P.,** Amino-terminal sequences of the tryptophan synthase α-chain of *Bacillus subtilis, J. Bacteriol.*, 118, 187, 1974.
27. **Reyes, G. R. and Rocha, V.,** Immunochemical comparison of phosphoribosylanthranilate isomerase-indoleglycerol phosphate synthase among the *Enterobacteriaceae, J. Bacteriol.*, 129, 1448, 1977.
28. **Shiio, I., Sugimoto, S., and Nakagawa, M.,** Microbial production of L-tryptophan. III. Production of L-tryptophan by mutants of *Brevibacterium flavum* resistant to both tryptophan and phenylalanine analogs, *Agric. Biol. Chem.*, 39, 627, 1975.
29. **Shiio, I. and Sugimoto, S.,** Altered regulatory mechanisms for tryptophan synthesis in fluorotryptophan resistance mutants of *Brevibacterium flavum, J. Biochem. (Tokyo),* 83, 879, 1978.
30. **Sugimoto, S. and Shiio, I.,** Enzymes of the tryptophan synthetic pathway in *Brevibacterium flavum, J. Biochem. (Tokyo),* 81, 823, 1977.
31. **Wu, A. M. and Platt, T.,** Transcription termination: nucleotide sequence at 3′ end of tryptophan operon in *Escherichia coli, Proc. Natl. Acad. Sci. U.S.A.*, 75, 5442, 1978.
32. **Hershfield, V., Boyer, H. W., Yanofsky, C., Lovett, M. A., and Helinski, D. R.,** Plasmid ColE1 as a molecular vehicle for cloning and amplification, *Proc. Natl. Acad. Sci. U.S.A.*, 71, 3455, 1974.
33. **Squires, C. L., Lee, F. D., and Yanofsky, C.,** Interaction of the *trp* repressor and RNA polymerase with the *trp* operon, *J. Mol. Biol.*, 92, 93, 1975.
34. **Lee, F., Bertrand, K., Bennett, G., and Yanofsky, C.,** Comparison of the nucleotide sequences of the initial transcribed regions of the tryptophan operons of *Escherichia coli* and *Salmonella typhimurium, J. Mol. Biol.*, 121, 193, 1978.
35. **Bertrand, K., Squires, C., and Yanofsky, C.,** Transcription termination in vivo in the leader region of the tryptophan operon of *E. coli, J. Mol. Biol.*, 103, 319, 1976.
36. **Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, C. L., Squires, C., and Yanofsky, C.,** New features in the regulation of the tryptophan operon, *Science,* 189, 22, 1975.
37. **Morse, D. and Morse, A. N. C.,** Dual control of the tryptophan operon is mediated by both tryptophanyl tRNA synthetase and the repressor, *J. Mol. Biol.*, 103, 209, 1976.
38. **Yanofsky, C. and Soll, L.,** Mutations affecting $tRNA^{Trp}$ and its charging and their effect on regulation of transcription termination at the attenuator of the tryptophan operon, *J. Mol. Biol.*, 113, 663, 1977.
39. **Zurawski, G., Elseivers, D., Stauffer, G. V., and Yanofsky, C.,** Translational control of transcription termination at the attenuator of the *Escherichia coli* tryptophan operon, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 5988, 1978.
40. **DiNocera, P. P., Blasi, F., DiLauro, R., Frunzio, R., and Bruni, C. B.,** Nucleotide sequence of the attenuator region of the histidine operon of *Escherichia coli* K12, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 4276, 1978.
41. **Barnes, W. M.,** DNA sequence from the histidine operon control region: seven histidine codons in a row, *Proc. Natl. Acad. Sci. U.S.A.*, 75, 4281, 1978.
42. **Zurawski, G., Brown, K., Killingly, D., and Yanofsky, C.,** Nucleotide sequence of the leader region of the phenylalanine operon of *Escherichia coli, Proc. Natl. Acad. Sci. U.S.A.*, 75, 4271, 1978.

71. **Adams, J., Kinney, T., Thompson, S., Rubin, L., and Helling, R. B.,** Frequency-dependent selection for plasmid containing cells of *Escherichia coli, Genetics,* 91, 627, 1979.
72. **Hoch, S. O.,** Tryptophan synthase from *Bacillus subtilis:* purification and characterization of the β_2 component, *J. Biol. Chem.,* 248, 2992, 1973.
73. **Patel, N., Holmes, W. M., and Kane, J. F.,** Intergeneric complementation of anthranilate synthase subunits, *J. Bacteriol.,* 114, 600, 1973.
74. **Middleton, R. B. and Mojica-a, T.,** Homology in the Enterobacteriaceae based on intercrosses between species, *Adv. Genet.,* 16, 53, 1971.