Evolution of the Tryptophan Synthetase of Fungi

ANALYSIS OF EXPERIMENTALLY FUSED ESCHERICHIA COLI TRYPTOPHAN SYNTHETASE \( \alpha \) AND \( \beta \) CHAINS*

(Received for publication, October 3, 1989)

Dennis M. Burns‡, Virginia Horn, Janet Paluh, and Charles Yanofsky§
From the Department of Biological Sciences, Stanford University, Stanford, California 94305-5020

During evolution of fungi, the separate tryptophan synthetase \( \alpha \) and \( \beta \) polypeptides of bacteria appear to have been fused in the order \( \alpha-\beta \) rather than the \( \beta-\alpha \) order that would be predicted from the order of the corresponding structural genes in all bacteria. We have fused the tryptophan synthetase polypeptides of Escherichia coli in both orders, \( \alpha-\beta \) and \( \beta-\alpha \), with and without a short connecting (con) sequence, to explore possible explanations for the domain arrangement in fungi. We find that proteins composed of any of the four fused polypeptides, \( \alpha-\beta, \beta-\alpha, \alpha-\text{con-}\beta, \) and \( \beta-\text{con-}\alpha, \) are highly active enzymatically. However, only the \( \alpha-\beta \) and \( \alpha-\text{con-}\beta \) proteins are as active as the wild type enzyme. All four fusion proteins appear to be less soluble in vivo than the wild type enzyme; this abnormal characteristic is minimal for the \( \alpha-\text{con-}\beta \) enzyme. The \( \alpha \) and \( \beta \) domains of the four fusion polypeptides were not appreciably more heat labile than the wild type polypeptides. Competition experiments with mutant tryptophan synthetase \( \alpha \) protein, and the fusion proteins suggest that in each fusion protein the joined \( \alpha \) and \( \beta \) domains have a functional tunnel connecting their \( \alpha \) and \( \beta \) active sites. Three tryptophan synthetase \( \beta'-\alpha \) fusion proteins were examined in which the carboxyl-terminal segment of the wild type \( \beta \) polypeptide was deleted and replaced by a shorter, unnatural sequence. The resulting deletion fusion proteins were enzymatically inactive and were found predominantly in the cell debris. Evaluation of our findings in relation to the three-dimensional structure of the tryptophan synthetase enzyme complex of Salmonella typhimurium (5) and the results of mutational analyses with E. coli suggest that tryptophan synthetase may have evolved via an \( \alpha-\beta \) rather than a \( \beta'-\alpha \) fusion because in \( \beta'-\alpha \) fusions the amino-terminal helix of the \( \alpha \) chain cannot assume the conformation required for optimal enzymatic activity.

Nucleotide and amino acid sequence comparisons have provided overwhelming evidence establishing that gene rearrangements have played a major role in protein evolution. In many instances, segments of genetic material encoding functionally distinct polypeptide domains appear to have been combined, thereby creating multifunctional polypeptides. Intron within coding regions and transposable elements undoubtedly facilitated such rearrangements (1, 2).

Tryptophan synthetase (TSase) exemplifies an enzyme that may have evolved via a gene (and polypeptide) fusion. In the many prokaryotic species in which TSase and its structural genes have been examined, the enzyme is composed of non-identical polypeptides encoded by adjacent genes arranged in the order \( \text{trpB-trpA} \), with transcription proceeding through \( \text{trpB} \) into \( \text{trpA} \) (Fig. 1). The polypeptides encoded in these genes, TSase \( \alpha \) and \( \beta \), respectively, form an \( \alpha-\beta \) enzyme complex containing two \( \beta \) active sites and two \( \alpha \) active sites (4, 5). The \( \alpha-\beta \) complex catalyzes the following three reactions: 1) indole-3-glycerol P + L-serine \( \rightarrow \) L-tryptophan + D-glyceraldehyde-3-P, 2) indole-3-glycerol P \( \rightarrow \) indole + D-glyceraldehyde-3-P, and 3) indole + L-serine \( \rightarrow \) L-tryptophan. The separate \( \alpha \) and \( \beta \) subunits each catalyze one of the above reactions, \( \alpha \) performing reaction 2, and \( \beta \), reaction 3. However, their rates of catalysis of these reactions is only a few percent that of the \( \alpha-\beta \) complex. TSase activity may be measured in any of the above reactions, in the presence or absence of an excess of either of the subunits.

Crystallographic studies with the TSase of Salmonella typhimurium have revealed that the individual polypeptide chains are arranged in the nearly linear order \( \alpha-\beta-\alpha \), with a tunnel connecting each \( \alpha \) active site to its adjacent \( \beta \) active site (5). This tunnel presumably facilitates diffusion of the hydrophobic reaction product, indole, from its site of synthesis in the \( \alpha \) subunit, to the active site of the \( \beta \) polypeptide, where it is combined with L-serine to form L-tryptophan (5). In some bacterial species, \( \text{trpB} \) and \( \text{trpA} \) overlap, and translational coupling insures equivalent synthesis of the two polypeptides (6, 7).

In fungi, the \( \text{trpB} \) and \( \text{trpA} \) coding regions are fused and the single TSase polypeptide dimerizes to form the TSase complex (8-12). Of particular evolutionary interest, the fungal genetic segments homologous to the two bacterial structural genes have the order \( \text{trpA-trpB} \), opposite that observed in all bacteria (Fig. 1) (8-15). In several bacterial species, such as Escherichia coli, insertion of a single base within the \( \text{trpB} \) stop codon would fuse the two polypeptides in the \( \beta-\alpha \) order (14, 16). Yet, comparison of the sequences of bacterial and fungal TSase polypeptides indicates that during TSase evolution, gene fusion produced an \( \alpha-\beta \) polypeptide rather than an \( \alpha-\beta-\alpha \) dimer (17).
a $\beta \alpha$ polypeptide. Interestingly, it has recently been shown that in at least one eukaryote, Arabidopsis thaliana, there is a gene encoding a discrete TSase $\beta$ polypeptide (17), indicating that its TSase enzyme complex probably consists of separate TSase $\alpha$ and $\beta$ polypeptides (18).

Previous investigations provided some insight into the apparent gene rearrangement and gene fusion that occurred during evolution of the fungal TSase. Comparison of the predicted amino acid sequences of the TSase polypeptides of Saccharomyces cerevisiae and Neurospora crassa with those of bacteria revealed that the fused fungal polypeptides contain an "amino acid connector" (9, 10). In the yeast and Neurospora TSase polypeptides, approximately 70 residues join regions homologous to the bacterial $\alpha$ and $\beta$ polypeptide chains (9, 10). Mutational studies performed with the yeast enzyme showed that deletion of an 18-residue segment of this amino acid connector inactivated the enzyme. However, replacing the deleted segment by an unrelated sequence partially restored enzymatic activity (15). On the basis of these and other findings, it was hypothesized that the connector's role in the yeast polypeptide was to allow the adjacent $\alpha$ and/or $\beta$ domain segments sufficient freedom to fold correctly, so they could catalyze their respective reactions (15). Thus, the fungal TSase polypeptide could have evolved with an $\alpha\beta$ rather than a $\beta\alpha$ domain order because fusion of bacterial $trpB$ and $trpA$ genes in their present order would have specified an inadequate connector. In the initial test of this prediction, the E. coli $trpB$ and $trpA$ genes were fused in their natural orientation (16). Two fusions were constructed, each containing a different residue joining the $\beta$ and $\alpha$ polypeptides. Both fusion $\beta\alpha$ proteins were enzymatically active in the various TSase reactions, but each had abnormal properties. They were less active than the wild type enzyme (16), and, following cell disruption, a significant amount of each was found associated with cell debris. In addition, large enzyme aggregates were observed, indicating that abnormal association of fusion polypeptides had occurred (16). Any of these unnatural properties could have been sufficiently deleterious to have excluded direct gene fusion as the path of TSase evolution.

In this report, we describe further studies examining the significance of TSase domain order. The E. coli $trpB$ and $trpA$ genes were fused in the order $trpA$-$trpB$, i.e. the order observed in fungi. In addition, we inserted a 10-residue segment between the two polypeptides in the $\alpha\beta$ orientation and a 13-residue segment between the two polypeptides in the $\beta\alpha$ orientation. We find that the TSase $\beta$-con-$\alpha$ protein behaves much like the TSase $\beta\alpha$ polypeptide. It has slightly reduced catalytic activity and it is partially associated with cell debris. In contrast, the TSase $\alpha\beta$ and $\alpha$-con-$\beta$ fusion proteins appear to be fully active. However, like the TSase $\beta\alpha$ fusion protein, they are partly associated with cell debris. The $\alpha$-con-$\beta$ is the most soluble of the fusion proteins that were examined. Heat inactivation studies with the fusion proteins show that with the exception of TSase $\beta$-con-$\alpha$, fusion does not radically alter the stabilities of the individual domains, i.e. the $\alpha$ domain is more labile than the $\beta$ domain, as is true for the wild type $\alpha\beta_2$ enzyme complex (4). However, the $\alpha$ domain of the TSase $\beta$-con-$\alpha$ polypeptide appears to be unusually stable. We also examined fused TSase polypeptides formed by three deletion derivatives in which removal of a segment near the end of $trpB$ placed $trpB$ in the same reading frame as $trpA$ (7, 51). These deletion derivatives direct the synthesis of large amounts of an inactive, debris-associated TSase fusion protein.

**EXPERIMENTAL PROCEDURES**

**RESULTS**

Construction of Fusion Strains—The deduced amino acid sequences of the fusion regions for the TSase $\beta\alpha$, $\alpha\beta$, $\beta$-con-$\alpha$, and $\alpha$-con-$\beta$ polypeptides are presented in Fig. 4. The TSase $\beta\alpha$ fusion was constructed previously (16) by introducing an Ncol site at the junction region, thereby replacing the $trpB$ stop codon by a Ser codon and inserting this single residue between the $\beta$ and $\alpha$ polypeptide sequences. To construct the TSase $\beta$-con-$\alpha$ fusion, a synthetic double-stranded DNA segment was inserted at the Ncol site (see "Experimental Procedures" and Figs. 2 and 4). The 10-residue "core" sequence of the added segment (Fig. 4) was copied from a...
The corresponding plasmid was cleaved at the unique ScaI residue, Thr, between the $\alpha$ and $\beta$ chain sequences (Fig. 4). In the product, TSase $\alpha$-con-$\beta$, 11 residues separate the carboxy-terminal Ser of the $\alpha$ chain sequence and the amino-terminal Met of the $\beta$ chain sequence. The various gene fusion regions were verified by DNA sequencing.

Analysis of Expression of the Various Gene Fusions—The trpB-trpA and trpA-trpB fusions were treated differently in experiments designed to examine their expression and the properties of fusion proteins. Inasmuch as the gene segment order in the trpB-trpA and trpB-con-trpA fusions was normal, we recombined these fusions into the chromosome by genetic exchange, replacing the corresponding wild type trpB, trpA region of the trp operon (16). The recipient strain used in these manipulations lacked a functional trp repressor (trpR), hence expression of the chromosomal trp operon containing the fusions was essentially elevated. TSase comprises approximately 0.7% of the soluble protein in extracts of this strain. TSase activity levels were compared with anthranilate synthase (the complex composed of trpE and trpD polypeptides) levels to determine whether the fusion proteins were produced in normal amounts, and the extent to which they were catalytically active. TSase activity also was measured in the presence of excess wild type TSase $\alpha$ or $\beta$, subunits. In addition, inhibition of enzyme activity by anti-$\beta$ antibodies was measured. As can be seen in Table I, the relative TSase and TSase $B$ activities of the trpB-trpA fusion strain were lower than those of the wild type strain, in agreement with previous findings (16). Consistent with this result, anti-$\beta$ antibodies were less effective in inhibiting the TSase $B$ activity of the preparation from this fusion strain (16). Presumably, there was less TSase $B$ activity per unit of immunologically reactive $\beta$. There was more TSase $B$ activity than TSase activity, again in agreement with previous studies which showed that this fusion strain produces some free wild type TSase $\alpha$ chain (16), presumably by initiation of translation at the trpA start codon. As observed before, a significant amount of the fusion protein was associated with the cell debris, and the soluble enzyme fraction catalyzed the

<table>
<thead>
<tr>
<th>Strain $^a$</th>
<th>TSase/ASase activity ratios $^b$</th>
<th>TSase in debris $^c$</th>
<th>TSase $B$ inhibition by anti-$\beta$ $^d$</th>
<th>Activity ratio $^e$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TSase</td>
<td>TSase $B$</td>
<td>TSase $A$</td>
<td>$2\mu_l$</td>
</tr>
<tr>
<td>trpB (Wild type trp operon)</td>
<td>1.22</td>
<td>1.47</td>
<td>1.53</td>
<td>3</td>
</tr>
<tr>
<td>trpB trpB-A</td>
<td>0.68</td>
<td>0.73</td>
<td>1.59</td>
<td>16</td>
</tr>
<tr>
<td>trpB trpB-con-A</td>
<td>0.70</td>
<td>0.82</td>
<td>0.74</td>
<td>30</td>
</tr>
</tbody>
</table>

$^a$ All strains are derivatives of E. coli W3110 trpA2. TSase $\rightarrow$ tryptophan synthetase; ASase $\rightarrow$ anthranilate synthase. TSase, TSase $B$, and TSase $A$, are tryptophan synthetase levels in the absence of added TSase $\alpha$ or $\beta$, with added TSase $\alpha$, and with added TSase $\beta$, respectively.

$^b$ Average of results of two or more experiments.

$^c$ Percent of total activity.

$^d$ Percent inhibition of 3 units of TSase $\beta$ activity (in the presence of excess TSase $\alpha$) by the indicated volume of antiserum.

$^e$ InGP $\rightarrow$ indole-3-glycerol phosphate; In $\rightarrow$ indole; Trp $\rightarrow$ tryptophan.

The gene segment order in the trpA-trpB and trpA-con-trpB fusions was the reverse of that in the chromosome, therefore, it was inappropriate to insert these fusions into the chromosomal trp operon. Instead, they were introduced into a multicopy plasmid between the lacZ promoter and the rpoC transcription termination region (see "Experimental Procedures" and Fig. 3). This configuration allows controlled expression from the lac promoter-operator. These plasmids were transformed into a W3110 trpA2 strain in which the entire trp operon had been deleted ($\Delta$trpEA2). Plasmid pMS421, bearing the lacZ gene, was introduced to allow controlled expression from the lac promoter.

Strains containing plasmids with the trpA-trpB and trpA-con-trpB gene fusions were grown in the presence and absence of IPTG and cell extracts prepared and assayed. Extracts of strains with these plasmids contain an excess of TSase $B$ activity (Table II) suggesting that translation may initiate at the trpB start codon as well as at the trpA start codon. A significant fraction of both fusion proteins is found in the cell debris, even when expression is low, i.e. in uninduced cultures (Table II). Note that less than 1% of the TSase activity of the control strain is in the debris. However, when the wild type enzyme is more highly expressed (pWS1) TSase activity does appear in the debris. The debris-associated activity is significantly above that of the control strain. Analysis of susceptibility to inhibition by anti-$\alpha$ antibodies suggest that both the $\alpha$-$\beta$ and $\alpha$-con-$\beta$ fusion proteins are as active as the

Insertion of a 13 residue "connector" segment into the fusion polypeptide did not increase the soluble fraction. TSase and TSase $B$ enzyme activities above those of the trpB-trpA fusion strain, and the TSase $A$ activity of the strain was slightly lower than that of the control. An appreciable fraction of the TSase $A$ activity (30%) was in the cell debris following sonication. In addition, anti-$\beta$ antibodies were slightly less effective in inhibiting the TSase $B$ activity in the fusion extracts than with wild type extracts. Activity ratios in the two enzymatic reactions were normal. If the soluble and debris-associated activities of the trpB-con trpA strain are combined the total enzymatic activity is closer to that of wild type than that of the trpB-A strain. These findings demonstrate that fusion of the two TSase polypeptides in the $\beta$-$\alpha$ order, with or without a 14-residue connector, results in an enzyme with some abnormal properties.
 Heat Inactivation

<table>
<thead>
<tr>
<th>Extract</th>
<th>% of initial activity surviving heat treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TSase</td>
</tr>
<tr>
<td>Wild type</td>
<td>24</td>
</tr>
<tr>
<td>trpB-A fusion</td>
<td>13</td>
</tr>
<tr>
<td>trpB-con-A fusion</td>
<td>63</td>
</tr>
<tr>
<td>trpA-B fusion</td>
<td>19</td>
</tr>
<tr>
<td>trpA-con-B fusion</td>
<td>24</td>
</tr>
</tbody>
</table>

* Values in parentheses indicate fold stimulation of TSase activity by added wild type TSase α (+A) or β ( + B) protein.
tunnel is formed by the interaction of the complementary domains of two fused polypeptide chains. Such an arrangement could explain the slightly lower activity of the α-β fusion protein and the tendency toward aggregation. To distinguish between these alternatives, we performed competition assays with the α chain of missense mutant trpA33 and each of the other fusion proteins. TSase a33 has the single change Glu49 → Met and binds substrate and β subunits normally (22). There is much evidence that Glu-49 is catalytically essential (5, 23). The rationale for this experiment is as follows: if each α-β tunnel is formed by the interaction of the α active site of one polypeptide with the β active site of a second polypeptide, as in the wild type enzyme complex, then added mutant a33 polypeptide should compete with the α domain of one fusion polypeptide for the β domain of a second fusion polypeptide. This would result in effective inhibition of catalysis of the α-β reaction (indoleglycerol phosphate + serine → tryptophan). On the other hand if covalent joining prevented dissociation of the active α and β domains, the mutant a33 polypeptide should be an ineffective competitor. Experiments of this type were performed, and the results obtained are presented in Table IV. It is obvious that the a33 polypeptide was appreciably more inhibitory to the wild type enzyme complex than to any of the fusion proteins. These findings strongly suggest that in each fusion protein a functional α-β tunnel is formed between segments of each covalently joined α and β domain. Comparable results for the TSase α-β fusion protein have been presented (16).

**DISCUSSION**

We have constructed in phase fusions of the *E. coli* TSase structural genes and have analyzed the characteristics of the resulting fusion polypeptides. The TSase structural genes were joined in both orders, with and without a synthetic

---

**TABLE IV**

<table>
<thead>
<tr>
<th>TSase assayed</th>
<th>Units addeda</th>
<th>Units TSase a33 addedb</th>
<th>% inhibitionc</th>
</tr>
</thead>
<tbody>
<tr>
<td>αβ2 (wt)</td>
<td>1.14</td>
<td>38</td>
<td>80</td>
</tr>
<tr>
<td>β-con-α</td>
<td>1.00</td>
<td>38</td>
<td>18</td>
</tr>
<tr>
<td>α-β</td>
<td>1.60</td>
<td>38</td>
<td>15</td>
</tr>
<tr>
<td>α-con-β</td>
<td>1.71</td>
<td>38</td>
<td>14</td>
</tr>
</tbody>
</table>

*TSase units, measured in the indoleglycerol phosphate + serine → tryptophan reaction.

*TSase A units, measured in the indole + serine → tryptophan reaction in the presence of an excess of TSase β2.

*The indicated units of active enzyme and inactive TSase a33 protein were mixed in 0.1 M Tris-HCl buffer, pH 7.8, and equilibrated for 1 h in an ice bath. The reaction components were then added, and the mixtures assayed in the indoleglycerol phosphate + serine → tryptophan reaction.
segment encoding approximately 10 connecting amino acid residues. These gene fusions produced TSase $\alpha$, $\beta$, $\alpha$, $\alpha$ con $\beta$, and $\beta$-con-$\alpha$ polypeptides. Each of these fusion polypeptides was enzymatically active. The order TSase $\alpha$-con-$\beta$ resembles the arrangement observed naturally in S. cerevisiae and N. crassa. We constructed these gene fusions to ask whether the properties of the corresponding fusion polypeptides would provide an explanation for the order of TSase structural gene segments that exists in fungi (8–14).

Previously, we fused TSase $\beta$ to TSase $\alpha$ by joining the carboxyl terminus of TSase $\beta$ to the amino terminus of TSase $\alpha$ through single connecting residues (16). The two fusion proteins that were produced were enzymatically active, but both were less active than the wild type protein. These proteins were abnormal in other respects: they appeared to be somewhat insoluble, and one was shown to form large, soluble aggregates (16). In this report, we show that insertion of a 14-amino acid residue connector or spacer at the junction of the TSase $\beta$ and $\alpha$ chain segments creates an active protein with properties resembling those of the TSase $\beta$-$\alpha$ fusion protein (16). Apparently, $\beta$-$\alpha$ fusion proteins, with or without the 14 residue connector, have one or more defects that modestly reduce catalytic activity and markedly increase aggregation or association with insoluble cell components. It is unlikely that association of the TSase $\beta$-$\alpha$ and $\beta$-con-$\alpha$ fusion proteins with cell debris is due to overexpression and the formation of inclusion bodies since insoluble proteins were observed in our single copy chromosomal constructs. The wild type enzyme produced at the same level was not associated with cell debris (Table I). The activity inhibition studies with anti-$\beta_2$ antisera provide additional evidence indicating that the TSase $\beta$-con-$\alpha$ fusion protein is altered; activity inhibition is not proportional to the amount of antisera added (Table I and other observations). The properties of these fusions suggest that joining the two E. coli polypeptides in the $\beta$-$\alpha$ order generates a somewhat abnormal enzyme.

We also examined three in phase trpB-trpA fusions in which segments near the carboxyl end of the $\beta$ chain have been deleted (Fig. 4). In the fusion studied most thoroughly, trpB'-trpA $\Delta 38$ (7), 86 base pair (1068–1153) were deleted near the end of trpB, placing the remainder of the gene in phase with the intact trpA coding region (Fig. 4). This frameshift changes the carboxyl-terminal residues of the TSase B segment of the fusion polypeptide (Fig. 4). Strains with this deletion have about 15% of the TSase A activity of comparable control strains, and no TSase B activity (7). The observed TSase A activity is probably due to a small amount of free TSase $\alpha$ chains that result from translation initiation at the normal trpA start codon. Most of the fusion proteins that is produced is found in the cell debris (Fig. 5B). The debris-associated protein has little or no TSase $\alpha$ activity, unlike the debris-associated TSase $\beta$-$\alpha$, $\beta$-con-$\alpha$, $\alpha$-$\beta$, $\alpha$-con-$\beta$, and wild type proteins, which have appreciable TSase $\alpha$ activity. The other two frameshift deletion strains should form TSase $\beta$-$\alpha$ fusion polypeptides with different residues at the end of the TSase $\beta$ chain segment (Fig. 4). These deletion strains behaved identically to trpB'-trpA $\Delta 38$. We conclude from these analyses that the catalytic capability of the TSase $\alpha$ domain is adversely affected by fusion to a TSase $\beta$ chain that has an altered carboxyl terminal segment.

Unlike the TSase $\beta$-$\alpha$, $\beta$-con-$\alpha$, and $\beta$'-con-$\alpha$ $\Delta 38$ proteins, the TSase $\alpha$-$\beta$ fusion protein containing a single connecting residue is fully active catalytically. However, this protein, when produced at low levels, also is partially associated with the cell debris of sonic extracts (Table II). Insertion of an 11-

C. Yanofsky, unpublished observation.

We would like to relate these and other findings with TSase to the three-dimensional structure of the TSase $\alpha\beta_2$ complex of S. typhimurium (5). Since the S. typhimurium and E. coli TSase $\alpha$ and $\beta$ chains are the same length, and their sequences are highly conserved, we will assume that their three-dimensional structures are identical. In this structure approximately 50 Å separate the amino terminus of the $\alpha$ chain and the carboxyl terminus of the $\beta$ chain, and about 70 Å separate the amino terminus of the $\beta$ chain and the carboxyl terminus of the $\alpha$ chain (5). Joining the terminal segments of the two S. typhimurium chains covalently without disturbing their natural structures would require addition of a 20–30-amino acid residue connector. Alignment of the TSase A and B domains of the E. coli, N. crassa, and S. cerevisiae polypeptides (Fig. 6) suggest that the N. crassa and S. cerevisiae polypeptides have 76 and 67 residue connectors joining their homologous A and B domains. Interestingly, the structural gene for the TSase of a second filamentous fungus, Coprinus cinereus, has recently been cloned and sequenced (12). From its deduced amino acid sequence it is evident that the domain order is identical to that of the yeast and N. crassa polypeptides, and a connector analogous to the yeast and N. crassa connector joins the TSase A and TSase B domains (12). This connector is not homologous to the connector of either of the other species (12).

We do not know the minimum required length of the connector. Deletion of 18 residues of the yeast connector inactivates the enzyme (15). If the E. coli chains were fused in the $\alpha$-$\beta$ order with no additional connecting residues the non-homologous $\alpha$ chain carboxyl-terminal segment, 31 residues, plus the non-homologous $\beta$ chain amino-terminal segment, 8 residues, would be 28 residues shorter than the S. cerevisiae connector (Fig. 6). However, if the carboxyl- and amino-terminal ends of the TSase polypeptides that were ancestral to the yeast polypeptide were as long as those of Brevibacterium lactofermentum and Caulobacter crescentus (Fig. 6), 45 and 19 residues, respectively, only 3 additional connector residues would be needed to equal the length of the 67-residue connector of the yeast polypeptide. Since the domain connectors constructed in this study are shorter than
that of the yeast polypeptide (40 and 50 versus 67 residues), it is likely that the structure of the amino or carboxyl-terminal segments of the E. coli α and β chain domains are distorted somewhat in each fusion polypeptide. Nevertheless, the α-β and α-con-β fusion proteins are fully active catalytically.

The lack of sequence conservation at the ends of the TSase polypeptides (and in the fungal connectors) suggests that residues in these segments do not have important functions. In the three-dimensional structure of the S. typhimurium enzyme, none of these segments appears to contribute to the active site of either of the polypeptides (5). Mutational studies with E. coli TSase α are consistent with this conclusion since among the many totally defective missense mutants that have been isolated only a few have inactivating amino acid replacements in these segments (24). The amino acid change Tyr-4 → Cys appears to radically reduce TSase α1 combining activity but the inherent TSase α activity of the Tyr-4 → Cys polypeptide is only slightly low (25). The change Phe-22 → Ser has been reported to eliminate inherent TSase α activity while having no effect on the TSase β1 combining activity (25). The importance of Phe-22 has also been suggested by one exceptional case in which the substitution Phe-22 → Leu inactivates TSase α when an unnatural residue, Ser, replaces the natural residue, Gly, at position 211 in the polypeptide (26). The polypeptide with Leu-22 and Gly-211 is partially active (20). These observations suggest that perhaps the first 20 residues of TSase α of E. coli are nonessential.

Deletion, gene fusion, and mutational analyses provide some information on the essentiality of the carboxyl-terminal segment of the TSase α polypeptide. Deletion ΔtonB-trpA229 removes about the last 30 residues of TSase α, probably replacing this segment with foreign residues; this deletion eliminates TSase A activity (24). Mutant trpA496, a nonsense mutant altered at codon 245 (the wild type TSase α chain contains 268 residues), also has no TSase A activity (24). However, trpA-lacZ gene fusions with TSase A activity have been constructed in which the sites of fusion are located within trpA but beyond the site of the A06 mutation (27). In addition, one functional trpA frameshift mutant is known in which the frameshift occurs distal to the site corresponding to codon 234 of trpA (28). These findings suggest that perhaps the last 10–20 residues of the TSase α polypeptide are not necessary for catalytic activity.

Mutational studies with TSase β have identified one inactivating missense change near one end of the polypeptide, 16 residues from the carboxyl terminus (29). In addition, introduction of a stop codon in trpB that should result in production of a polypeptide lacking the last 23 residues also eliminates enzyme activity (7). However, one uncharacterized deletion mutant exists which lacks an unknown number of residues from the carboxyl end of the β chain, but produces an active, though labile polypeptide. Inactivating mutations altering the amino-terminal end of the β chain are not known, but replacement of the first 11 residues of TSase β of Saccharomyces cerevisiae by fusion to other proteins yields an active TSase β fusion polypeptide. The first 25 residues of TSase β of E. coli and S. typhimurium are identical (30). These findings suggest that at least the first 11 and last 10 or so residues of TSase β are not essential for enzyme activity.

Evidently, to be able to answer the fundamental question posed in this study, we must learn the specific structural or functional roles played by the ends of the E. coli TSase α and β polypeptides. TSase α is a member of a family of proteins with an 8-fold α/β barrel structure, called the TIM barrel (5, 31), in which α helices form the external staves of the barrel and parallel β strands occupy its interior. The amino- and carboxyl-terminal segments of the TSase α chain form external helices 0 and 8, respectively (5). Helix 0 is not present in most structurally similar TIM barrel polypeptides (5, 32); however, indoleglycerol phosphate synthase, the enzyme catalyzing the immediately preceding reaction in the tryptophan pathway, also has a TIM barrel structure with an extra amino-terminal helix (33). In TSase α, helix 0 caps the bottom of the barrel and may shield Glu-49, the key catalytic residue, from solvent (5, 22). This may raise the pK of Glu 49 by placing it in a hydrophobic environment, which would allow Glu-49 to act as a base during catalysis (22). In this regard, we determined the pH optimum of the TSase β-con-α protein in the indoleglycerol phosphate + serine → tryptophan reaction, but it was indistinguishable from that of the wild type enzyme. Interestingly, helix 0 is encoded by a separate exon in the TSase structural gene of N. crassa (10). The role of the carboxyl-terminal segment of TSase α, helix 8, is not apparent (5).

In the TSase β polypeptide, the amino-terminal segment has a random coil configuration while the carboxyl-terminal segment is helical. The role of the amino-terminal segment is unknown. The carboxyl-terminal segment may contribute to interactions between the two β chains in stabilizing the β2 subunit.

From the preceding discussion it is apparent that covalent joining of the two E. coli TSase chains, in either order, without addition of an appropriate connector, would probably distort the structure of the ends of either or both of the connected chains. The findings that both TSase β-α and α-β fusion proteins are enzymatically active in the indole → tryptophan and indoleglycerol phosphate → tryptophan reactions suggests that any structural distortion resulting from either fusion does not radically affect either active site. Furthermore, the competition experiments with mutant TSase α33 indicate that each fusion polypeptide probably can form a tunnel connecting its α and β active sites.

The properties of the fusion proteins suggest that the order α-β, observed naturally, is compatible with a more normal TSase enzyme complex than the order β-α. Perhaps, as we suspect, the amino-terminal segment of the α chain, in its normal conformation, plays an important role in establishing a structurally desired feature of this polypeptide. In the E. coli TSase β-α and α-β fusion proteins, this amino-terminal arm may be incapable of properly performing this function. Additional conjecture concerning the advantages and disadvantages of the two fusion orders must await additional information on the biological essentiality of the terminal segments of the bacterial TSase polypeptides, particularly helix 0 of TSase α.

Acknowledgments—We are grateful to Barry Hurlburt and Paul Gollnick for their critical comments on the manuscript. We would also like to thank Edith Miles and C. Craig Hyde for sharing their special insight, and the latter for permission to cite his interpretations of the three-dimensional model of the enzyme.

REFERENCES

C. Hyde, personal communication.
39. Rev. Microbial. 40, 55–77
Tryptophan Synthetase Evolution

As part of the construction of a dual-light-gene fusion (see below), a 452 bp segment between the coding sequences of the trypt and tetA genes contained on plasmid pBR322 was inserted into the TSM1/MTHF1 hybrid plasmid pMTHF1. The plasmid pMTHF1 also contained the 3′- untranslated region (UTR) of the trypt gene. The construction of the dual-light-gene fusion was by ligating the 452 bp fragment into the 3′-flanking DNA of the trypt gene. The resulting construct was used to transform competent E. coli cells, and the resulting plasmid DNA was isolated. The plasmid DNA was then used to transform competent E. coli cells, and the resulting plasmid DNA was isolated. The plasmid DNA was then used to transform competent E. coli cells, and the resulting plasmid DNA was isolated. The plasmid DNA was then used to transform competent E. coli cells, and the resulting plasmid DNA was isolated. The plasmid DNA was then used to transform competent E. coli cells, and the resulting plasmid DNA was isolated.
Tryptophan Synthetase Evolution

by guest on September 13, 2017

http://www.jbc.org/ Downloaded from

Figure 1. Construction of the tryptophan and tRNA-utilizing tRNA genes and the organization of expression plasmids. See Experimental Procedures for details not included in the figure. Plasmids are not drawn to scale. The sizes of pGEM3 (6.0 kb), pGRM (9.0 kb), pGEM2 (9.67 kb), pGRM2 (9.5 kb), pGRM3 (9.64 kb), and pGRM4 (9.57 kb) are shown. The thin lines represent the vector moieties of each plasmid, while the thick lines represent the insert DNA. Only relevant restriction sites at each step of the construction are shown. C = SacI, K = KpnI, S = SalI, C1 = C1, N = Ncol, H = HindIII, and E = EcoRV. The sequence of the connector-encoding segment, formed by annealing two complementary 30-mers to give the 60-mers, is

Note that when this 60-mer was inserted into pGRM, three sites were not generated in the pGRM derivative due to the lack of appropriate nucleotides at either the first or sixth positions of the hexamers at the junctions.

Figure 2. Construction of the trpC-control fusion gene. See Experimental Procedures for details not included in the figure. Plasmids are not drawn to scale. Their sizes are pGEM3 (6.0 kb), pGRM (9.0 kb), pGEM2 (9.67 kb), pGRM2 (9.5 kb), pGRM3 (9.64 kb), and pGRM4 (9.57 kb). The thin lines represent the vector moieties of each plasmid, while the thick lines represent the insert DNA. Only relevant restriction sites at each step of the construction are shown. C = SacI, K = KpnI, S = SalI, C1 = C1, N = Ncol, H = HindIII, and E = EcoRV. The sequence of the connector-encoding segment, formed by annealing two complementary 30-mers to give the 60-mers, is

Note that when the DNA-encoding 60-mer was inserted into pGRM, a SalI site was generated at one junction. The outcome of this event was that the 60-mer was inappropriately ligated to the 30-mer resulting in a 60-mer ligated to the 30-mer and containing a non-tRNA. The sequence of the connector-encoding segment, formed by annealing two complementary 30-mers to give the 60-mers, is
Evolution of the tryptophan synthetase of fungi. Analysis of experimentally fused
Escherichia coli tryptophan synthetase alpha and beta chains.
D M Burns, V Horn, J Paluh and C Yanofsky


Access the most updated version of this article at http://www.jbc.org/content/265/4/2060

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/265/4/2060.full.html#ref-list-1