Article

Potential DNA slippage structures acquired during evolutionary divergence of Acinetobacter calcoaceticus chromosomal benABC and Pseudomonas putida TOL pWW0 plasmid xylXYZ, genes encoding benzoate dioxygenases

HARAYAMA, Shigeaki, *et al*.

Abstract

The xylXYZ DNA region is carried on the TOL pWW0 plasmid in Pseudomonas putida and encodes a benzoate dioxygenase with broad substrate specificity. The DNA sequence of the region is presented and compared with benABC, the chromosomal region encoding the benzoate dioxygenase of Acinetobacter calcoaceticus. Corresponding genes from the two biological sources share common ancestry: comparison of aligned XylX-BenA, XylY-BenB, and XylZ-BenC amino acid sequences revealed respective identities of 58.3, 61.3, and 53%. The aligned genes have diverged to assume G+C contents that differ by 14.0 to 14.9%. Usage of the unusual arginine codons AGA and AGG appears to have been selected in the P. putida xylX gene as it diverged from the ancestor it shared with A. calcoaceticus benA. Homologous A. calcoaceticus and P. putida genes exhibit different patterns of DNA sequence repetition, and analysis of one such pattern suggests that mutations creating different DNA slippage structures made a significant contribution to the evolutionary divergence of xylX.

Reference

HARAYAMA, Shigeaki, *et al*. Potential DNA slippage structures acquired during evolutionary divergence of Acinetobacter calcoaceticus chromosomal benABC and Pseudomonas putida TOL pWW0 plasmid xylXYZ, genes encoding benzoate dioxygenases. *Journal of bacteriology* , 1991, vol. 173, no. 23, p. 7540-8

PMID : 1938949

JOURNAL OF BACTERIOLOGY, Dec. 1991, p. 7540–7548
0021-9193/91/237540-09$02.00/0

Vol. 173, No. 23

Potential DNA Slippage Structures Acquired during Evolutionary Divergence of *Acinetobacter calcoaceticus* Chromosomal *benABC* and *Pseudomonas putida* TOL pWW0 Plasmid *xylXYZ*, Genes Encoding Benzoate Dioxygenases

SHIGEAKI HARAYAMA,[1] MONIQUE REKIK,[1] AMOS BAIROCH,[1] ELLEN L. NEIDLE,[2]† AND L. NICHOLAS ORNSTON[2]*

Department of Medical Biochemistry, Faculty of Medicine, University of Geneva, 1211 Geneva 4, Switzerland,[1] and Department of Biology, Yale University, New Haven, Connecticut 06511[2]

Received 7 June 1991/Accepted 17 September 1991

The *xylXYZ* DNA region is carried on the TOL pWW0 plasmid in *Pseudomonas putida* and encodes a benzoate dioxygenase with broad substrate specificity. The DNA sequence of the region is presented and compared with *benABC*, the chromosomal region encoding the benzoate dioxygenase of *Acinetobacter calcoaceticus*. Corresponding genes from the two biological sources share common ancestry: comparison of aligned XylX-BenA, XylY-BenB, and XylZ-BenC amino acid sequences revealed respective identities of 58.3, 61.3, and 53%. The aligned genes have diverged to assume G+C contents that differ by 14.0 to 14.9%. Usage of the unusual arginine codons AGA and AGG appears to have been selected in the *P. putida xylX* gene as it diverged from the ancestor it shared with *A. calcoaceticus benA*. Homologous *A. calcoaceticus* and *P. putida* genes exhibit different patterns of DNA sequence repetition, and analysis of one such pattern suggests that mutations creating different DNA slippage structures made a significant contribution to the evolutionary divergence of *xylX*.

Benzoate dioxygenase (30) catalyzes the first of two step reactions that give rise to catechol from benzoate. Many bacterial species, exemplified by *Acinetobacter calcoaceticus*, carry chromosomal genes for enzymes that convert catechol to citric acid cycle intermediates via β-ketoadipate (20, 21, 23, 31). These enzymes do not act effectively upon methylsubstituted substrates. Methylcatechols, formed from methylbenzoates by enzymes with relatively broad substrate specificity, are utilized by a different metabolic pathway (2). Genes for this pathway frequently are carried on plasmids such as the TOL pWW0 plasmid from *Pseudomonas putida* (8, 36).

Isofunctional enzymes for dissimilation of aromatic compounds in *A. calcoaceticus* and *P. putida* generally exhibit close evolutionary ancestry reflected in identities of amino acid sequence close to or exceeding 50% (9, 23, 38, 39, 41). In contrast, genes for such isofunctional enzymes have diverged substantially, as indicated by differences in G+C content of about 15% (9, 20, 27). Such differences in G+C content may be attributed in part to directional pressure exerted by mutations within divergent cell lines (32). In addition, selection for cognate tRNAs characteristic of the host (1, 11) appears to have contributed to divergence of the G+C content in DNA because genes from either *A. calcoaceticus* or *P. putida* tend to possess a pattern of codon usage that is similar to that of other genes from closely related organisms (27, 35).

A question that remains is the nature of mutations that gave rise to divergent DNA sequences encoding homologous genes in the two bacterial species. Comparison of amino acid sequences suggested that gene conversion events causing substitution of oligonucleotide sequences within (29) and among (40) coevolving genes created sequence repetitions that were a major source of genetic divergence. The data were puzzling because they implied that DNA sequences encoding dipeptides and tripeptides were shuffled within genes during their evolutionary divergence. The evidence suggested that sequence exchange between misaligned DNA strands might have been a source of the shuffling mutations (25, 29), but in the absence of DNA sequences for the divergent genes, it was not possible to suggest specific mechanisms that created and maintained the divergent sequences.

Analysis of a range of biological systems has suggested that sequence exchange among slipped DNA strands caused mutations that have been a significant source of evolutionary divergence (17, 34). The interaction of slipped DNA strands during mutation has been documented by demonstration that sequence-directed mutations create both deletions (3–6) and repetitions of DNA sequence. Comparison of DNA sequences for oxygenative enzymes from *A. calcoaceticus* and *P. putida* suggested that slippage structures formed between misaligned DNA strands were formed by mutation and are maintained by mismatch repair during evolution (9, 21).

Elsewhere, we have reported the DNA sequences of the chromosomal *A. calcoaceticus benABC* genes, which encode a benzoate dioxygenase with relatively narrow substrate specificity; the enzymes encoded by these genes proved to be members of widely extended families of proteins associated with oxygenative processes and electron transport (22). Here we present the DNA sequences of *P. putida* TOL plasmid DNA containing the *xylXYZ* genes which encode a benzoate dioxygenase with broad substrate specificity. Comparison of the *benABC* and *xylXYZ* genes demonstrates their overall homology, reveals different patterns of sequence repetition that were acquired during their divergence, and suggests specific, potentially interactive,

* Corresponding author.
† Present address: Microbiology Department, University of Texas Medical School, Houston, TX 77225.

DNA slippage structures that were acquired during the evolution of *xylZ*.

MATERIALS AND METHODS

Sources of DNA for sequencing were pPL392 (7) and subclones derived from it (8). DNA fragments generated by cleavage with different restriction enzymes were subcloned into mp18 or mp19 and propagated in *Escherichia coli* JM101 (37). DNA sequencing with M13 derivatives as templates was done with a kit from Pharmacia. Overlapping sequences from both strands were determined. Sequence analysis was performed with PC/Gene (Intelligenetics) and Microgenie (Beckman) software packages.

Nucleotide sequence accession number. The *xyl* DNA sequence presented in this report has been deposited with GenBank under accession number M64747.

RESULTS

Overall sequence comparisons. Figure 1 presents DNA sequences containing the *P. putida xylXYZ* genes and the *A. calcoaceticus benABC* genes accompanied by the aligned amino acid sequences of the gene products. Three comparisons, XylX-BenA, XylY-BenB, and XylZ-BenC, revealed respective amino acid sequence identities of 58.3, 61.3, and 53%. Regions of amino acid sequence similarity were distributed fairly evenly through the compared regions, and thus it appears likely that the *xylXYZ* and *benABC* genes were derived from a common ancestral DNA fragment.

DNA sequences lying between the aligned genes were not subjected to selection at the level of protein, and these regions have diverged. Whereas the *benA* and *benB* genes overlap by a single nucleotide, the corresponding *xylX* and *xylY* genes do not overlap (Fig. 1). The distance between *benB* and *benC* is 66 nucleotides longer than the distance between *xylY* and *xylZ*. The *ben* and *xyl* sequences show no similarity in this region, but it would be incorrect to conclude that this segment of DNA escaped selection for structure because the *ben* sequence contains an inverted repetition 10 bp in length (Fig. 1).

G+C content and codon usage. Despite their similarity at the level of translation, the *P. putida* and *A. calcoaceticus* genes differ substantially in G+C content. Whereas the respective G+C contents of the *P. putida xylX*, *xylY*, and *xylZ* genes are 59.3, 56.1, and 61.3%, the G+C contents of the corresponding *A. calcoaceticus benA*, *benB*, and *benC* genes are 45.0, 41.2, and 47.3%, respectively. Thus, the differences in the G+C content of the *xyl* genes range over 4.2%, and the differences in the G+C content of the *ben* genes range over 6.1%. Differences in the G+C content emerging from *xylX-benA*, *xylY-benB*, and *xylZ-benC* comparisons fall within the relatively narrow range between 14.0 and 14.9%.

Variation in the G+C content of genes from different organisms has been attributed to a combination of directional mutation pressure toward a defined G+C content (32) and selection for codon usage as determined by the relative pool levels of cognate tRNAs (1, 11). Comparison of codon usage in the *xyl* and *ben* genes showed that differences are found most frequently in the third nucleotide position, which allows codons with different G+C contents to encode identical amino acid residues. For example, 60 arginyl residues expressed by the *P. putida xylXYZ* genes are encoded with respective frequencies of 57 and 17% by CGC and CGT. Conversely, the 49 arginyl residues in the *A. calcoaceticus* genes are encoded by CGC and CGT with relative frequencies of 22.4 and 59.2%, respectively.

DISCUSSION

Conservation of genes and divergence of intergenic sequences. Amino acid sequence comparisons provide strong evidence for common ancestry of the *xylXYZ* and *benABC* regions: identity of amino acid sequence in the aligned gene products (Fig. 1) exceeds 53%. Intergenic regions of DNA have not been subjected to selection at the level of protein function, and substantial divergence of DNA sequence occurred in these noncoding regions. The single-base overlap of *benA* and *benB* prompts speculation that their expression might be concerted by translational coupling (24). Whatever the merits of this proposal, it must be recognized that the putative control has not been conserved in the closely homologous *xylX* and *xylY* genes, which do not overlap (Fig. 1). Possible genetic or physiological significance might be ascribed to the 10-bp inverted repetition lying between *benB* and *benC* (Fig. 1), but selective forces favoring the inverted repetition were not conserved during divergence of *xylY* and *xylZ* between which the corresponding region of DNA is not present (Fig. 1).

Use of exceptional arginine codons in *xylX*. An unusual property of the *xylX* gene is its use of two AGA and three AGG codons for 5 of the 35 arginyl residues. These codons are employed rarely by structural genes from fluorescent *Pseudomonas* species (35). Nor are the five exceptional codons used by any of the 17 arginyl residues conserved in *xylX* during its divergence from *benA* (Fig. 1). Thus, the rare AGA and AGG codons represent 5 of 18 arginyl codons that were either acquired or maintained during divergence of *xylX* from the ancestor it shared with *benA*. It therefore is reasonable to conclude that the unusual arginyl codons in the *xylX* gene were selected. The basis for the selection is unknown. Remarkable enrichment of the exceptional arginyl codons is also found in other *xyl* genes from the TOL plasmid pWW0 (10, 12, 13, 19, 33). These observations raise the possibility that the *xyl* structural genes have been selected in an organism in which tRNA corresponding to AGA and AGG codons was relatively high.

Evolutionary acquisition of DNA slippage structures. Different patterns of DNA sequence repetition distinguish the *xyl* and *ben* regions. As discussed below, it appears that the sequence repetitions arose as a consequence of mutations that superimposed DNA sequences from different chromosomal regions during evolutionary divergence. The question that must be addressed is how to account for both the origin and the maintenance of complex patterns of DNA sequence repetition.

Acquired DNA sequence repetitions have been interpreted as components of three-dimensional structures formed by hybridization between slipped strands (25–29). According to this view, the genetic basis for the formation and the conservation of the DNA slippage structures is mismatch repair between the hybridizing regions (16, 18). The model predicts that DNA forced into single strandedness in one slippage structure may hybridize with a complementary DNA sequence in another region and thus may foster formation of an alternative slippage structure. Thus, interactive slippage structures could contribute to genetic continuity and, when disrupted, generate cascades of mutation through mismatch repair. Therefore, the model accounts for how rapid genetic divergence could occur and, once achieved, how it might be maintained.

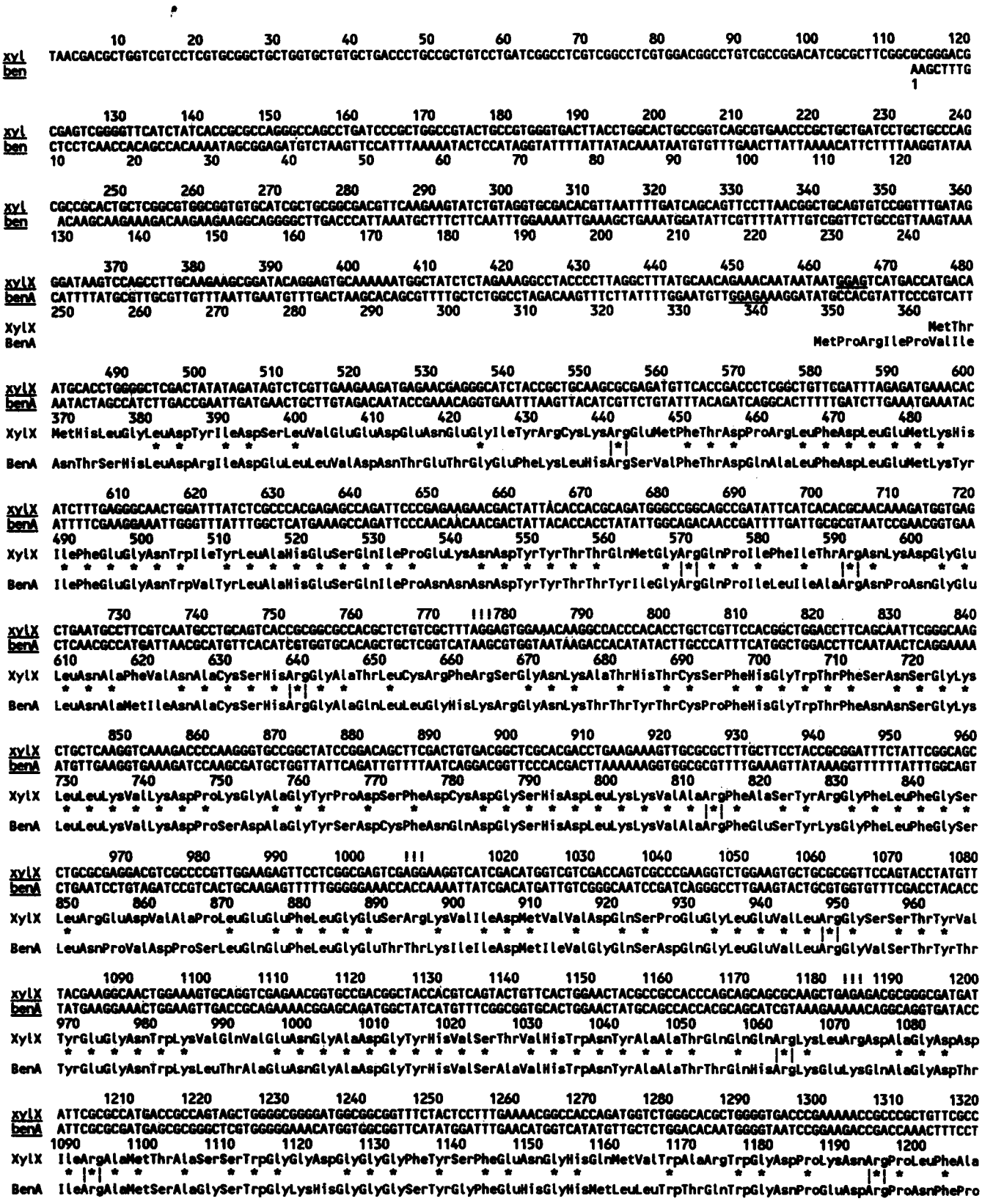

FIG. 1. Aligned DNA sequences of *P. putida xylXYZ* and *A. calcoaceticus benABC*. Potential ribosome binding sites preceding each gene are underlined. Asterisks mark amino acid residues shared in the protein products of the aligned genes. Vertical lines flank asterisks marking arginyl residues conserved in XylX and BenA. Exclamation points mark the exceptional AGA and AGG arginyl codons that appear to have been selected in *xylX*. Arrows indicate a perfect 10-bp repetition in the DNA sequence lying between *benB* and *benC*.

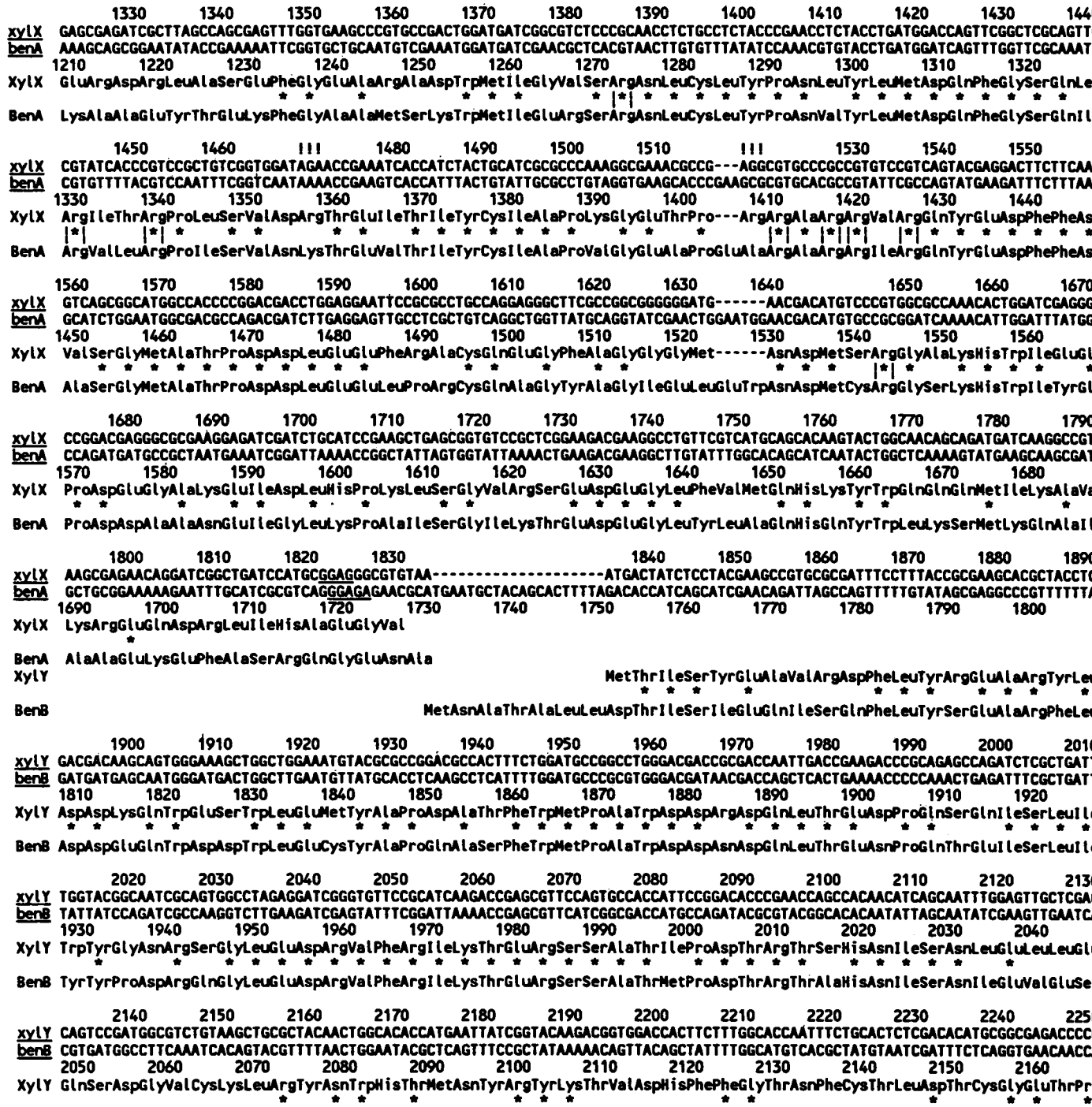

FIG. 1—*Continued.*

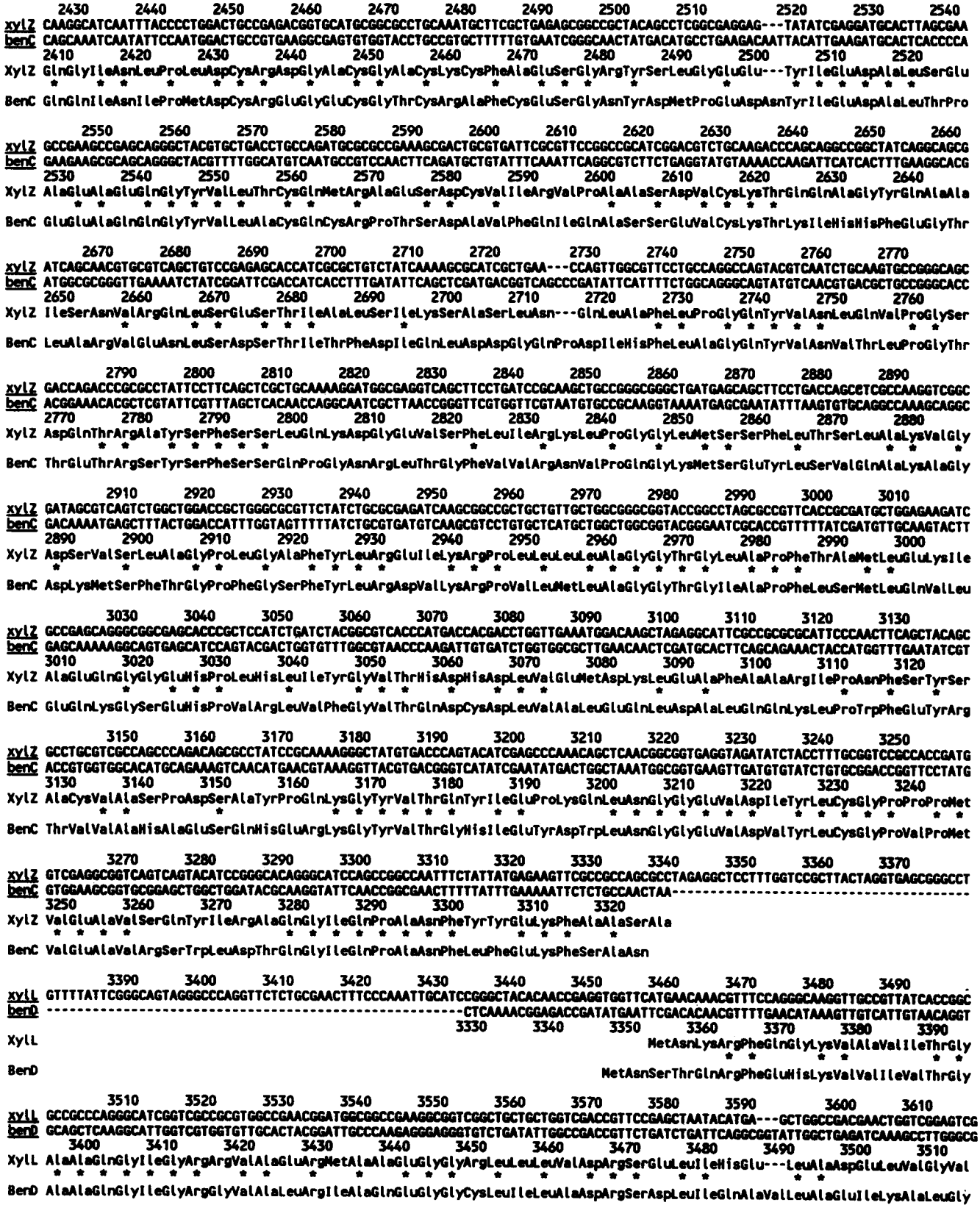

FIG. 1—*Continued.*

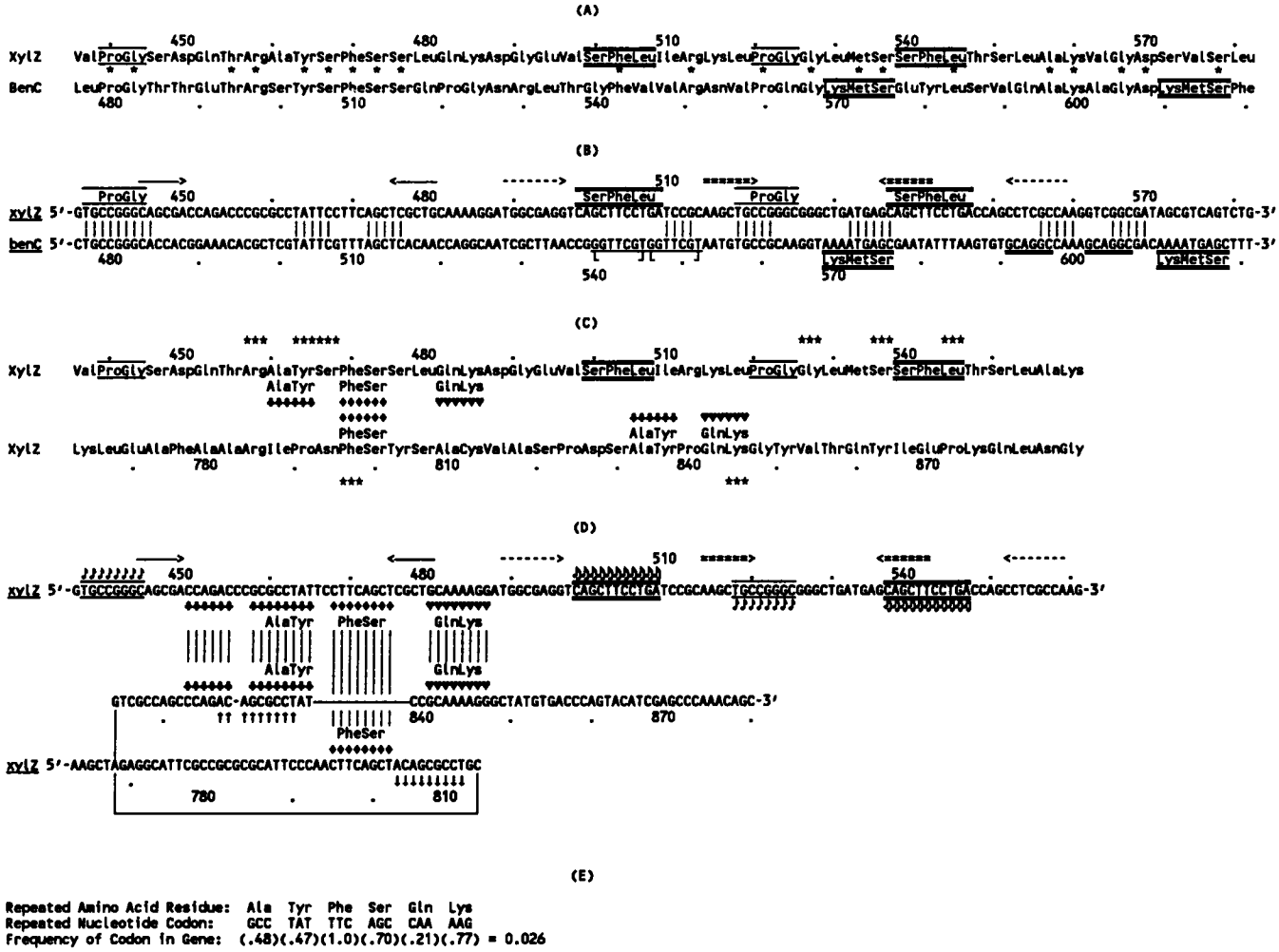

FIG. 2. Inferences concerning common ancestry of DNA sequence repetitions acquired during evolution of upstream (positions 436 to 561) and downstream (positions 763 to 888) regions of *xylZ*. Numerals correspond to positions within the nucleotide sequence of the genes. (A) Amino acid sequence repetitions within the upstream region of XylZ and within the homologous region of BenC. Asterisks mark amino acids conserved in the two gene products. Repeated peptides within each divergent amino acid sequence are marked above and below by single or double lines. (B) DNA sequence repetitions acquired within regions of *xylZ* and *benC* and corresponding to the amino acid sequence comparisons shown in panel A. Vertical lines connect regions in which four or more contiguous bases in the aligned *xylZ* and *benC* DNA sequences are identical. Horizontal single or double lines indicate direct DNA sequence repetitions. Inverted DNA sequence repetitions within *xylZ* are distinctively marked with arrows. (C) Dipeptides found in the upstream and downstream regions of XylZ are marked with clubs (AlaTyr), diamonds (PheSer), and hearts (GlnLys). Stars above and below the amino acid sequences indicate amino acids that have been conserved during divergence of XylZ and BenC from ancestors they share with chloroplast ferredoxin reductase (22). (D) Introduction of a loop into the downstream *xylZ* sequence suggests how DNA encoding peptides conserved between the upstream and downstream sequences could be traced to a slippage structure acquired during divergence of the gene. Clubs, diamonds, and hearts indicate DNA corresponding to peptides depicted in panel C. Single or double musical notes mark DNA sequences corresponding to peptides repeated in the upstream region as shown in panel B, and vertical arrows indicate a DNA sequence repetition acquired in the downstream region. Vertical lines connect DNA sequence identities clustered in the aligned regions of the upstream and downstream sequences. (E) Frequency of usage within *xylZ* of codons for peptides repeated within the upstream and downstream regions.

The complex nature of DNA slippage structures and their potential ability to interact makes them difficult to elucidate. Procedures that may reveal such structures are illustrated by analysis of repetitions acquired in *xylZ* during its divergence from an ancestral gene it shared with *benC*. The *xylZ* sequence repetitions presented in Fig. 2 and 3 encode the C-terminal region of XylZ and were acquired during divergence of the protein from predecessors shared with other oxidoreductases including ferredoxin reductase and, more recently, BenC.

DNA sequences presented as slippage structures in Fig. 2 and 3 were selected after screening the entire *xylXYZ* region for DNA sequence repetitions exceeding 8 bp and falling within 120 bp of each other. Only two DNA segments, the upstream and downstream sequences in Fig. 2 and 3, met this stringent criterion for evidence of localized strand slippage. This inference was strengthened by discovery of three nonoverlapping pairs of inverted repetitions of 6 bp or longer within the upstream region. Further analysis revealed three different 8-bp sequences in the upstream region that were repeated, albeit in different order, in the downstream region. Potential hybridization patterns between comple-

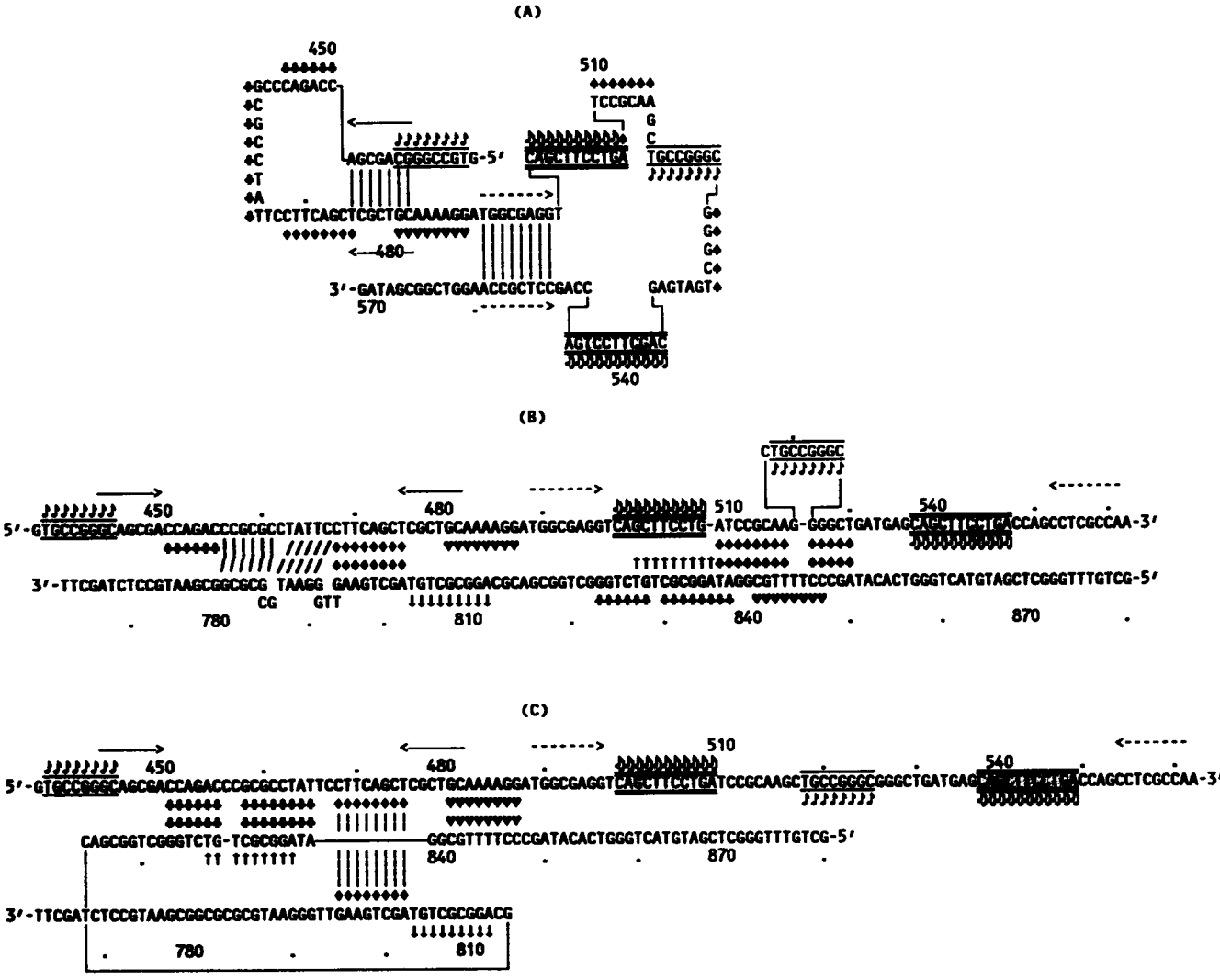

FIG. 3. Slippage structures suggested by repeated DNA sequences in the upstream and downstream regions of *xylZ*. (A) The 5′ (coding) sequence of the upstream region (Fig. 2) is folded so that vertical lines illustrate hybridization that might occur between inverted repetitions. As in Fig. 2, single or double musical notes indicate direct DNA sequence repetitions that might allow hybridization between the coding and noncoding strands of the upstream region after slippage. Largely single stranded within the coding strand slippage structure are sequences (clubs, diamonds, and hearts) that possess the potential ability to hybridize with the noncoding strand of the downstream region of *xylZ* (Fig. 2D and panel C, below). Spades indicate DNA which, pinched into single strandedness in this representation of the upstream region, has the potential ability to initiate formation of a slippage structure by hybridizing with DNA in the downstream region as shown in panel B. (B) One possible DNA slippage structure in which hybridization between the coding strand of the upstream region and the noncoding strand of the downstream region might predominate. Horizontal arrows indicate regions contributing to the intrastrand loops depicted in panel A. Vertical arrows mark regions of direct sequence repetition, separated by 24 bases, that could allow hybridization between coding and noncoding strands of the downstream region. This interaction might contribute to formation of the alternative slippage structure shown in panel C. (C) Slippage structure that could account for conservation of DNA sequence repetitions within *xylZ* as depicted in Fig. 2D.

mentary components of repeated DNA sequences are presented in Fig. 3.

A portion of the amino acid sequence of XylZ is compared with the corresponding sequence of BenC in Fig. 2A. The comparison illustrates the similarities that run throughout the aligned sequences (Fig. 1) and demonstrates internal repetitions, expressed at the level of peptide sequence, that were acquired as XylZ and BenC diverged. Figure 2B shows that the acquired repetitions observed at the level of peptide in Fig. 2A are expressed more extensively at the level of direct DNA sequence repetitions which are as long as 11 bp (in the *xylZ* sequences encoding SerPheLeu, Fig. 2B). In addition, the *xylZ* DNA contains inverted repetitions lying in regions divergent from *benC* (Fig. 2B).

The XylZ sequence contains three dipeptides which are repeated in a different order in a downstream portion of the same gene product (Fig. 2C). As shown in Fig. 2D, introduction of a loop into the 5′ strand encoding the downstream set of dipeptides allows alignment of its DNA with a nearly identical sequence encoding the upstream set of peptides. Also shown in Fig. 2D are direct and inverted DNA sequence repetitions that have the potential ability to hybridize and thus to contribute to slippage structures. The frequency of codon usage for the repeated dipeptides, shown in Fig.

2E, indicates that the probability that this particular DNA sequence would be used to encode the peptides is 0.026. This value is a substantial overestimate because it overlooks the fact that DNA encoding the dipeptides represents repetitions extending 8 rather than 6 bp (Fig. 2D). Furthermore, the 8-bp repetition corresponding to the peptide AlaTyr is flanked by a perfect 6-bp repetition between the upstream and downstream sequences (Fig. 2D).

Clusters of DNA sequence identity shared between the upstream and downstream regions of *xylZ* (Fig. 2D) are more extensive than those shared by the upstream region of *xylZ* and its unmistakable *benC* homolog (Fig. 2B): the former comparison contains three clusters in which sequence identities exceed seven contiguous base pairs, and only one such cluster is found in the latter comparison. This evidence points to the conclusion that the upstream and downstream regions of *xylZ* share common ancestry, and this common ancestry was established by genetic exchange between the regions subsequent to their divergence from the ancestor shared with *benC*. The genetic exchange must have been based on hybridization between complementary sequences in the upstream and downstream regions of *xylZ*; examination of the DNA sequences in these regions (Fig. 3) suggests how such hybridization might be formed.

In Fig. 3A, the 5′ (coding) strand from the upstream region of *xylZ* is depicted in a form it would possess if hybridization between inverted repetitions were allowed to take place. The 3′ (noncoding) strand is not shown in this depiction, but it should be recognized that it could enter into a slippage structure through hybridization between complementary components of direct DNA sequence repetitions. The coding portions of these repetitions, represented by single or double musical notes in Fig. 3A, are largely single stranded and therefore accessible for interaction with their misaligned noncoding counterparts in the upstream region. Also largely single stranded in the depicted structure of the upstream coding region are DNA sequences (symbolized by clubs, diamonds, and hearts in Fig. 3A) that are free to enter into hybridization with noncoding counterparts in the downstream region (Fig. 3C).

The remaining portion of the upstream coding strand is represented largely by the sequence marked by spades in Fig. 3A. As shown in Fig. 3B, this region could enter into hybridization with a portion of the noncoding downstream strand to form an alternative slippage structure. This structure would pinch into single strandedness the coding component of the upstream direct repetition that is represented by single musical notes (Fig. 3B). Also forced into single strandedness by the slippage structure suggested in Fig. 3B would be noncoding DNA (symbolized by hearts in the downstream region) that would have potential ability to hybridize with coding DNA (symbolized by hearts in the upstream region) and thus contribute to formation of the slippage structure shown in Fig. 3C.

Existence of the potentially interactive slippage structures shown in Fig. 2 and 3 is proposed on the basis of patterns of sequence identity that appear to defy coincidence. Regions of sequence identity in misaligned DNA strands extend from 6 to 8 bp and fit into three-dimensional jigsaw puzzles in which very few pieces appear to be missing (Fig. 2 and 3). The repeated DNA sequences, acquired during divergence of *xylZ* from the ancestor it shared with *benC* (Fig. 2), have been maintained even against selective pressure for usage of preferred codons. It therefore seems possible that mismatch repair between slipped DNA strands contributes to maintenance of the DNA sequence repetitions. This possibility can be explored by examination of the stability of mutations that disrupt potentially interactive DNA sequences and also cause substitution of essential amino acid residues in proteins. The contribution of slippage structures to repair of such mutations could be assessed by examination of the reversion frequency of strains in which they are carried.

Site-directed mutagenesis (15) can be used to prepare bacterial strains carrying mutations that alter amino acid residues essential for *xylZ* function. Targets for such mutagenesis emerge from identification of amino acid residues conserved within the broad oxidoreductase family encompassing chloroplast ferredoxin reductase (22). Targets thus identified are marked with stars in the slippage structure suggested in Fig. 2. The specific contribution made by the targeted residues can be inferred because the crystal structure of ferredoxin reductase has been determined (14). The likelihood that these residues were conserved because they make an important contribution to enzyme function can be tested by their genetic modification through site-directed mutagenesis. If the modifications result in an inactive enzyme, the possibility that slippage structures contribute to DNA sequence maintenance can be explored by examining frequencies of reversion of the mutated DNA.

ACKNOWLEDGMENTS

This research was supported by grants to L.N.O. from the Army Research Office, the Celgene Corporation, and the National Science Foundation. Research in the laboratory of S.H. was supported by the Swiss National Science Foundation.

REFERENCES

1. **Andersson, S. G. E., and C. G. Kurland.** 1990. Codon preference in free-living microorganisms. Microbiol. Rev. **54**:198–210.
2. **Dagley, S.** 1986. Biochemistry of aromatic hydrocarbon degradation in pseudomonads, p. 527–555. *In* J. R. Sokatch (ed.), The bacteria, vol. 10. The biology of Pseudomonas. Academic Press, Inc., Orlando, Fla.
3. **de Boer, J. G., and L. S. Ripley.** 1984. Demonstration of the production of frameshift and base-substitution mutations by quasipalindromic DNA sequences. Proc. Natl. Acad. Sci. USA **81**:5528–5531.
4. **Fix, D. F., P. A. Burns, and B. W. Glickman.** 1987. DNA sequence analysis of spontaneous mutants in the PolA1 strain of *Escherichia coli* indicates sequence specific effects. Mol. Gen. Genet. **207**:267–272.
5. **Glickman, B. W., and L. S. Ripley.** 1984. Structural intermediates of deletion mutagenesis: a role for palindromic DNA. Proc. Natl. Acad. Sci. USA **81**:512–516.
6. **Golding, G. B., and B. W. Glickman.** 1985. Sequence directed mutagenesis: evidence from a phylogenetic history of human α-interferon. Proc. Natl. Acad. Sci. USA **82**:8577–8581.
7. **Harayama, S., P. R. Lehrbach, and K. N. Timmis.** 1984. Transposon mutagenesis analysis of *meta*-cleavage pathway operon genes of the TOL plasmid of *Pseudomonas putida* mt-2. J. Bacteriol. **160**:251–255.
8. **Harayama, S., and M. Rekik.** 1990. The *meta* cleavage operon of TOL degradative plasmid pWW0 comprises 13 genes. Mol. Gen. Genet. **221**:113–120.
9. **Hartnett, C. S., E. L. Neidle, K.-L. Ngai, and L. N. Ornston.** 1990. DNA sequences of genes encoding *Acinetobacter calcoaceticus* protocatechuate 3,4-dioxygenase: evidence indicating shuffling of genes and of DNA sequences within genes during their evolutionary divergence. J. Bacteriol. **172**:956–966.
10. **Horn, J. M., S. Harayama, and K. N. Timmis.** DNA sequence determination of the TOL plasmid *xylGFJ* genes of *Pseudomonas putida*: implications for the evolution of aromatic catabolism. Mol. Microbiol., in press.
11. **Ikemura, T.** 1981. Correlation between the abundance of *E. coli* transfer tRNAs and the occurrences of the repetitive codon in protein genes. J. Mol. Biol. **146**:389–409.

12. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1986. Nucleotide sequence of the regulatory gene *xylS* on the *Pseudomonas putida* TOL plasmid and identification of the protein product. Gene **44**:235–242.
13. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1988. Nucleotide sequence of the regulatory gene *xylR* of the TOL plasmid from *Pseudomonas putida*. Gene **66**:301–306.
14. **Karplus, P. A., M. J. Daniels, and J. R. Herriott.** 1991. Atomic structure of ferredoxin reductase: prototype for a structurally novel flavoenzyme family. Science **251**:60–66.
15. **Kunkel, T. A.** 1985. Rapid and efficient site specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
16. **Kunkel, T. A., and A. Soni.** 1988. Mutagenesis by transient misalignment. J. Biol. Chem. **263**:14784–14789.
17. **Levinson, G., and G. A. Gutman.** 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. **4**:203–221.
18. **Meselson, M.** 1988. Methyl directed repair of DNA mismatches, p. 91–110. *In* K. B. Low (ed.), Recombination of genetic material. Academic Press, Inc., New York.
19. **Nakai, C., H. Kagamiyama, M. Nozaki, T. Nakazawa, S. Inouye, Y. Ebina, and A. Nakazawa.** 1983. Complete nucleotide sequence of the metapyrocatechase gene on the TOL plasmid of *Pseudomonas putida* mt-2. J. Biol. Chem. **258**:2293–2298.
20. **Neidle, E. L., C. S. Hartnett, S. Bonitz, and L. N. Ornston.** 1988. DNA sequence of the *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase I structural gene *catA*: evidence for evolutionary divergence of intradiol dioxygenases by acquisition of DNA sequence repetitions. J. Bacteriol. **170**:4874–4880.
21. **Neidle, E. L., C. S. Hartnett, and L. N. Ornston.** 1989. Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. J. Bacteriol. **171**:5410–5421.
22. **Neidle, E. L., C. S. Hartnett, L. N. Ornston, A. Bairoch, M. Rekik, and S. Harayama.** 1991. Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-dioxygenase reveal evolutionary relationships among multicomponent oxygenases. J. Bacteriol. **173**:5385–5395.
23. **Neidle, E. L., and L. N. Ornston.** 1986. Cloning and expression of the *Acinetobacter calcoaceticus* catechol 1,2-oxygenase structural gene, *catA*, in *Escherichia coli*. J. Bacteriol. **168**:815–820.
24. **Normark, S., S. Bergstrom, T. Edlund, T. Grundstrom, B. Jaurin, F. P. Lindberg, and O. Olsson.** 1983. Overlapping genes. Annu. Rev. Genet. **17**:499–525.
25. **Ornston, L. N.** 1982. Recurring themes and repeated sequence in metabolic evolution, p. 105–126. *In* A. Chakrabarty (ed.), Biodegradation and detoxification of environmental pollutants. CRC Press, Inc., Boca Raton, Fla.
26. **Ornston, L. N., J. E. Houghton, E. L. Neidle, and L. A. Gregg.** 1990. Subtle selection and novel mutation during evolutionary divergence of the β-ketoadipate pathway, p. 207–225. *In* S. Silver, A. M. Chakrabarty, B. Iglewski, S. Kaplan, and S. Silver (ed.), *Pseudomonas*: biotransformations, pathogenesis, and evolving biotechnology. American Society for Microbiology, Washington, D.C.
27. **Ornston, L. N., and E. L. Neidle.** 1991. Evolution of genes for the β-ketoadipate pathway in *Acinetobacter calcoaceticus*, p. 201–237. *In* K. J. Towner, E. Bergogne-Berezin, and C. A. Fewson (ed.), The biology of *Acinetobacter*. FEMS Symposium Series. Plenum Press, New York.
28. **Ornston, L. N., E. L. Neidle, and J. E. Houghton.** 1990. Gene rearrangements, a force for evolutionary change; DNA sequence arrangements, a source of genetic constancy, p. 325–334. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
29. **Ornston, L. N., and W.-K. Yeh.** 1979. Origins of metabolic diversity: evolutionary divergence by sequence repetition. Proc. Natl. Acad. Sci. USA **76**:3996–4000.
30. **Reiner, A. M.** 1971. Metabolism of benzoic acid by bacteria: 3,5-cyclohexadiene-1,2-diol-1-carboxylic acid is an intermediate in the formation of catechol. J. Bacteriol. **108**:89–94.
31. **Shanley, M. S., E. L. Neidle, R. E. Parales, and L. N. Ornston.** 1986. Cloning and expression of *catBCED* genes in *Pseudomonas putida* and *Escherichia coli*. J. Bacteriol. **165**:557–563.
32. **Sueoka, N.** 1988. Directional mutation pressure and neutral molecular evolution. Proc. Natl. Acad. Sci. USA **85**:2653–2657.
33. **Suzuki, M., T. Hayakawa, J. P. Shaw, M. Rekik, and S. Harayama.** 1991. Primary structure of xylene monooxygenase: similarities to and differences from the alkane hydroxylation system. J. Bacteriol. **173**:1690–1695.
34. **Tautz, D., M. Trick, and G. A. Dover.** 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature (London) **322**:652–656.
35. **West, S. E. H., and B. Iglewski.** 1988. Codon usage in *Pseudomonas aeruginosa*. Nucleic Acids Res. **16**:9323–9335.
36. **Worsey, M. J., and P. A. Williams.** 1975. Metabolism of toluene and xylenes by *Pseudomonas putida* (*arvilla*) mt-2: evidence for a new function of the TOL plasmid. J. Bacteriol. **124**:7–13.
37. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–109.
38. **Yeh, W.-K., P. Fletcher, and L. N. Ornston.** 1980. Evolutionary divergence of coselected β-ketoadipate enol-lactone hydrolases in *Acinetobacter calcoaceticus*. J. Biol. Chem. **255**:6342–6346.
39. **Yeh, W.-K., P. Fletcher, and L. N. Ornston.** 1980. Homologies in the NH_2-terminal amino acid sequence of τ-carboxymuconolactone decarboxylases and muconolactone isomerases. J. Biol. Chem. **255**:6347–6354.
40. **Yeh, W.-K., and L. N. Ornston.** 1980. Origins of metabolic diversity: substitution of homologous sequences into genes for enzymes with different catalytic activities. Proc. Natl. Acad. Sci. USA **143**:5365–5369.
41. **Yeh, W.-K., and L. N. Ornston.** 1981. Evolutionarily homologous $\alpha_2\beta_2$ oligomeric structures in β-ketoadipate succinyl CoA transferases from *Acinetobacter calcoaceticus* and *Pseudomonas putida*. J. Biol. Chem. **256**:1565–1569.