The aromatic amino acids phenylalanine, tyrosine and tryptophan are formed by a sequence of reactions starting with synthesis of 3-deoxy-d-arabino-heptulosonate 7-phosphate (DAHP), which occurs by the enzymically catalysed condensation of D-erythrose 4-phosphate and phosphoenolpyruvate (Henner et al., 1973). The synthesis is catalysed by DAHP synthase (EC 4.1.2.15), encoded in the chromosome (Anagnostopoulos et al., 1993). DAHP synthase of B. subtilis 168 is sensitive to feedback inhibition by chorismate and prephenate and is also responsible for the chorismate mutase activity in this strain (Huang et al., 1974a,b). The chorismate mutase activity in B. subtilis 168 is assigned to the gene aroA (Hoch & Nester, 1973). The aroA and aroG genes have been placed in one locus on the chromosome map of B. subtilis 168 (Anagnostopoulos et al., 1993). The enzyme encoded by aroA(G) has been partially purified and data on its amino acid composition obtained (Huang et al., 1974a). It was shown (Llewellyn et al., 1980) that DAHP synthase of B. subtilis 168 was distinct from the monofunctional DAHP synthase of B. subtilis ATCC 6051 Marburg strain. It had an additional chorismate mutase activity, and it was postulated that the bifunctional enzyme arose from the monofunctional enzyme by conversion of an allosteric site for binding prephenate into an active site for chorismate mutase activity (Llewellyn et al., 1980).

In this paper we present our data on the cloning and structural analysis of the aroA(G) gene of B. subtilis 168 and the aroA gene of B. subtilis ATCC 6051 Marburg strain.

The aroA(G) gene of B. subtilis 168 was cloned on a 2 kb DNA fragment by partial digestion of the chromosome with Sau3A, using the plasmid pIJ2 (Yomantas et al., 1979) as the vector. The selection was carried out by complementation of the aroG32 mutation of an araG932 leuA8 recE4 strain of B. subtilis (Yu. Yomantas, GNIIgenetika, Moscow). The cloned fragment also complements the aroA6 mutation of B. subtilis. The level of DAHP synthase and chorismate mutase activities in crude extracts of a B. subtilis strain carrying the aroA(G) fragment on a multicopy plasmid was about 10 times higher than in the recipient B. subtilis strain without the plasmid (data not shown).

The nucleotide sequence of the fragment is shown in Fig. 1. It contains two open reading frames, designated orf324 and orf1102. Both have putative ribosome-binding sites.

**Abbreviations:** DAHP, 3-deoxy-d-arabino-heptulosonate 7-phosphate.

The EMBL accession number for the nucleotide sequence reported in this paper is X65945.
A. BOLOTIN and OTHERS

Fig. 1. Nucleotide sequence of the cloned DNA fragment. The sites for Avall, EcoRI, EcoRV and PstI are indicated. The amino acid sequences of the products of orf324 and orf1102 (aroA(G)) are indicated in single-letter code. Paired arrows under the nucleotide sequence denote inverted repeats. The transcriptional start point of aroA(G) mRNA (which was determined in primer extension experiments with oligonucleotide 5′Agtagatagcttttagaattttgaaagaa-S') is indicated by an asterisk. The -35 and -10 regions upstream of these proteins are presented in Fig. 2(a). Multiple alignments of these proteins are shown by capital letters. The oligonucleotides AG-N and AG-C2 were used for synthesis of the PCR product of the aroA gene of B. subtilis Marburg strain.

A. BOLOTIN and OTHERS

Fig. 1. Nucleotide sequence of the cloned DNA fragment. The sites for Avall, EcoRI, EcoRV and PstI are indicated. The amino acid sequences of the products of orf324 and orf1102 (aroA(G)) are indicated in single-letter code. Paired arrows under the nucleotide sequence denote inverted repeats. The transcriptional start point of aroA(G) mRNA (which was determined in primer extension experiments with oligonucleotide 5′Agtagatagcttttagaattttgaaagaa-S') is indicated by an asterisk. The -35 and -10 regions upstream of these proteins are presented in Fig. 2(a). Multiple alignments of these proteins are shown by capital letters. The oligonucleotides AG-N and AG-C2 were used for synthesis of the PCR product of the aroA gene of B. subtilis Marburg strain.

and ρ-independent transcription terminator structures. The start point of orf1102 mRNA was determined by primer extension analysis; −35 and −10 regions upstream of it resemble the consensus structure of ρ-independent promoters (Fig. 1). The protein encoded by orf1102 has a molecular mass and amino acid composition similar to purified DAHP synthase (Huang et al., 1974a). Disruption of orf1102 by inserting a DNA fragment into the PstI or EcoRI site resulted in the loss of the complementation of the araA mutation in B. subtilis. These data support the identity of orf1102 as the araA(G) gene.

The amino acid sequence of the araA(G) gene product (DAHP synthase-chorismate mutase) was compared with the PIR/NBRF protein data bank. The N-terminal part of the protein (amino acids 1–60) shows homology with the chorismate mutase parts of bifunctional enzymes encoded by the pheA and tyrA genes of E. coli (Hudson & Davidson, 1984) and monofunctional chorismate mutase encoded by the araH gene of B. subtilis ATCC 6051 Marburg strain (Gray et al., 1990). Multiple alignments of these proteins are presented in Fig. 2(a). The C-terminal part of the amino acid sequence of DAHP synthase-
chorismate mutase has no significant similarities to sequences of known E. coli DAHP synthases (Davies & Davidson, 1982; Schulz et al., 1984; Zurawski et al., 1981), but has some similarity to the amino acid sequence of 3-deoxy-d-manno-octulosonate 8-phosphate synthase of E. coli (EC 4.1.2.16) encoded by the kdsA gene (Woisteschlag & Hogenauer, 1987) (Fig. 2b). This enzyme was responsible for similar aldol condensation of d-arabino 5-phosphate (inside d-erythrose 4-phosphate) and phosphoenolpyruvate (Levin & Racker, 1959), which resulted in the 3-deoxy-d-manno-octulosonic acid—a sugar which links the hydrophilic polysaccharide chains of the cell wall in Gram-negative bacteria.

The 1116 base PCR product of the aroA gene (amplified with oligonucleotides AG-N and AG-C2: Fig. 1), which encodes monofunctional DAHP synthase in B. subtilis Marburg strain, was used for DNA sequencing and was cloned in the BamHI site of plasmid pCB22 (Sorokin & Khazak, 1990) downstream of expression unit EU19035. The recombinant plasmid complemented aroA(G1) of strain 168, which was responsible for similar aldole condensation of 3-deoxy-~-manno-octulosonic acid and mimic effects of high active chorismate mutase encoded by aroH in the Marburg strain.

Fig. 2. Protein sequence alignments of DAHP synthase-chorismate mutase (aroA(G)) encoded by the aroA(G) gene of B. subtilis with: (a) monofunctional chorismate mutase (aroH) encoded by the aroH gene of B. subtilis ATCC 6051 Marburg strain (Gray et al., 1990), chorismate mutase-prephenate dehydratase (pheA), chorismate mutase-prephenate dehydrogenase (tyrA) encoded by the pheA and tyrA genes of E. coli (Hudson & Davidson, 1984); and (b) 3-deoxy-d-manno-octulosonate 8-phosphate synthase (kdsA) encoded by the kdsA gene of E. coli (Woisteschlag & Hogenauer, 1987). Similarities in amino acid sequences are marked by asterisks (same amino acid), colors (strong similarity) and dots (family similarity). Alignments were done with the DNA-SUN software (A. Mironov, GNigentetika).

REFERENCES


Received 3 April 1995; accepted 11 May 1995.