**Communication**

**Tandem Repeat of the Genes for Protein S, a Development-specific Protein of Myxococcus xanthus**

(Received for publication, September 10, 1982)

Sumiko Inouye†, Yoshimasa Ika§ and Masayori Inouye†

From the †Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, New York 11794 and the §Division of Biology, City of Hope Research Institute, Duarte, California 91010

Protein S, a development-specific protein of Myxococcus xanthus is produced only during fruiting body formation. More than 15% of total protein synthesis during this period is accounted for by the production of protein S. The genes for protein S were identified and cloned with the use of mixed probes consisting of eight synthetic oligodeoxyribonucleotides (tetradecamers) which correspond to a carboxyl-terminal portion of protein S. The two genes are oriented in the same direction and are separated approximately 1.2 kilobases. The DNA sequences of the carboxyl-terminal portions of the two genes reveal that both can code for the identical eleven amino acid sequence which corresponds to the carboxyl-terminal end of protein S. However, there are a few base substitutions upstream of these regions. This duplication of genes in **M. xanthus** may facilitate the extremely rapid synthesis of protein S during fruiting body formation.

**Myxococcus xanthus** is a rod-shaped Gram-negative bacterium, which undergoes a unique developmental cycle forming fruiting bodies when starved of nutrients (see review Ref. 1). During fruiting body formation, several development-specific proteins have been identified (2–4). Among them, protein S is of particular interest since an extremely large amount of protein S is produced during differentiation (2, 3). Protein S assembles on the surface of myxospores, and the mechanism of its self-assembly has been investigated (3). It has been purified and crystallized (5, 6), and its partial structure has been determined (6). In the present paper, we have attempted to clone the gene for protein S using synthetic oligonucleotides as probes. It was found that the gene for protein S is duplicated in **Myxococcus xanthus**. The significance of the gene duplication will be discussed.

**EXPERIMENTAL PROCEDURES**

**DNA Preparation—** **M. xanthus** DNA was prepared as described previously (7).

**Other Materials and Methods—** Synthetic oligonucleotides were obtained from BioLogicals (Ottawa, Canada). Restriction enzymes were obtained from either Bethesda Research Laboratories or New England Biolabs.

**RESULTS**

**Cloning of the Gene for Protein S—** In order to identify and clone the gene for protein S, the following oligodeoxyribonucleotides consisting of 14 nucleotides were synthesized:

```
Met Asn Asn Asn Thr
ATGAATAATAATAC
```

These eight tetradecamers represent all possible DNA sequences which would code for the amino acid sequence, Met-Asn-Asn-Asn-Thr, derived from a part of the carboxyl-termi-
FIG. 2. Restriction enzyme mapping of a clone (pSI001) of the 9.7-kb HindIII fragment carrying the genes for protein S. a, the mapping of the entire plasmid DNA of pSI001. The boxed portion represents the 9.7-bp HindIII fragment from M. xanthus. The clone was screened with use of 3P-labeled mixed probes consisting of eight synthetic tetradecamers as described in the text. The shaded regions represent the DNA fragments which hybridized with the probes. The solid lines represents the DNA fragment from pBR322. Amp with an arrow indicates the position and the direction of transcription of the gene for 3-lactamase. b, location of the genes for protein S. Only the 3.2-kb DNA fragment from the HindIII site to the second BamHI site, which includes both regions hybridizing with the probes, is shown. Triangles indicate the positions where the probes hybridized. Solid lines under the map represent the two DNA fragments (600-bp HindIII-Hinfl fragment, and 640-bp Hinfl-Hinfl fragment) used for the sequencing. These fragments were fitted in with [a-3P]dXTP using Klenow enzyme and digested with AvelI. The right-hand half of each fragment was purified by acrylamide gel electrophoresis, and its DNA sequence was determined using the method of Maxam and Gilbert (10). The coding regions and orientations of the two genes (gene 1 and 2) were indicated by thick arrows.

Fig. 3. Sequencing gels of the DNA fragments derived from gene 1 and gene 2 for protein S (see Fig. 2b). In both cases, the sequences were determined from the right Hinfl site of the 600-bp HindIII-Hinfl fragment of gene 1 and the 640-bp Hinfl-Hinfl fragment of gene 2 (see Fig. 2b).

Gene 1

1. The abbreviation used is: kb, kilobases.
Genes for Protein S of M. xanthus

sequences which hybridized with the probes were ATGAA-GAACAAACAC in both cases and the amino acid sequences deduced from the DNA sequences at this region (-Ala-Met-Asn-Asn-Thr-Ser-Ser-Ile-) are identical with that of the carboxyl-terminal end of protein S (6).

DISCUSSION

These results strongly suggest that there are two genes for protein S in the M. xanthus chromosome which are tandemly repeated in the same orientation. Such gene duplication is possibly very important for the production of an extremely large amount of protein S during fruiting body formation. Since protein S consists of approximately 210 amino acid residues, the coding region of protein S should be at least 630 base pairs in length. Therefore, the space between the two genes (the termination codon for gene 1 and the initiation codon of gene 2) is estimated to be approximately 1.2 kb. How is the expression of the tandemly-repeated genes regulated? Do the amino acid substitutions in protein S play any role in its structure and function? How have these two genes been evolved? The determination of the DNA sequence of the entire region including genes 1 and 2 will shed light onto these questions.

At this stage, it is also possible that one of the two genes is a pseudogene which is not expressed or that they may code two different products which share the same carboxyl-terminal structure. The DNA partial sequences of gene 1 and gene 2 (Fig. 4) indicate (a) base substitutions occur always at the first or second base of triplet codons, resulting in alteration of amino acids coded by them; (b) all 32 codons have C or G at the third base position, which is consistent with the high GC content of the M. xanthus chromosomal DNA (69%; see review, Ref. 1); and (c) 53% of the amino acid residues in this region are accounted for by asparagine, serine, and theonine.

Acknowledgments—We are grateful to Dr. T. Komano for his valuable discussion and to T. Franceschini for his skillful assistance throughout this work.

REFERENCES