Protein U, a Late-Developmental Spore Coat Protein of *Myxococcus xanthus*, Is a Secretory Protein

RACHEL GOLLOP, MASAYORI INOUEY, AND SUMIKO INOUEY*

Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854

Received 9 November 1990/Accepted 16 March 1991

Protein U is a spore coat protein produced at the late stage of development of *Myxococcus xanthus*. This protein was isolated from developmental cells, and its amino-terminal sequence was determined. On the basis of this sequence, the gene for protein U (pru) was cloned and its DNA sequence was determined, revealing an open reading frame of 179 codons. The product from this open reading frame has a typical signal peptide of 25 amino acid residues at the amino terminal end, followed by protein U of 154 residues. This result indicates that protein U is produced as a secretory precursor, pro-protein U, which is then secreted across the membrane to assemble on the spore surface. This is in sharp contrast to protein S, a major spore coat protein produced early in development, which has no signal peptide, indicating that there are two distinct pathways for trafficking of spore coat proteins during the differentiation of *M. xanthus*.

*Myxococcus xanthus* is a gram-negative bacterium that lives in soil and migrates by gliding on solid surfaces. Upon nutritional starvation, cells aggregate and form fruiting bodies, within which cells differentiate into myxospores (for reviews, see references 9, 14, and 18). Dramatic changes in the pattern of protein synthesis have been reported (4) during development and fruiting-body formation. The most significant changes in the pattern of protein synthesis is the appearance of two proteins, S and U (4). The synthesis of protein S increases dramatically at a very early stage of development and reaches a maximum of 15% of total soluble protein synthesized (4, 5). On the other hand, the synthesis of protein U starts at a late stage of development, when spore formation begins (4). Both proteins S and U are assembled on spore surfaces and are considered to be spore coat proteins (5). Thus these two proteins provide ideal systems for studying differential gene expression during cellular differentiation of *M. xanthus*.

In addition, there is an intriguing question of how these spore coat proteins are excrated from the cell to be assembled on the spore surface. It has been demonstrated that protein S has no signal peptide (6) in spite of the fact that it is excrated from the cell. To determine the amino-terminal sequence of protein U, *M. xanthus* FB(DZF1) cells (7) were harvested from eight plates at 130 h of development on CF agar (3) and suspended in 10 mM Tris-HCl buffer (pH 7.4). The suspension was centrifuged at 9,300 × g for 15 min at 4°C. The pellet was suspended in 100 µl of the same buffer and sonicated. This pellet mainly consists of myxospores, and the sonication treatment described above was found to release protein U into the supernatant after centrifugation at 100,000 × g for 30 min at 4°C. The supernatant fraction was then loaded on a 17.5% sodium dodecyl sulfate–polyacrylamide gel; after electrophoresis, proteins were blotted to a polyvinylidene difluoride membrane by using a semidry blotter apparatus (Sartorius, Göttingen, Germany). The polyvinylidene difluoride membrane was then stained with Coomassie blue, and the band corresponding to protein U was cut out. Approximately 600 pmol was used for sequencing. The amino acid sequence was obtained on an Applied Biosystems, Inc. 470A protein sequencer.

The amino-terminal sequence of protein U was determined. The sequence thus obtained was X-Thr-Ala-Thr-Ala-Val-Leu-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Ser-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,

Asn-Leu-Asn-Val-Thr-Ala-Asn-Val-Gly-Gly-,
(lane 2), 0.42-kb RsaI (lane 3), and 1.0-kb NruI (lane 4) fragments were hybridized with the probe. On the basis of these results, the 1.8-kb SmaI (XmaI was used for cloning) fragment was then cloned into the XmaI site of pUC9 (21). Using the same fragment as a probe, colonies that hybridized with the probe were identified. One of the plasmids isolated from the positive colonies was then designated pRGUX60. Then the rest of the DNA sequence was determined (Fig. 3). The DNA sequence revealed the existence of an open reading frame of 179 amino acid residues. The amino-terminal amino acid sequence was determined by microsequencing of protein U to extend from Thr at position 27 (note that the threonine residue was the second amino acid of protein U and that the first amino acid could not be identified) to Gly at position 40. This indicates that protein U is produced from a precursor, pro-protein U, with an extra amino-terminal sequence of 25 amino acid residues. The extra sequence shows features characteristic of the prokaryotic signal peptide (13), such as a positively charged amino-terminal region, an extremely hydrophobic core region, a less hydrophobic cleavage region, and an alanine residue at the cleavage site (13). It is interesting to note that there is a glutamic acid residue at the second residue from the cleavage site, which is unusual for a signal peptide. This may have some role in modulating the efficiency of the signal peptide cleavage for pro-protein U.

The existence of the signal peptide in pro-protein U is quite different from the major early spore coat protein, protein S, which has no signal peptide (6). Earlier we proposed an exospore hypothesis (20) to explain how protein S is excreted into the medium and assembled on the surface of the myxospores. This hypothesis is based on the following facts. (i) Approximately 90% of the developing cells lyse in an early stage of development (22). Note that the percentage of the developmental cells that lyse appears to depend on experimental conditions (12). (ii) There are two homologous genes for protein S, designated ops and tps, which are tandemly located on the M. xanthus chromosome with a short spacer sequence (6). (iii) The tps gene is responsible for the early production of protein S during development, whereas the ops gene is expressed in a very late stage of development. (iv) The large amount of the tps product (protein S) is assembled on the spore surface, whereas the ops product (protein S1) accumulates only inside the spore. (v) The amount of protein S1 and the amount of protein S inside the mature spores are of the same order of magnitude. (vi) The amount of protein S on the outer surface of the spores is approximately 10 times greater than that found inside the spores. (vii) Protein S produced by a particular strain can be assembled on the spore surface of another strain (20). On the basis of these facts, we have proposed that neither protein S nor protein S1 is secreted across the membrane and that protein S assembled on the spore surface is released by the 90% of the cells that lyse during development (20).

In contrast to the exospore hypothesis, Nelson and Zusman (11) have proposed that protein S can be translocated across the membrane without a cleavable signal peptide. In their proposal it has not been shown how protein S can be

---

**FIG. 1.** Southern blot hybridization of the *M. xanthus* chromosomal DNA. (A) Hybridization with the 23-mer mixed oligonucleotides as a probe. Restriction enzymes used (lanes): 1, BamHI; 2, PstI; 3, Sall; 4, XhoI. The 23-mer mixed probes were synthesized on the basis of the amino-terminal sequence of protein U. (B) Hybridization with the PstI(b)-RsaI(a) fragment as a probe. Restriction enzymes used (lanes): 1, PstI; 2, SmaI; 3, RsaI; 4, NruI. Approximately 3 µg of DNA was loaded in each lane.

---

**FIG. 2.** Restriction map of the 6.4-kb DNA fragment containing the pru gene. The solid arrows represent the open reading frame and its orientation. Solid bars on the top indicate the DNA fragments that were cloned in pRGU68 and pRGUX60. DNA manipulations were performed as described by Maniatis et al. (10). Abbreviations: P, PstI; S1, Sall; X, XhoI; Sm, SmaI; N, NruI; R, RsaI. NruI sites within the PstI(a)-SmaI(b) fragment and RsaI sites within the PstI(a)-RsaI(a) fragment are not shown.
translocated across the membrane without a signal peptide. In either proposal, however, it is clear that there are two distinct pathways for excretion of spore coat proteins: one without the signal peptide and the other with the signal peptide.

Protein U shows no significant homology to known proteins. Analysis of the secondary structure of protein U with the Pplotstructure computer program revealed that the protein has a high content of β structure, as does protein S (7), which may be an important feature for the structure of spore coat proteins. It should be noted that protein U is a rather hydrophobic protein with 45% hydrophobic amino acid residues. The mature protein contains only five basic residues (arginine) and five acidic residues (aspartic acid) and has pl of 6.3. The molecular weight of the mature protein was calculated to be 15,555, which is in good agreement with the apparent molecular weight (16,700) determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

The codon usage of the open reading frame shows that 90% of codons use G or C at the third position of the codon. The sequence GAG, 8 bases upstream to the initiation codon, is likely to be the ribosome-binding site (19). A stable stem-loop structure can be formed between bases 835 and 858 (with a ΔG, calculated by the method of Salser [16], of

**Transmembrane peptides:**

```
VOL. 173, 1991

TGG AAT CTG GTT GCC AAC GGA TTT GCG TGG CCT CTT GCG CTT GTT 45
CGT CCC TCC CTG TGA GTG CAT GTG CAT GTG TGC CGG TAA AAA TAG 90
TTA ATG TTT CGC TTT CCC CTA TGC TCT GGA GAC GCA TGG CAG GCA 135
TGC GTG CGA CAT CGG CCG AAA AGA GGA GGG GCA GTG AGG ACT GTT TGA 180
GAG CTG CGC GCC TGG TTA TGA TGC GCC GCC GCC GGA ATC TAC 225
CGG CGC CCC GAT TGA GCA TGC TCA ATG ACG AAC ATC AAC ACC GCC 270
V A A V T A A A A G L V A P S F P G T C G C G G T C T G C T G C T C G C T C T G C C C 315
A E A V A T A T A N L H V T A N GCC GAG GCC GCC ACC GCC ACC AAC ACC AAC ACC CTG AAC ATC ACC ACC AAC AAC 360
V G G A C S I G S G A G G G T GTC GGA GTT GCC TGG ACC ATC GAT GCC TCC GCC GGG GGG GGG GGG GGG GGG AAC 405
L N F G T Y D P V V V N S A L CGT AAC TAC GCC AGC TAC GAC CCC GCC GTC GTC GTC GTC GTC AAC TCC GCC GTG 450
G V D L F G T G S L S V Q C T GGC GTC GAC GTG TCC GGG AGC GCC TCG TGG AGT CGA TGC ACC 495
L L S T A V I T A L G L Q G L Y P CGT CTA AGC ACC GGC GTC ATC ACC TTT GCG CAC GCC CTG TCG TAC ACC 540
A A G S T A A A V P L R R M R N GCC GCG CGC TGG ACC GGC GCT GTG CCG CTT GGG CTT CTG ATG GCC AAA 585
A A S T D Y L S F Y L Y M D V GCG GCC TGG ACT GAC TAC CTC TCG TAC TCG TAT ATG GAC GTC 630
T R L I A W G N T S G T G L P ACC CGG CTC ATC GCC TGG GGC AAC ACC TCC GCC ACG GCC CTC CGG 675
F L G L G L P V P V P V Q V Y G T TCC CTT GCC CCT GCG CTT CTC GCC CCC GTC CGT CAG GTG TAC GCC AGC 720
V P R I G Q N V P S G T Y N D T GGT CTC GCC GCC CAG AAC GTC CCC TCT GCC ACC TAC AAC GAC ACC 765
V V A T I T F GTG GTT GCC ACC ATC AGC TCC TGC TCA TGC GGA GGG GCA TAG ACC ATT 810
CGT AAC CGT CCT TCG TGC ACT GCC GCC GGC CTT GCT AGC GCA AGC GCC 855
GCC CTG GAA TGG GCC ATC GCG AGC GCC CCC GAG CTC GAT GCC 900
AGT CTG GTC CGG CTT GAG CTG CTT GTT GCC GCC GCC C 934
```

**FIG. 3.** Nucleotide sequence of the gene for protein U of *M. xanthus* and its deduced amino acid sequence. The small arrowhead indicates the location of the *Parl* site (P[b]) used for cloning (see the text). The inverted repeat sequences at the 3' end of the *pru* gene are shown by long arrows. A putative ribosome binding site is underlined. The large arrowhead indicates the cleavage site of the signal peptide.
−13.2 kcal [ca. −55.2 kJ]), which can function as a transcription termination signal (15).

It would also be interesting to know how pru expression is regulated. Recently, it has been suggested that there is a sigma factor cascade for the expression of developmental proteins in M. xanthus (1). Experiments are now in progress to identify the specific sigma factor responsible for the pru gene expression.

We thank Bert Lampson and David Apelian for their reading of the manuscript.

This work was supported by Public Health Service grant 5M26843 from the National Institutes of Health.

REFERENCES