JOURNAL OF BACTERIOLOGY, Nov. 1991, p. 6773–6782
0021-9193/91/216773-10$02.00/0

Vol. 173, No. 21

Murein-Metabolizing Enzymes from *Escherichia coli*: Sequence Analysis and Controlled Overexpression of the *slt* Gene, Which Encodes the Soluble Lytic Transglycosylase

HENK ENGEL, BERT KAZEMIER, AND WOLFGANG KECK*

Department of Biochemistry, BIOSON-Research Institute, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands

Received 5 June 1991/Accepted 2 September 1991

The complete nucleotide sequence of the *slt* gene encoding the soluble lytic transglycosylase (Slt; EC 3.2.1.-) from *Escherichia coli* has been determined. The largest open reading frame identified on a 2.5-kb *Pvu*II-*Sal*I fragment indicates that the enzyme is translated as a preprotein of either 654 or 645 amino acids, depending on which of two potential start codons is used. The two possible translation products differ only in the lengths of their predicted signal peptides, 36 or 27 amino acids, respectively. In both cases, processing results in a soluble mature protein of 618 amino acids (M_r = 70,468). The deduced primary structure of the mature protein was confirmed by N-terminal sequencing and determination of the amino acid composition of the isolated transglycosylase. The *slt* gene contains a high percentage of rare codons, comparable to other low-expressed genes. A hairpin structure that could serve as a transcriptional terminator is located downstream of the *slt* coding region and precedes the *trpR* open reading frame at 99.7 min on the *E. coli* chromosomal map. A computer-assisted search did not reveal any significant sequence similarity to other known carbohydrate-degrading enzymes, including lysozymes. Interestingly, a stretch of 151 amino acids at the C terminus of the transglycosylase shows similarity to the N-terminal portion of the internal virion protein D from bacteriophage T7. Overexpression of the *slt* gene, under the control of the temperature-inducible phage lambda p_R promoter, results in a 250-fold overproduction of the mature transglycosylase, whereas after deletion of the signal peptide a 100-fold overproduction of the enzyme is observed in the cytoplasm.

The murein polymer of the bacterial cell wall is composed of glycan strands of variable length which are cross-linked by short peptide bridges to form one macromolecule around the cell.

A whole set of murein-metabolizing enzymes in *Escherichia coli* has been identified (for a review, see reference 19). The balanced action of these murein-synthesizing and -degrading enzymes determines the specific shape of the murein sacculus and consequently the shape of the bacterium (12, 38, 54).

In *E. coli*, four of the nine murein-degrading enzymes that have been characterized so far are potentially capable of degrading the intact polymer and thereby lysing the cell (20). Two of these enzymes are endopeptidases, able to cleave the peptide cross bridges of murein (47, 49), whereas the other two are glycosylases which are able to degrade the glycan strands (18, 22, 33).

The two glycosylases have been reported to show the same enzymatic activity but to differ with respect to their cellular localization, one of them being a soluble enzyme with a molecular mass of 65 kDa and the other being membrane bound with a molecular mass of 35 kDa (18, 22, 33).

Both glycosylases catalyze the cleavage of the β-1,4-glycosidic bond between *N*-acetylmuramic acid and *N*-acetylglucosamine residues, as do various lysozymes. However, the bacterial glycosylases also catalyze an intramolecular transglycosylation reaction, conserving the energy of the glycosidic bond by the synthesis of a 1,6-anhydrobond in the muramic acid residue (Fig. 1). They are therefore known as soluble and membrane-bound transglycosylases (18, 33). This study concerns the soluble lytic transglycosylase (Slt; EC 3.2.1.-).

Studies on bacterial transglycosylases are interesting for at least two reasons. In the first place, such studies will contribute to a better understanding of the metabolism of murein and might consequently lead to the design of selective inhibitors and thus to the development of a new class of antibiotics. The second reason is that these enzymes can be used for the production of pharmacologically active compounds by specific degradation of murein in vitro. Pharmacological activities of degradation products of murein have been demonstrated in several studies. A sleep-inducing factor, isolated from human urine (31), turned out to be identical to one of the monomeric products (*N*-acetylglucosaminyl-1,6-anhydro-*N*-acetylmuramyl-L-alanyl-D-isoglutamyl-*m*-diaminopimelyl-D-alanine) obtained by degradation of murein with the lytic transglycosylases (18). Picomole amounts of this compound, injected cerebro-intraventricularly, induced excess slow-wave sleep in rabbits (24).

Furthermore, the human pathogens *Bordetella pertussis* and *Neisseria gonorrhoeae* were found to release the structurally identical 1,6-anhydromuropeptides, eliciting cytotoxic effects on ciliated epithelial cells in higher organisms (7, 8, 32, 41).

Recently, Betzner and Keck (3) have published the cloning of the *slt* gene from *E. coli*, encoding the soluble lytic transglycosylase. They obtained a 30-fold overexpression of the *slt* gene. A three-step purification procedure was used to isolate the enzyme to homogeneity, and X-ray-quality crystals were obtained for the determination of the three-dimensional structure of the protein (36). The work described in the present paper was carried out to support studies to

* Corresponding author.

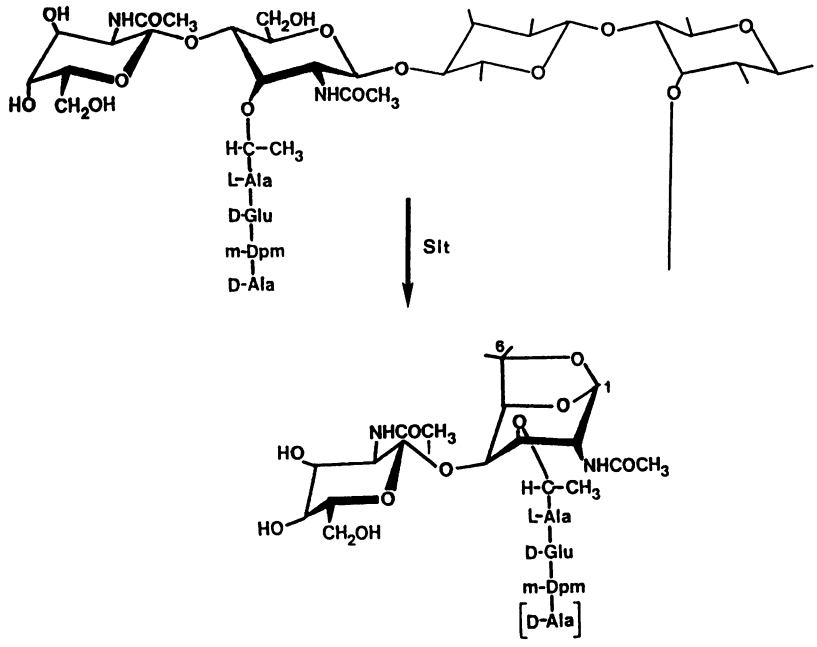

FIG. 1. Enzymatic activity of soluble lytic transglycosylase. Soluble lytic transglycosylase catalyzes the cleavage of the β-1,4-glycosidic bond between *N*-acetylmuramic acid and *N*-acetylglucosamine residues, thereby conserving the energy in a newly synthesized 1,6-anhydrobond. The monomeric products (*N*-acetylglucosaminyl-1,6-anhydro-*N*-acetylmuramyl-L-alanyl-D-isoglutamyl-*m*-diaminopimelyl with or without D-alanine) are shown.

determine the three-dimensional structure and, especially, the structure-function relationship of the enzyme. Since large amounts of highly purified enzyme are needed for both the crystallographic work and future kinetic studies on the interesting enzymatic activity, this study includes work on overexpression of the *slt* gene which enabled us to isolate 100-mg amounts of the enzyme in a one-step purification procedure.

MATERIALS AND METHODS

Plasmids and bacterial strains. Plasmid pAB58, carrying the structural gene for soluble lytic transglycosylase, has been described by Betzner and Keck (3). Plasmid vectors M13mp18 and M13mp19 were those used by Vieira and Messing (50). Phasmids pMa5-8 and pMc5-8 were gifts from H.-J. Fritz (43), and plasmid pJRD187 was obtained from J. Davison (9). Strain CJ236 was purchased from Bio-Rad Laboratories, Richmond, Calif. The helper phage M13KO7 (51) was obtained from Pharmacia LKB Biotechnology Inc., Uppsala, Sweden. *E. coli* strains were grown in Luria-Bertani medium (30) supplemented with the appropriate antibiotics.

Materials. [^{3}H]diaminopimelic acid-labeled murein (specific activity, 10,000 cpm/5 μg of murein) was a kind gift of J.-V. Höltje (Max-Planck-Institut, Tübingen, Germany). Deoxyadenosine 5′-([α-^{35}S]thio)triphosphate was purchased from Amersham, Aylesbury, United Kingdom. Unlabeled deoxynucleotide triphosphates were purchased from Pharmacia Laboratories (Uppsala, Sweden), and dideoxynucleotides were obtained from Boehringer, Mannheim, Germany. All restriction endonucleases, exonuclease *Bal* 31, T4 DNA ligase, Klenow DNA polymerase, and T4 polynucleotide kinase were purchased from Boehringer. Sequenase was obtained from United States Biochemical Corp., Cleveland, Ohio. All enzymes were used under the conditions recommended by the suppliers.

DNA isolation and transformation. Plasmid DNA was purified by cesium chloride-ethidium bromide density gradient centrifugation (30). Small-scale plasmid DNA preparations were purified according to the alkaline extraction procedure (4). Plasmids were introduced into *E. coli* host strains as described previously (29).

DNA sequencing. As a preliminary to the nucleotide sequencing of the 2.5-kb *Pvu*II-*Sal*I chromosomal insert from pAB58, restriction endonuclease fragments or fragments being created by exonuclease *Bal* 31 treatment were subcloned into M13mp18 or M13mp19 (50). The subclones were sequenced by the dideoxyribonucleotide chain termination reaction (37) by using Sequenase. Both strands were completely sequenced. DNA sequences were analyzed with the PC GENE program (Genofit, Geneva, Switzerland). Amino acid sequences were compared with the Swiss-Prot protein data base (EMBL, Heidelberg, Germany) by using the FastP program (27) and the method of Goad and Kanehisha (13).

Construction of Slt overproducers. The 0.9-kb *Eco*RI-*Sal*I fragment of pJRD187, carrying the thermosensitive phage lambda repressor (*c*I857), p_R promoter, Shine-Dalgarno sequence, and ATG triplet of the phage lambda *cro* gene, together with a multiple cloning site, was cloned into the 3.8-kb phasmid pMa5-8 linearized with *Eco*RI and *Sal*I. This resulted in a 4.6-kb plasmid (pHEMa99). The 2.5-kb *Pvu*II-*Sal*I fragment of pAB58, carrying the *slt* gene, was cloned into pHEMa99, which was linearized with *Nru*I and *Sal*I, resulting in a 7.1-kb plasmid (pHEMa112).

Site-directed mutagenesis was performed as described previously (25). Strain CJ236 was used as a host for plasmid pHEMa112 for the incorporation of uridine, with the helper phage M13KO7 for the production of single-stranded DNA (51). Four synthetic oligonucleotides carrying *Asp*718 re-

striction sites were synthesized on an Applied Biosystems (Warrington, United Kingdom) model 380B DNA synthesizer by Eurosequence, Groningen, The Netherlands. The first oligonucleotide was used to mutate the *Bam*HI site of the multiple cloning site into a unique *Asp*718 site, resulting in the plasmid pHEMa112A. The second and third oligonucleotides facilitated cloning of the *slt* gene encoding the transglycosylase with a 28- or 37-amino-acid signal peptide, respectively. The fourth oligonucleotide was used for cloning the *slt* gene without a signal peptide-encoding sequence. The *Asp*718 restriction sites were characterized by restriction and nucleotide sequence analysis. The individual *Asp*718 fragments were then independently deleted in pHEMa112A to facilitate cloning of the *slt* gene downstream of the Shine-Dalgarno sequence, in frame with the ATG start codon of the *cro* gene, and under the control of the temperature-inducible phage lambda p_R promoter. The in-frame cloning was verified by sequence analysis and resulted in the plasmids pHEMa115 (encoding Slt carrying a signal peptide of 37 amino acids), pHEMa140 (encoding Slt carrying a signal peptide of 28 amino acids), and pHEMa116 (encoding Slt without a signal peptide).

Overexpression of the *slt* gene. *E. coli* HB101, harboring pHEMa99, pHEMa115, pHEMa116, or pHEMa140, was grown in 200 ml of Luria-Bertani medium containing 100 μg of ampicillin per ml at 28°C. When the cultures had reached the mid-exponential growth phase (optical density at 600 nm = 0.4), the temperature was shifted to 42°C. Samples (20 ml) were withdrawn 45 and 30 min before and 0, 15, 30, and 45 min after the temperature shift and cooled on ice. Cells were collected by centrifugation (15 min, 4,000 × *g*, 4°C), resuspended in 1 ml of 10 mM Tris-maleate-NaOH buffer (pH 6.8), and lysed by ultrasonication (Vibra cell sonifier equipped with a microtip; Sonics and Materials, Danbury, Conn.) for 4 × 10 s at 100 W. Soluble fractions were obtained by centrifugation of the sonicated cell preparations (homogenates) for 1 h at 10,000 × *g* and 4°C.

The degree of overproduction was estimated by Western blot (immunoblot) analysis with various dilutions of the soluble fractions as described below. Host strain HB101 carrying the vector pHEMa99 served as a control.

The soluble fractions were assayed for enzymatic activity as described below.

Protein concentrations were determined by the method of Bradford (6) with bovine serum albumin as a standard.

SDS-polyacrylamide gel electrophoresis and Western blot analysis. Proteins from homogenates or soluble fractions were separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (28). Either the proteins were stained with Coomassie brilliant blue R-250 or they were transferred to nitrocellulose filters and soluble lytic transglycosylase was stained immunochemically (5). Polyclonal antibodies against soluble lytic transglycosylase were prepared as described previously (22).

Isolation of soluble lytic transglycosylase. Soluble lytic transglycosylase was isolated from *E. coli* HB101 harboring the plasmid pHEMa115. The culture was grown in a 65-liter mobile plant fermentor (New Brunswick, Edison, N.J.) in Luria-Bertani medium supplemented with 0.5% glucose. The bacteria were stirred (55 rpm) and aerated (30 liters/min). After reaching an optical density at 600 nm of 0.4 at 28°C, the temperature was shifted to 42°C and growth was continued for an additional 45 min to a final optical density at 600 nm of 0.7. Cells were then harvested by centrifugation in a continuous centrifuge. The yield from one fermentation was about 70 g of cells (wet weight). The cells were resuspended in 10 mM Tris-maleate-NaOH buffer (pH 6.8) at a concentration of 1 g (wet weight) per ml, disrupted in a French press at 10,000 lb/in^2, and centrifuged at 100,000 × *g* for 1 h. Transglycosylase was isolated from the soluble fraction by a modification of previously described protocols (18, 26). Briefly, the following procedure was used. The soluble fraction (375 ml, 47.0 mg of protein per ml) was dialyzed against 10 mM Tris-maleate-NaOH buffer (pH 6.8) and loaded on a CM-Sepharose CL-6B column (volume of 500 ml) equilibrated with the same buffer. The enzyme was eluted at a concentration of 0.3 M NaCl when a linear gradient from 0.01 to 0.6 M NaCl in 10 mM Tris-Maleate-NaOH buffer (total volume, 1 liter) was applied. To concentrate the enzyme, the Slt-containing fractions (volume, 325 ml; 0.56 mg of protein per ml) were dialyzed against 10 mM potassium phosphate buffer (pH 6.8), and loaded on a hydroxylapatite column (volume, 20 ml) equilibrated with the dialysis buffer. Fractions containing the concentrated enzyme were eluted at a concentration of 0.1 M when a gradient of 0.01 to 0.4 M potassium phosphate buffer (total volume, 100 ml) was applied, yielding 175 mg of transglycosylase up to a concentration of 9.3 mg of protein per ml.

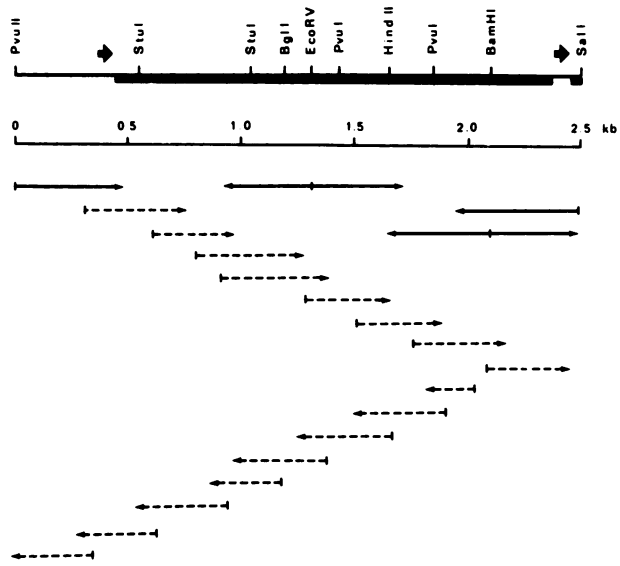

FIG. 2. Strategy used for sequencing the *slt* gene. The 2.5-kb *Pvu*II-*Sal*I DNA fragment from plasmid pAB58, containing the complete *slt* gene and the N-terminal encoding sequence of the *trpR* gene, is shown. The two coding regions are indicated by thick bars. Initially, bases were determined from the *Pvu*II and *Sal*I sites in one direction and from *Bam*HI and *Eco*RV sites in both directions (thin solid arrows). To read bases that were not obtained from these fragments, *Bal* 31 exonuclease digestions were performed from the *Pvu*II, *Bam*HI, and *Eco*RV sites to produce fragments which allowed further reading (broken arrows). The two thick arrows indicate the direction of transcription of the *slt* and *trpR* genes.

Enzyme assay. Murein degradation was assayed by using isolated murein labeled with [^{3}H]diaminopimelic acid (18) as a substrate. The release of low-molecular-weight muropeptides from the murein polymer was determined essentially as described earlier (2), except for omission of 0.2% Triton X-100 from the reaction mixture. Release of labeled murein fragments was not strictly proportional to the amount of protein present in the assay.

Amino acid analysis and N-terminal sequence analysis of transglycosylase. Amino acid analysis of the purified protein

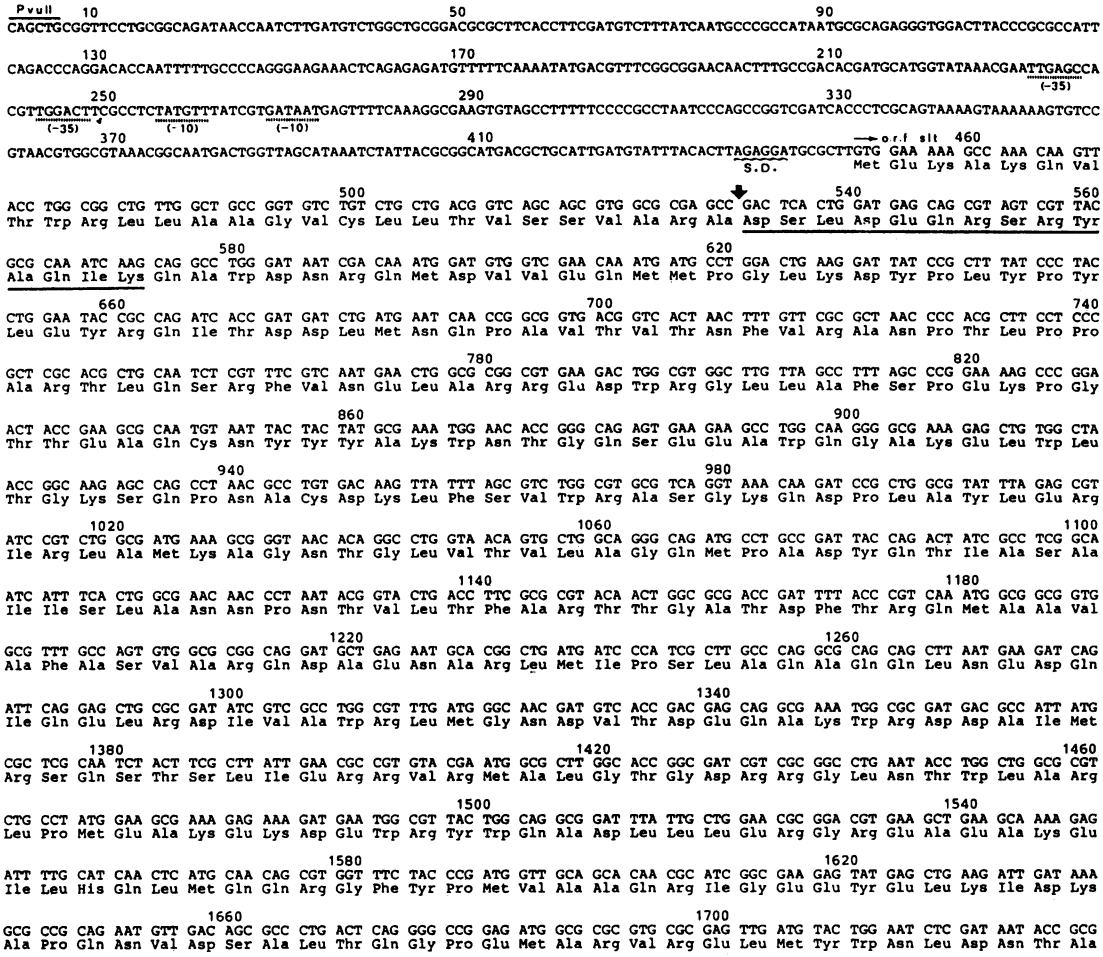

FIG. 3. Nucleotide sequence of the *slt* gene and deduced amino acid sequence. Shown is the 2,548-bp *Pvu*II-*Sal*I fragment of plasmid pAB58. The *slt* coding region starts at nucleotide 450 (GTG) and ends at nucleotide 2385 (TGA). The *trpR* open reading frame (o.r.f) starts at nucleotide 2477 (ATG). The corresponding amino acids are shown directly below the triplet codons in the three-letter code. The proposed Shine-Dalgarno ribosome binding site (S.D.) is wavy underlined, and a vertical arrow indicates the signal peptidase cleavage site between amino acids 27 and 28. The N-terminal amino acid sequence of the mature enzyme, indicated by a solid line, was verified by sequencing of the purified protein. The two proposed −35 and −10 regions are indicated by stippled lines. The putative transcriptional terminator sequence has been indicated with horizontal arrows, and the nucleotides marked with dots indicate the mismatch in the stem of the proposed transcriptional terminator structure.

TABLE 1. Codon usage derived from the sequence of mature soluble lytic transglycosylase

Residue and codon[a]	Total no. of codons	(%)[b] Frequency
Phe UUU	7	41
Phe UUC	10	59
Leu UUA	4	7
Leu UUG	8	15
Leu CUU	9	17
Leu CUC	2	4
Leu CUA	2	4
Leu CUG	29	54
Ile AUU	12	50
Ile AUC	12	50
Ile AUA	0	0
Met AUG	22	100
Val GUU	5	17
Val GUC	7	24
Val GUA	4	14
Val GUG	13	45
Ser UCU	3	9
Ser UCC	2	6
Ser UCA	6	17
Ser UCG	4	11
Ser AGU	6	17
Ser AGC	14	40
Pro CCU	8	29
Pro CCC	5	18
Pro CCA	2	7
Pro CCG	13	46
Thr ACU	6	17
Thr ACC	15	41
Thr ACA	5	14
Thr ACG	10	28
Ala GCU	11	16
Ala GCC	18	26
Ala GCA	10	14
Ala GCG	30	44
Tyr UAU	14	48
Tyr UAC	15	52
Ter UAA	0	0
Ter UAG	0	0
His CAU	3	100
His CAC	0	0
Gln CAA	18	43
Gln CAG	24	57
Asn AAU	16	52
Asn AAC	15	48
Lys AAA	14	54
Lys AAG	12	46
Asp GAU	24	75
Asp GAC	8	25
Glu GAA	19	50
Glu GAG	19	50
Cys UGU	2	100
Cys UGC	0	0
Ter UGA	1	100
Trp UGG	19	100
Arg CGU	23	49
Arg CGC	15	32
Arg CGA	4	8
Arg CGG	5	11
Arg AGA	0	0
Arg AGG	0	0
Gly GGU	7	20
Gly GGC	13	37
Gly GGA	5	14
Gly GGG	10	29

[a] The amino acids corresponding to the eight rare codons are underlined.
[b] Percentage of codons used for each amino acid.

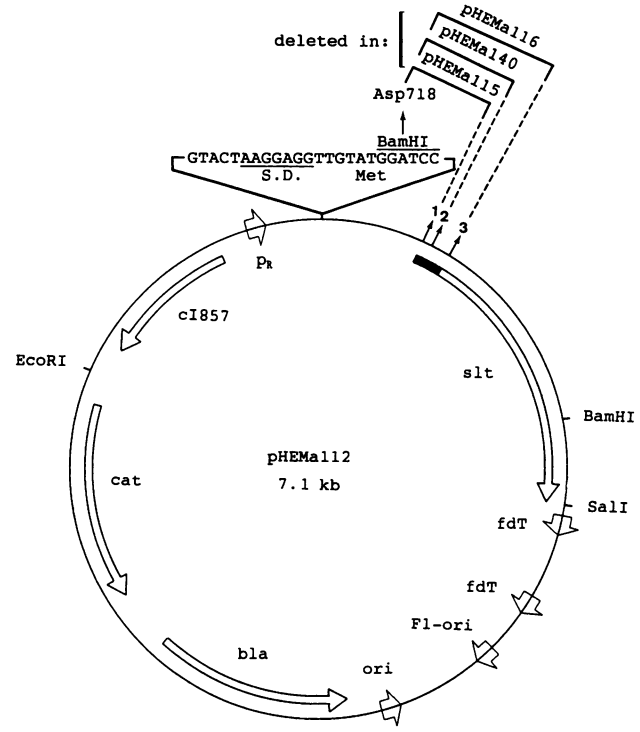

FIG. 4. Map of the plasmid pHEMa112 that was used for the construction of the Slt-overproducing plasmids. Some restriction sites are indicated. Open boxes show the relative positions of the genes coding for soluble lytic transglycosylase (*slt*), phage lambda repressor (*c*I857), β-lactamase (*bla*), and chloramphenicol acetyltransferase (*cat*). The filled box represents the nucleotide sequence encoding the largest possible signal peptide of Slt. The arrows indicate the direction of transcription. The origin of replication of phage f1 (F1-*ori*), the central transcription terminator of phage fd (fdT), the ColE1 origin of replication (*ori*), and the phage lambda p_R promoter are also indicated. The nucleotide sequence of the Shine-Dalgarno ribosome binding site (S.D.) and the triplet of the start codon (ATG) of the phage lambda *cro* gene as well as the flanking sequence of the ATG are given. The positions of the independently created *Asp*718 restriction sites are numbered from 1 to 3. The primers used are described in Materials and Methods and in Fig. 5. The different constructs, obtained after independently deleting the small *Asp*718 fragments, are described in Fig. 5.

was performed on a Liquimat III analyzer (Kontron, Zürich, Switzerland) after hydrolysis in 6 M HCl at 110°C in evacuated sealed glass tubes for 24 h. Tryptophan was quantitively determined after hydrolysis with 3 M mercaptoethanesulfonic acid.

N-terminal sequence analysis was performed on an Applied Biosystems model 477A protein sequencer (pulse-liquid sequenator) connected on line with a 120A PTH analyzer.

Nucleotide sequence accession number. Primary nucleotide and amino acid sequence data reported in this paper have been deposited in the GenBank Submissions of the Los Alamos National Laboratory and are available under accession number M69185.

RESULTS

Nucleotide sequence of the *slt* gene. The gene encoding the soluble lytic transglycosylase from *E. coli* had been cloned

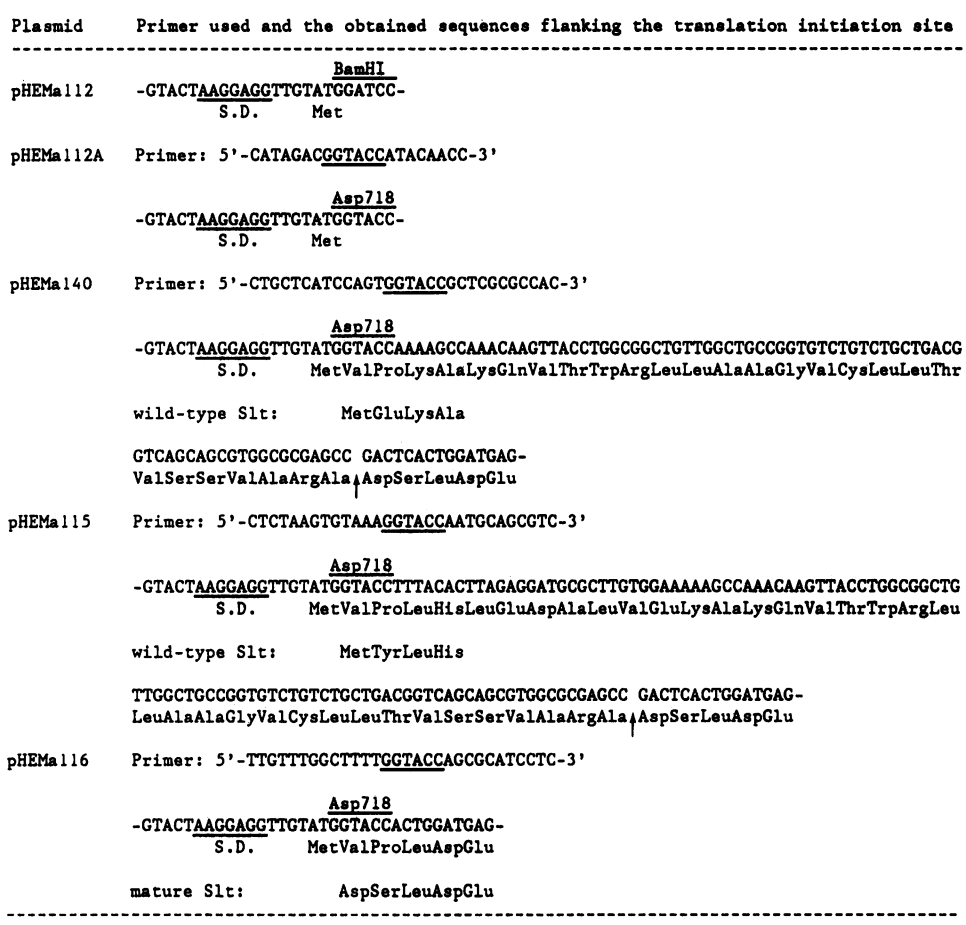

FIG. 5. Nucleotide sequences flanking the translation initiation signals in the constructed plasmids. The primers that were used to create *Asp*718 restriction sites as well as the obtained nucleotide sequences and their corresponding amino acid sequences are shown. S.D. indicates the Shine-Dalgarno sequence of the phage lambda *cro* gene. The arrow indicates the signal peptidase cleavage site. The N-terminal amino acid sequences of the signal peptide and the mature Slt of wild-type *E. coli* are given below the artificial amino acid sequences.

previously on the plasmid pAB58 (3). In order to determine its nucleotide sequence, subfragments of the 2.5-kb *Pvu*II-*Sal*I fragment of pAB58 were subcloned into M13 vectors and then sequenced by the Sanger dideoxy method. The sequence strategy used is depicted in Fig. 2. The complete nucleotide sequence and the corresponding open reading frames are shown in Fig. 3. One large open reading frame and one truncated open reading frame were identified. The latter was identified as the N-terminal sequence of the *trpR* gene product, and the large open reading frame belongs to Slt. The initiation of translation can start either at nucleotide 423 (AUG) or at nucleotide 450 (GUG) and terminates in both cases at nucleotide 2385 (UGA), encoding a polypeptide of 654 or 645 amino acid residues, respectively. A computer-assisted search for a signal peptide revealed that the N-terminal 36 amino acids of the polypeptide consisting of 654 amino acids as well as the N-terminal 27 amino acids of the polypeptide consisting of 645 amino acids are recognized as cleavable signal peptides, resulting in both cases in a mature transglycosylase of 618 residues.

Promoter sequence, ribosome binding site, and transcription terminator. We have identified two possible promoter sequences upstream of the encoding region, as shown in Fig. 3, showing significant similarity to −35 and −10 consensus promoter sequences and classified as medium-strength promoters (17, 34). We do find a putative ribosome binding site upstream of the 27-amino-acid-encoding signal peptide sequence which shows homology to the consensus Shine-Dalgarno sequence (39). However, no ribosome binding site was found upstream of the 36-amino-acid-encoding signal peptide sequence. A strong transcriptional terminator structure consisting of a GC-rich stem of 11 bp and a loop of 3 bases, followed by an AT-rich sequence, as depicted in Fig. 3, has also been proposed as the promoter-operator region of the *trpR* gene, suggesting that expression of the *trpR* gene is autoregulatory (16). We propose that this stable secondary structure might exhibit both functions simultaneously.

Rare codon usage in the *slt* gene. It has been shown previously (23) that codon usage of 25 nonregulatory *E. coli* genes is not random. The percentage of usage of the eight rare codons (ATA [Ile], TCG [Ser], CCU and CCC [Pro], ACG [Thr], CAA [Gln], AAT [Asn], and AGG [Arg] [23]) in the reading frame encoding the mature Slt (Table 1) is the same as that reported for the *dnaG* (42), *lacI* (11), *araC* (44), and *trpR* (16, 40) genes mentioned above. In the *slt* gene reading frame, rare codons account for 9.9% of all codons; for the *dnaG*, *lacI*, *araC*, and *trpR* genes, they account for 11.3, 10.0, 9.2, and 6.7%, respectively. The combination of unusual codons, a poor ribosome binding site, and the

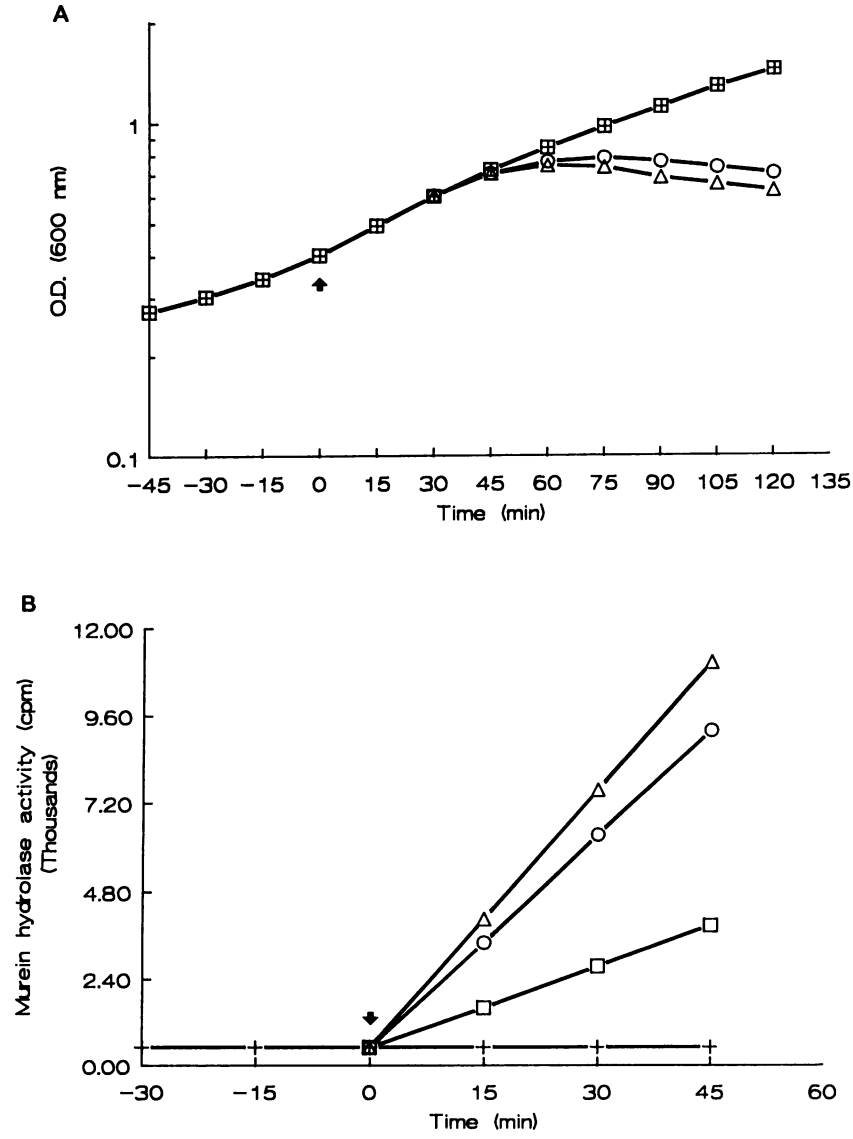

FIG. 6. Growth curves and in vitro murein hydrolase activity of Slt-overproducing cells and controls. Arrows indicate the temperature shift to 42°C. (A) Growth curves for HB101 cells carrying constructed plasmids: pHEMa99 (+, vector), pHEMa140 (△, Slt with 28-amino-acid signal peptide), pHEMa115 (○, Slt with 37-amino-acid signal peptide), and pHEMa116 (□, Slt without a signal peptide). O.D., optical density. (B) In vitro murein degradation with soluble fractions from cells carrying the same constructed plasmids. Soluble protein (500 ng) (10,000 × *g* supernatant fraction of homogenate) was assayed for enzymatic activity. The release of labeled murein fragments is not strictly proportional to the amount of protein used in the assay.

translation initiation codon GUG reduces both the frequency of initiation and the rate of passage of ribosomes along the mRNA and could account for the relatively low expression rate of this putative autolysin found in wild-type cells (36).

Rare codon usage in the noncoding reading frames of the *slt* gene. In most *E. coli* proteins, the frequency of rare codons in noncoding reading frames is approximately three times that in coding reading frames (1, 23). This relationship does not hold for the *slt* gene (9.9% rare codon usage in the reading frame versus 12.8 and 9.5% in the noncoding reading frames), as has been found for the *dnaG* (42), *lacI* (11), *araC* (44), and *trpR* (16, 40) genes.

Regulated high-level overexpression of the *slt* gene. Overproducers of Slt were constructed by using the plasmid pHEMa112 (Fig. 4). Because of the presence of the two possible translation initiation sites, we constructed Slt overproducers carrying an either 37- or 28-amino-acid signal peptide as well as Slt without a signal peptide (Fig. 5).

A 250-fold overproduction of Slt carrying a signal peptide of 28 amino acids was achieved 45 min after temperature induction (42°C). When growth was continued at 42°C, the optical density of the culture decreased somewhat, suggesting growth inhibition and lysis of some of the cells (Fig. 6 and 7). The overproduced Slt was almost completely processed to the mature form and localized in the soluble enzyme fraction (Fig. 7, lanes 2 and 4). An *slt* gene that carried a signal peptide of 37 amino acids also resulted in a 250-fold overproduction of soluble enzyme and decrease of the optical density of the culture 45 min after temperature

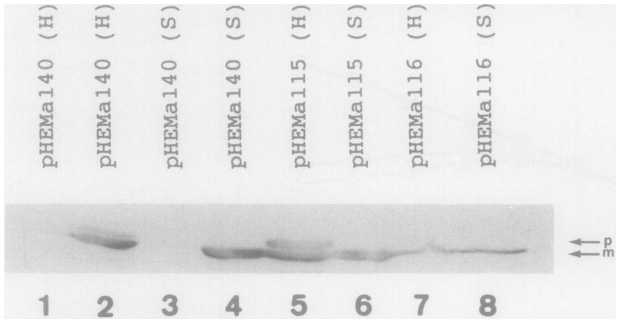

FIG. 7. Western blot analysis of homogenates and 10,000 × *g* soluble fractions of Slt overproducers. Slots contained either homogenates (H) or soluble fractions (S) of cells transformed with pHEMa140 (codes for Slt with 28-amino-acid signal peptide), pHEMa115 (codes for Slt with 37-amino-acid signal peptide), or pHEMa116 (codes for Slt without a signal peptide). Lanes 1 and 3 contained material from cells before induction; the other lanes were loaded with protein obtained 45 min after induction. All lanes contained 10 μg of protein. The position of the preprotein (p) and mature Slt (m) are indicated by arrows.

induction (Fig. 6 and 7). However, in this case the cells contained about 50% unprocessed Slt (Fig. 7, lanes 5 and 6).

In the cytoplasm, Slt can be overproduced up to 100-fold, as is also shown in Fig. 6 and 7. The cytoplasmic Slt is enzymatically active and localized in the soluble enzyme fraction (Fig. 7, lanes 7 and 8).

N-terminal sequence and amino acid analysis of the protein. The single-step isolation of Slt is shown in Fig. 8. Sequencing of the first 14 N-terminal amino acids confirmed that the identified open reading frame represents the structural gene of the soluble lytic transglycosylase. Furthermore, it confirmed that a signal peptide is cleaved off during the maturation and transport of the transglycosylase across the cytoplasmic membrane, as predicted from the nucleotide sequence. The amino acid composition was determined after acid hydrolysis of the purified protein. The determined amino acid composition is in very close agreement with the composition as deduced from the nucleotide sequence, supporting the assignment of the open reading frame.

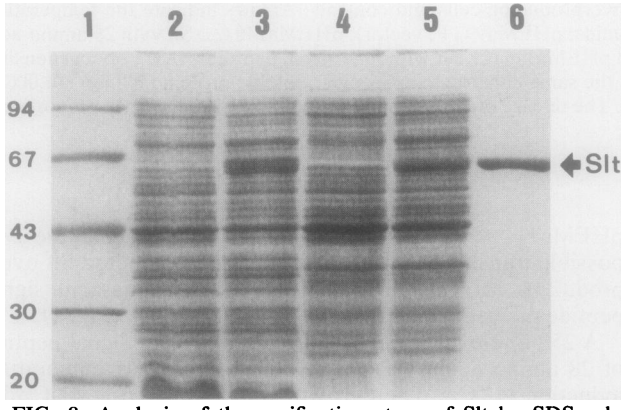

FIG. 8. Analysis of the purification steps of Slt by SDS-polyacrylamide gel electrophoresis. Lane 1, molecular mass markers (sizes given in kilodaltons); lanes 2 and 3, total protein before and 45 min after temperature induction, respectively; lanes 4 and 5, soluble protein before and 45 min after temperature induction, respectively; lane 6, enzyme pool obtained after eluting Slt from CM-Sepharose CL-6B. Lanes 2 to 5 contain 20 μg of protein; lane 6 contains 1 μg of protein. The 10% polyacrylamide gel was stained with Coomassie brilliant blue R-250.

DISCUSSION

Nucleotide sequence and initiation of translation of the *slt* gene. The largest open reading frame that might contain the coding region for the structural gene of soluble lytic transglycosylase codes for a protein of 654 amino acids. Analysis of the primary structure revealed that the N-terminal 36 amino acids of this polypeptide, despite being unusually long, contained most of the typical features of a signal peptide (53), including (i) an initial group of charged, hydrophilic amino acids followed by (ii) a group of 8 to 15 hydrophobic amino acids and then (iii) small hydrophobic amino acids such as Ala at positions −1 and −3, counting from the cleavage site, and (iv) a sequence that can be cleaved by signal peptidase I. However, all signal peptides in gram-negative bacteria reported in the literature are smaller than 30 amino acids. This makes the presence of such a long signal peptide in Slt less likely and prompted us to search for an alternative translational initiation site.

Since it is known that in *E. coli* occasionally a translational initiation site different from AUG is used (45), we suggest that translation starts at nucleotide 450 (GUG), which fits the same reading frame. The triplet GUG normally codes for an internal Val, but it is known that the tRNA carrying formylmethionyl can recognize the GUG codon in vivo (52). This would result in a signal peptide of 27 amino acids that still exhibits all the typical features of a signal peptide. Moreover, we did not find a consensus ribosome binding site upstream of the nucleotide sequence encoding the putative 36-amino-acid signal peptide, but we do find one upstream of the 27-amino-acid-encoding signal peptide nucleotide sequence. The putative cleavage site stays the same for both signal peptides.

The Perceptron algorithm of Stormo et al. (46) was used to predict which of the two codons was the more likely initiation site for the translation of *slt* mRNA. Although both the triplet for AUG at nucleotides 423 to 425 and the triplet for GUG at nucleotides 450 to 452 gave low scores, the program points to GUG as the more likely position.

Controlled overexpression of the *slt* gene. On the basis of the discovery of two possible translation initiation sites, two Slt overproducers with an either 37- or 28-amino-acid signal peptide were constructed. The signal peptides carry, because of the creation of suitable restriction sites by means of site-directed mutagenesis, one additional and two mutated amino acids at the N terminus compared with the wild-type sequence (Fig. 5). The 28-amino-acid signal peptide is efficiently processed, even when the transglycosylase is 250-fold overproduced (Fig. 7), suggesting that minor changes at the N terminus have no effect on the processing of this signal peptide. However, the 37-amino-acid signal peptide is not efficiently processed at about the same level of overproduction (Fig. 7), and about 50% of the overproduced enzyme is present as a preform in the sedimentable fraction. These data support the suggestion made in the previous section, i.e., that translation initiates at nucleotide 450 and not at nucleotide 423.

Physiological function of the transglycosylase. Slt has been suggested to be a cytoplasmic enzyme (18, 19, 22, 33). Our results clearly show that a signal peptide is cleaved off during maturation, suggesting transport of the transglycosylase into the periplasmic space.

```
442  WDHLEERFPLAYNDLFKRYTSGKEIPQSYAMAIARQESAWNPKVKSPVGAS  492
     :+ +++ : +:++::++ +++++++ +  + +:+ :: + :+++: +:+
1    MDKYDKNVPSDYDGLFQKAADANGVSYDLLRKVAWTESRFVPTAKSKTGPL   51

493  GLMQIMPGTATHTVKMFSIPGYSSPGQLLDPETNINIGTSYLQYVYQQFGN  543
     :+::+  +::+ ++ + ++++ ++ ++ :+:: +:: +++ :+ +  +:++
52   GMMQFTKATAK-ALGL-RVTD-GPDDDRLNPELAINAAAKQLAGLVGKFDG   99

544  NRIFSSAAYNAGLGRVRTWLGNSAGRIDAVAFVESIPFSETRGYVKNVLA  593
     + + ++ :::+: ::    :::++   + +  +::+ +:+:+:++:+:
100  DELKAALAYNQGEGR----LGNPQLEAYSKGDFASIS-EEGRNYMRNLLD  144
```

FIG. 9. Comparison of the amino acid sequences of the soluble lytic transglycosylase from *E. coli* (upper sequence) and the internal virion protein D from bacteriophage T7 (lower sequence). Numbers denote the amino acid positions in the respective primary structures. Amino acid identities are indicated by double dots, conservative residues are indicated by plusses and gaps in the amino acid sequence alignment are indicated by dashes.

The physiological function of Slt remains unclear. The enzyme might play a role in recycling of muropeptides (14, 15, 20) during cell elongation and/or cell division, when cleavage of covalent bonds is required. The energy, conserved in the 1,6-anhydrobond, might be used for the formation of a new β-1,4-glycosidic bond (18).

Cells grow normally when the *slt* gene has been deleted from the chromosome (35). It is possible that in these cells, the function of the soluble enzyme is taken over by the membrane-bound lytic transglycosylase (33). It will be interesting to investigate whether *E. coli* can survive in the absence of both transglycosylases.

The massive overproduction of Slt by cells carrying the pHEMa140 or pHEMa115 plasmid does not lead to lysis of the majority of the cells (Fig. 6). This indicates that in these cells the enzyme must, in some way, be prevented from degrading the murein polymer.

Comparison of the primary structure of transglycosylase with those of other proteins. The amino acid sequence of the soluble lytic transglycosylase from *E. coli* was compared with those of other proteins present in the Swiss-Prot protein sequence data base (16,941 entries) by using the FastP program (27). Alignment with the program of Goad and Kanehisha (13) gave essentially the same result (Fig. 9).

The primary structure of transglycosylase does not show significant similarity to those of other known carbohydrate-degrading enzymes, including phage lambda endolysin (M_r = 17,500), the only protein possessing an identical enzymatic activity whose primary structure is known (21, 48). The *E. coli* transglycosylase does show similarity to the internal virion protein D from bacteriophage T7 (10). A stretch of 151 amino acids at the C terminus of Slt shows significant similarity to the N-terminal part of the internal virion protein D (Fig. 9). About 26% of the amino acid residues in the two sequences are identical, and 71% are similar if conservative substitutions are included. The function or enzymatic activity of this protein, late expressed in the phage cycle, is not known, but it is feasible that it plays a role in lysis of the host cell by the bacteriophage. Whether the C terminus of the transglycosylase is necessary for the cleavage of the β-1,4-glycosidic bond will be studied by constructing N- and C-terminal deletions.

ACKNOWLEDGMENTS

We are grateful to P. Terpstra for help with the computer analysis, P. Jekel for the amino acid composition analyses, and J. M. W. Bouma for carefully reading the manuscript.

REFERENCES

1. **Alff-Steinberger, C.** 1984. Evidence for a coding pattern of the non-coding strand of the *E. coli* genome. Nucleic Acids Res. **12**:2235–2241.
2. **Beachy, E. H., W. Keck, M. A. de Pedro, and U. Schwarz.** 1981. Exoenzymatic activity of transglycosylase isolated from *Escherichia coli*. Eur. J. Biochem. **116**:355–358.
3. **Betzner, A., and W. Keck.** 1989. Molecular cloning, overexpression and mapping of the *slt* gene encoding the soluble lytic transglycosylase of *Escherichia coli*. Mol. Gen. Genet. **219**:489–491.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
5. **Bittner, M., P. Kupferer, and C. F. Morris.** 1980. Electrophoretic transfer of proteins and nucleic acids from slab gels to diazobenzyloxymethyl cellulose or nitrocellulose sheets. Anal. Biochem. **102**:459–471.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
7. **Cookson, B. T., H.-L. Cho, L. A. Herwaldt, and W. E. Goldman.** 1989. Biological activities and chemical composition of purified tracheal cytotoxin of *Bordetella pertussis*. Infect. Immun. **57**: 2223–2229.
8. **Cookson, B. T., A. N. Tyler, and W. E. Goldman.** 1989. Primary structure of the peptidoglycan-derived tracheal cytotoxin of *Bordetella pertussis*. Biochemistry **28**:1744–1749.
9. **Davison, J., M. Heusterpreute, N. Chevalier, and F. Brunel.** 1987. A 'phase-shift' fusion system for the regulation of foreign gene expression by lambda repressor in gram-negative bacteria. Gene **60**:227–235.
10. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. **166**:477–535.
11. **Farabough, P. J.** 1978. Sequence of the *lacI* gene. Nature (London) **274**:765–769.
12. **Ghuysen, J. M., and G. D. Schockman.** 1973. Biosynthesis of peptidoglycan, p. 37–130. *In* L. Leive (ed.), Bacterial membranes and walls, 1st ed. Marcel Dekker, Inc., New York.
13. **Goad, W. B., and M. J. Kanehisha.** 1982. Pattern recognition in nucleic acid sequences. I. A general method for finding local homologies and symmetries. Nucleic Acids Res. **10**:247–263.
14. **Goodell, E. W.** 1985. Recycling of murein by *Escherichia coli*. J. Bacteriol. **163**:305–310.
15. **Goodell, E. W., and U. Schwarz.** 1985. Release of cell wall peptides into culture medium by exponentially growing *Escherichia coli*. J. Bacteriol. **162**:391–397.
16. **Gunsalus, R. P., and C. Yanofsky.** 1980. Nucleotide sequence and expression of *Escherichia coli trp*R, the structural gene for the trp aporepressor. Proc. Natl. Acad. Sci. USA **77**:7117–7121.
17. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
18. **Höltje, J.-V., D. Mirelman, N. Sharon, and U. Schwarz.** 1975. Novel type of murein transglycosylase in *Escherichia coli*. J. Bacteriol. **124**:1067–1076.
19. **Höltje, J.-V., and U. Schwarz.** 1985. Biosynthesis and growth of the murein sacculus, p. 77–119. *In* N. Nanninga (ed.), Molecular cytology of *Escherichia coli*. Academic Press Ltd., London.
20. **Höltje, J.-V., and E. I. Tuomanen.** 1991. The murein hydrolases of *Escherichia coli*: properties, functions and impact on the course of infections in vivo. J. Gen. Microbiol. **137**:441–454.
21. **Imada, M., and A. Tsugita.** 1971. Amino-acid sequence of λ phage endolysin. Nature (London) **233**:230–231.
22. **Keck, W., F. Wientjes, and U. Schwarz.** 1985. Comparison of two murein transglycosylases of *Escherichia coli*. Eur. J. Biochem. **148**:493–497.
23. **Konigsberg, W., and G. N. Godson.** 1983. Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80**:687–691.
24. **Krueger, J. M., M. L. Karnovsky, S. A. Martin, J. R. Pappenheimer, J. Walter, and K. Biemann.** 1984. Peptidoglycans as

promoters of slow-wave sleep. II. Somnogenic and pyrogenic activities of some naturally occuring muramyl peptides; correlations with mass spectrometric structure determination. J. Biol. Chem. **259**:12659–12662.
25. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
26. **Kusser, W., and U. Schwarz.** 1980. *Escherichia coli* murein transglycosylase. Purification by affinity chromatography and interaction with polynucleotides. Eur. J. Biochem. **103**:277–285.
27. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
28. **Lugtenberg, B., J. Meyers, R. Peters, P. van der Hoek, and L. van Alphen.** 1975. Electrophoretic resolution of the 'major outer membrane protein' of *Escherichia coli* K12 into four bands. FEBS Lett. **58**:254–258.
29. **Mandel, M., and A. Higa.** 1970. Calcium-dependent bacteriophage DNA infection. J. Mol. Biol. **53**:159–162.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Martin, S. A., M. L. Karnovsky, J. M. Krueger, J. R. Pappenheimer, and K. Biemann.** 1984. Peptidoglycans as promoters of slow-wave sleep. I. Structure of the sleep-promoting factor isolated from human urine. J. Biol. Chem. **259**:12652–12658.
32. **Melly, M. A., Z. A. McGee, and R. S. Rosenthal.** 1984. Ability of monomeric peptidoglycan fragments from *Neisseria gonorrhoeae* to damage human fallopian-tube mucosa. J. Infect. Dis. **149**:378–386.
33. **Mett, H., W. Keck, A. Funk, and U. Schwarz.** 1980. Two different species of murein transglycosylases in *Escherichia coli*. J. Bacteriol. **144**:45–52.
34. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict in vitro RNA polymerase selectivity. Nucleic Acids Res. **12**:789–800.
35. **Roeder, W., and R. L. Sommerville.** 1979. Cloning the *trpR* gene. Mol. Gen. Genet. **176**:361–368.
36. **Rozeboom, H. J., B. W. Dijkstra, H. Engel, and W. Keck.** 1990. Crystallization of the soluble lytic transglycosylase from *Escherichia coli*. J. Mol. Biol. **212**:557–559.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1979. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
38. **Schwarz, U., A. Asmus, and H. Frank.** 1969. Autolytic enzymes and cell division of *Escherichia coli*. J. Mol. Biol. **41**:419–429.
39. **Shine, J., and L. Dalgarno.** 1976. The 3′-terminal sequence of *Escherichia coli* 16 S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
40. **Singleton, C. K., W. D. Roeder, G. Bogosian, R. L. Sommerville, and H. L. Weith.** 1980. DNA sequence of the *Escherichia coli trp*R gene and prediction of the amino acid sequence of the Trp repressor. Nucleic Acids Res. **8**:1551–1560.
41. **Sinha, R. K., and R. S. Rosenthal.** 1980. Release of soluble peptidoglycan from growing gonococci: demonstration of anhydro-muramyl-containing fragments. Infect. Immun. **29**:914–924.
42. **Smiley, B. L., J. R. Lupski, P. S. Svec, R. McMacken, and G. N. Godson.** 1982. Sequences of the *Escherichia coli dna*G primase gene and regulation of its expression. Proc. Natl. Acad. Sci. USA **79**:4550–4554.
43. **Stanssens, P., C. Opsomer, Y. M. McKeown, W. Kramer, M. Zabeau, and H.-J. Fritz.** 1989. Efficient oligonucleotide-directed construction of mutations in expression vectors by the gapped duplex DNA method using alternating selectable markers. Nucleic Acids Res. **17**:4441–4454.
44. **Stoner, C. M., and R. F. Schleif.** 1982. Is the amino acid but not the nucleotide sequence of the *Escherichia coli ara*C gene conserved? J. Mol. Biol. **154**:649–652.
45. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *Escherichia coli*. Nucleic Acids Res. **10**:2971–2996.
46. **Stormo, G. D., T. D. Schneider, L. Gold, and A. Ehrenfeucht.** 1982. Use of the 'Perceptron' algorithm to distinguish translational initiation sites in *E. coli*. Nucleic Acids Res. **10**:2997–3011.
47. **Tamura, T., Y. Imae, and J. L. Strominger.** 1976. Purification to homogeneity and properties of two D-alanine carboxypeptidases I from *Escherichia coli*. J. Biol. Chem. **251**:414–423.
48. **Taylor, A., B. C. Das, and J. van Heijenoort.** 1975. Bacterial-cell-wall peptidoglycan fragments produced by phage λ or Vi II endolysin and containing 1,6-anhydro-N-acetylmuramic acid. Eur. J. Biochem. **53**:47–54.
49. **Tomioka, S., and M. Matsuhashi.** 1978. Purification of penicillin-insensitive DD-endopeptidase, a new cell wall peptidoglycan-hydrolyzing enzyme in *Escherichia coli*, and its inhibition by deoxyribonucleic acids. Biochem. Biophys. Res. Commun. **84**:978–984.
50. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
51. **Vieira, J., and J. Messing.** 1987. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. Methods Enzymol. **100**:468–500.
52. **Volckaert, G., and W. Fiers.** 1973. Studies on the bacteriophage MS2. G-U-G as the initiation codon of the A-protein cistron. FEBS Lett. **35**:91–96.
53. **von Heyne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
54. **Weidel, W., and H. Pelzer.** 1964. Bagshaped macromolecules—a new outlook on bacterial cell walls, p. 193–232. *In* F. F. Nord (ed.), Advances in Enzymology, vol. 26. Interscience Publishers, New York.