Identification of the *Escherichia coli tonB* Gene Product in Minicells Containing *tonB* Hybrid Plasmids

Kathleen Postle† and William S. Reznikoff

Department of Biochemistry, University of Wisconsin-Madison Madison, Wisc. 53706, U.S.A.

Sodium dodecyl sulfate/polyacrylamide gel analysis of proteins encoded by a series of *tonB*+ plasmids in minicells has identified the *tonB* gene product as a protein with an apparent molecular weight of 36,000. A parallel analysis of seven *tonB* mutations which have been genetically crossed onto a *tonB*+ plasmid supports this identification; the 36,000 M_r protein is absent from the set of proteins encoded by each *tonB*⁻ plasmid. Four of the *tonB* mutations are apparently IS1 insertions. The locations of these insertions within *tonB* have been determined by restriction endonuclease mapping. Correlation of these IS1 insertion sites with the molecular weights of prematurely terminated *tonB* polypeptides, suggests that *tonB* is transcribed in the direction opposite to that of the nearby tryptophan operon. In addition, a protein encoded by one of the inverted repeat sequences of the transposable element Tn5 has been tentatively identified.

1. Introduction

The product of the *tonB* gene in *Escherichia coli* appears to play a key role in a number of cell surface-related phenomena. Mutations in *tonB* are pleiotropic, affecting colicin sensitivity (Pugsley & Reeves, 1976,1977; Konisky *et al.*, 1976), bacteriophage sensitivity (Matsushiro, 1963), three different high affinity iron transport systems (Wang & Newton, 1969; Frost & Rosenberg, 1973,1975; Wayne & Neilands, 1975; for a review see Neilands, 1977), vitamin B_{12} transport (Bassford *et al.*, 1976) and iron-dependent modification of certain transfer RNAs (Rosenberg & Gefter, 1969; McCray & Herrmann, 1976). Aside from these observations concerning the phenotypes of *tonB* mutations, little is known about the *tonB* gene product itself. There is evidence that in the cases of bacteriophage T1 and ϕ80 infections and vitamin B_{12} transport, that the *tonB* gene product is involved in the energy-requiring steps for these processes (Hancock & Braun, 1976; Bassford *et al.*, 1976). Recent data suggest that *tonB* mutations affect an element in the inner membrane (Wookey & Rosenberg, 1978). It also appears that the *tonB* gene product is functionally unstable (Bassford *et al.*, 1977) and that it is consumed during its action (Kadner & McElhaney, 1978).

In a previous communication we described the *Hind*II and *Hind*III restriction endonuclease map of the *att*ϕ80-*tonB*-*trp* region and localized the *tonB* gene within a 1730-base-pair *Hind*II fragment (Postle & Reznikoff, 1978). In this paper we have

utilized minicells to analyze the proteins encoded by several cloned $tonB^+$ restriction fragments as well as fragments carrying $tonB$ mutations. Such analyses allow us to identify the $tonB$ gene product. Indirect evidence suggests the direction of $tonB$ transcription. In addition a potential promoter region for the $tonB$ gene has been located.

2. Materials and Methods

(a) Media

Growth medium for minicells is the low-phosphate medium of Landy et al. (1967) supplemented to 0·5% (w/v) glucose, 10^{-3} M-MgSO$_4$, 1 μg B$_1$/ml, 100 μg/ml each of methionine, isoleucine, histidine, and valine, 40 μg adenine/ml, and 2 μg pyridoxine/ml.

Labelling medium for minicells is sterile Methionine Assay Medium (Difco), diluted 1 : 3 with M9 salts (per l: 6 g Na$_2$HPO$_4$; 3 g KH$_2$PO$_4$; 0·5 g NaCl; 1·0 g NH$_4$Cl) and the mixture supplemented to 0·5% glucose, 40 μg adenine/ml, and 2 μg pyridoxine/ml.

(b) Bacteria and bacteriophage strains

Bacterial strain Ymel ΔA is F$^-$ $supF$ transduced to $\Delta(tonBtrpA)_{553}$ (obtained from C. Yanofsky) with P1vir. The minicell-producing strain χ984 is F$^-$$minA1$, tsx63, $purE$41, $supE$42, $pdxC$3, $minB$2, his53, $metC$65, $rpsL$97, T3^R, xyl-14, ilv277, $cycB$2, $cycA$1 and was obtained from R. Curtiss III. Plasmids in the background Ymel ΔA are described in Table 1. Strain m524 Mu$^R\lambda^R$($\lambda cI^{857}S$7) used as a source of λ morphogenic proteins was obtained from H. Kondoh.

Bacteriophage strains used are all $tonB^-trpA^+$ transducing derivatives of the λ-ϕ80 hybrid phage f2 and have been described previously (Postle & Reznikoff, 1978).

Selection of $tonB$ mutants with ϕ80vir and colicins V + B has been described previously (Gottesman & Beckwith, 1969).

(c) Transformations and ligations

The chosen recipient bacterial strain was made competent (Mandel & Higa, 1970) and transformed with 0·2 μg of ColE1 Neo^r $tonB$ plasmid DNA prepared according to the method of Blair et al. (1972). Transformed colonies were selected on agar plates after 1·5 h outgrowth. Transformants were screened for the presence of the correct plasmids by agarose gel electrophoresis as previously described; ligation reactions have also been described previously (Postle & Reznikoff, 1978; Barnes, 1977; Sgaramella, 1972). Colicin E1 used in some selections was kindly supplied by L. V. Wray.

(d) Minicell analysis

Minicells were prepared essentially according to the methods of Roozen et al. (1971) and Frazer & Curtiss (1975). Minicell-producing strain χ984 was grown to early stationary phase in growth medium at 37°C. (All further manipulations were at room temperature unless otherwise noted.) The majority of bacterial cells were removed by a low speed spin (2000 revs/min, 5 min in a Sorvall GS3 rotor). The supernatant which consisted mostly of minicells was then centrifuged at 9000 revs/min for 15 min in a Sorvall GS3 rotor. The cell pellets were resuspended in 2 ml each of growth medium, vortexed vigorously for 2 min, and loaded onto 24-ml sucrose step gradients (8 ml 5% sucrose, 8 ml 12·5% sucrose, and 8 ml 20% sucrose, all in low phosphate medium without supplements). After centrifuging 10 min at 5000 revs/min in a Sorvall HB4 swinging bucket rotor, the upper one-half of the minicell band was removed and diluted with an equal volume of growth medium. This dilution was carefully done in a dropwise fashion to avoid drastic changes which could lyse the minicells. The minicells were then centrifuged 15 min at 15,000 revs/min in a Sorvall SS34 rotor, resuspended in 10 ml growth medium and the centrifugation was repeated. The pellet was resuspended vigorously (2 min of vortexing) in growth medium and layered onto a 12-ml sucrose step gradient (4 ml each of 5%, 12·5%, 20% sucrose in low phosphate medium). The gradients were centrifuged 10 min at

5000 revs/min in a Sorvall HB4 rotor. The top half of the minicell band was collected and carefully diluted drop by drop with 5 ml of growth medium. The optical densities at a 550 nm wavelength of these cell suspensions ranged from 0·24 to 0·60 on a Bausch and Lomb Spectronic 20 instrument.

The minicells were centrifuged 15 min at 15,000 revs/min in the Sorvall SS34 rotor and washed twice with labelling medium. The final pellets were resuspended in 1 to 2 ml of labelling medium. (At this point an effort was made to resuspend the pellets in volumes such that the concentrations as judged by optical density were all equal.) 1 ml of the minicells in labelling medium was incubated at 37°C. After 60 to 90 min the culture was made 2 mg/ml in ampicillin and incubated for a further 90 min. At this point 10 μCi [^{35}S]methionine (Amersham-Searle) were added. After 30 min the minicells were pelleted (20 min, 30,000 revs/min in a Beckman Ti59 rotor) and frozen. For sodium dodecyl sulfate /polyacrylamide gel analysis the pellets were resuspended in 50 to 100 μl sample buffer (Laemmli, 1970) and heated for 2 min at 90°C. 20 μl of sample were loaded per slot. 12·5% sodium dodecyl sulfate/polyacrylamide slab gels (20 cm $\times$ 20 cm $\times$ 1·3 mm) were run according to the method of Laemmli (1970). The gels were dried and autoradiographed at room temperature using Kodak X-Omat R film for 3 to 4 days.

(e) *Preparation of* [^{35}S]*methionine-labelled* λ *structural proteins*

A 10-ml culture of m524 Mu$^{R}\lambda^{R}$ lysogenic for λcI^{857}S7 was grown in minimal M9 medium to early logarithmic phase at 37°C and then shifted to 42°C for 10 min. After the 42°C heat pulse, 100 μCi [^{35}S]methionine (Amersham-Searle) was added and the culture was incubated 2·5 h at 37°C. The cells were pelleted at 4°C, lysed in chloroform, the cellular debris pelleted and the supernatant added to non-radioactive carrier phage of the same type. The lysate was banded on a 10-ml CsCl equilibrium density gradient. The visible band of phage was removed from the gradient, dialyzed into 10 mM-Tris (pH 7·9), 20 mM-MSO$_4$, 5 mM-CaCl$_2$ and a portion counted for radioactivity. $\sim$ 10,000 cts/min per 10 μl of phage lysate were obtained. Whole phage were mixed with non-radioactive bacteria and sodium dodecyl sulfate sample buffer (Laemmli, 1970) and heated 3 min at 95°C. 20 μl of this mixture were loaded per gel slot. Molecular weights and relative amounts of the λ structural proteins were taken from Murialdo & Siminovitch (1972) and Casjens & Hendrix (1974).

(f) *Inhibition of* BglII *cleavage by RNA polymerase*

1 μl *E. coli* RNA polymerase at 4 mg/ml (generously provided by L. Maquat and S. Rothstein) was incubated with 2 μg of pRZ546 DNA for 30 min at 37°C in PBBX (Jones & Reznikoff, 1977) plus 100 mM-KCl. After incubation, the mixture was diluted 1 : 2 with 10 mM-Tris (pH 7·9), 6·6 mM-MgCl$_2$. All 4 ribonucleoside triphosphates were added to 400 μM each or else omitted and the mixtures incubated at 37°C for 15 min. Restriction enzymes were added and the mixtures incubated further at 37°C for 60 min. The DNAs were extracted with phenol once, extracted twice with ether, and analyzed on 3·5% polyacrylamide gels.

As a control pRZ546 was treated as above but without added polymerase or ribonucleoside triphosphates.

(g) *DNA restriction digests and polyacrylamide gels*

The DNA digestions and their visualization on 3·5% polyacrylamide slab gels have already been described (Postle & Reznikoff, 1978).

3. Results

(a) *Minicell analysis of proteins encoded by* tonB$^+$ *hybrid plasmids*

We have previously described a series of Neo^r (neomycin-resistant) *tonB*$^+$ hybrid plasmids (Postle & Reznikoff, 1978) which carry different extents of the *tonB* region between *attϕ*80 and the tryptophan operon. The properties of these *tonB*$^+$ plasmids
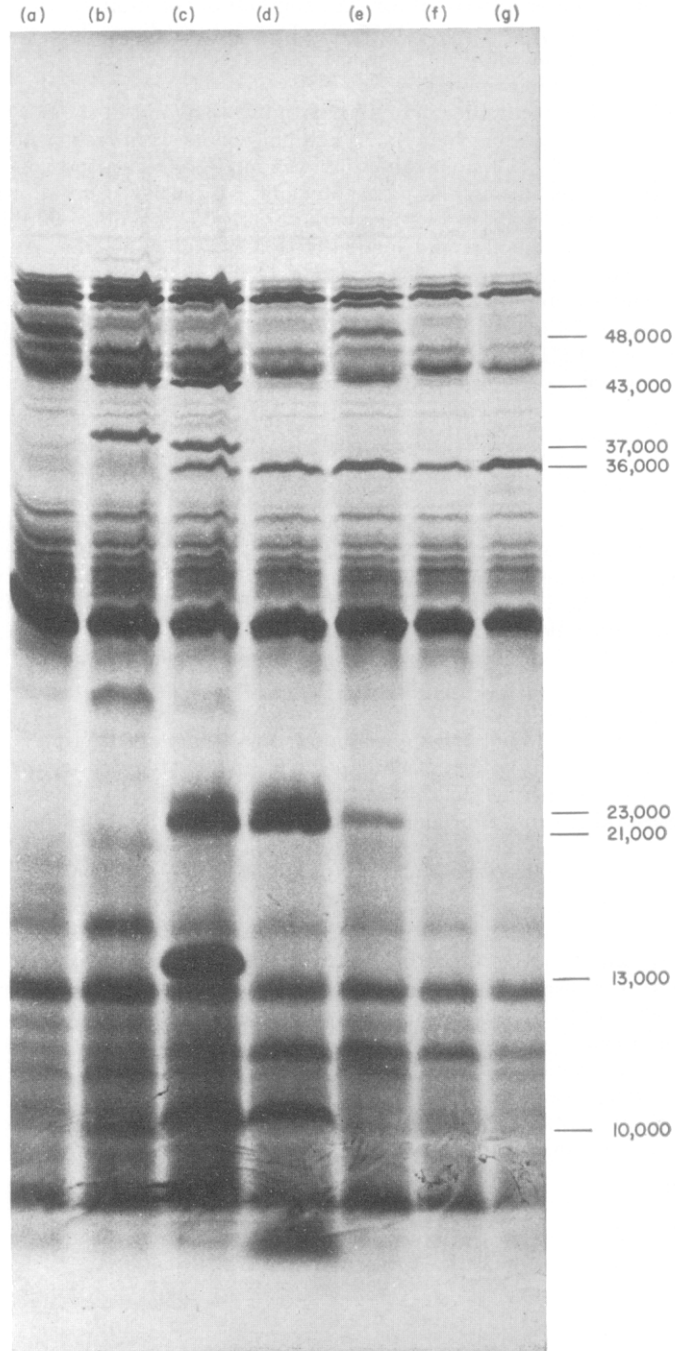
FIG. 1. [^{35}S]methionine-labelled proteins produced by $tonB^+$ plasmids in minicells.
Labels on the right side of photograph indicate novel proteins produced by the 10,000 bp
HindIII fragment as well as the 48,000 M_r protein from the inverted repeat of transposable
element Tn5. All plasmids are in strain χ984. (a) pRZ112: the parental plasmid; (b) pRZ429:
the $trpA^+$ plasmid; (c) pRZ510; the cloned 10,000 bp HindIII fragment; (d) pRZ526: the cloned

TABLE 1

Hybrid plasmids used in this study

Plasmid	Fragment cloned	Cloning site in pRZ112†	Plasmid phenotype
pRZ540	1730 bp HindII	HpaI	tonB$^+$
pRZ546	1730 bp HindII‡	HpaI	tonB$^+$
pRZ531	2760 bp PvuII	EcoRI	tonB$^+$
pRZ526	4900 bp HpaI	HpaI	tonB$^+$
pRZ510	10,000 bp HindIII	HindII	tonB$^+$
pRZ551	4900 bp HpaI	HpaI	tonB$^-$ (f2#23)
pRZ552	4900 bp HpaI	HpaI	tonB$^-$ (f2#31)
pRZ553	4900 bp HpaI	HpaI	tonB$^-$ (f2#22)
pRZ554	4900 bp HpaI	HpaI	tonB$^-$ (f2#29)
pRZ555	4900 bp HpaI	HpaI	tonB$^-$ (f2#28)
pRZ556	4900 bp HpaI	HpaI	tonB$^-$ (f2#24)
pRZ557	4900 bp HpaI	HpaI	tonB$^-$ (f2#17)
pRZ429	2990 bp HindIII	HindIII	trpA$^+$

The trpA$^+$ and tonB$^+$ plasmids were constructed previously (Postle & Reznikoff, 1978). The tonB$^-$ plasmids were constructed by homogenotization, as described in Results, with the f2 phages indicated in the 4th column. The structure of the vehicle pRZ112 used in these constructions is diagrammed in Fig. 8.

† pRZ112 is a Neo^r ColE1 derivative constructed by R. Jorgensen.

‡ Opposite orientation to pRZ540.

are summarized in Table 1. In order to examine the proteins encoded by the tonB region these plasmids were transformed into the minicell-producing strain χ984. [^{35}S]methionine-labelled proteins from minicell extracts were separated on 12·5% sodium dodecyl sulfate/polyacrylamide slab gels, the gels were dried, and autoradiographed. An autoradiogram of proteins produced by the tonB$^+$ plasmids is shown in Figure 1.

Plasmids of pRZ540 and pRZ546 (Fig. 1(f) and (g)) each carry the smallest (1730 bp†) cloned tonB$^+$ fragment and differ only in the orientation of the cloned fragment. Comparison of the proteins encoded by the parental plasmid, pRZ112, with the proteins encoded by either pRZ540 or pRZ546 reveals the presence of only one unique protein with an apparent molecular weight of 36,000. This protein is seen in the two minicell extracts regardless of the orientation of the 1730 bp HindII fragment which apparently encodes it. The unique protein is also encoded by the other tonB$^+$ hybrid plasmids (Fig. 1(c), (d) and (e)), all of which carry greater extents of the tonB region. These observations suggest that the 36,000 M_r protein is not an artifact of the cloning process in which fusion proteins are created by transcription across the boundaries of cloned restriction fragments. Its presence as the sole protein encoded by the 1730 bp tonB$^+$ HindII fragment suggests that it is the tonB gene product.

† Abbreviation used: bp, base-pair.

4900 bp HpaI fragment; (e) pRZ531: the cloned 2760 bp PvuII fragment; (f) pRZ540: the cloned 1730 bp HindII fragment; and (g) pRZ546: the cloned 1730 bp HindII fragment in opposite orientation to pRZ540. Molecular weights of proteins were determined using [^{35}S]methionine-labelled λ head and tail proteins. The molecular weights below 32,000 are uncertain due to a small number of λ proteins in that region of the gel. 12·5% sodium dodecyl sulfate/polyacrylamide slab gels were run 5 h at 30 mA, dried and autoradiographed. The control lane of χ984 with no plasmid (not shown) was essentially blank.

(b) *Introduction of* tonB *mutations onto* tonB$^+$ *plasmid pRZ526*

In order to confirm that the 36K protein was the *tonB* gene product, a number of *tonB* mutations were genetically crossed on to *tonB*$^+$ plasmid pRZ526. Bacterial strain Ymel$\varDelta$A was lysogenized with each of the following *tonB*$^-$*trpA*$^+$ transducing phages by selecting for *trpA*$^+$ colonies at 30°C: f2#22, f2#23, f2#24, f2#28, f2#29, f2#31, and f2#17. Each of these lysogens was screened for *tonB* phenotype (ϕ80vir resistance) and subsequently transformed with pRZ526, the Neo^r hybrid plasmid which carries the 4900 bp *tonB*$^+$ *Hpa*I fragment, by selecting for Chr^r (chromium resistance) and Neo^r. *tonB*$^+$ cells are resistant to 100 μM-chromium on agar plates. The selections for Chr^r, Neo^r, and *trpA*$^+$ have been previously described (Postle & Reznikoff, 1978; Jones & Reznikoff, 1976).

tonB$^+$ transformants were grown to saturation in nutrient broth. Neo^r colonies simultaneously resistant to ϕ80vir and colicins V+B were selected on agar plates. *tonB* Neo^r colonies were recovered with frequencies of 10^{-4} to 10^{-5} per cell, indicating that the homogenotization between plasmid, pRZ526, and the chromosome was efficient. Candidates were screened for the presence of plasmid DNA which was subsequently analyzed with restriction enzymes *Hind*II + III. The *tonB* mutations in the plasmids appeared to be the same as those in the *tonB*$^-$*trp*$^+$ transducing phages from which they were derived (Fig. 2). In the cases of f2#22, f2#23, f2#29 and f2#31, where the 1730 bp *Hind*II fragment 6 was missing from a *Hind*II + III restriction digest and replaced in each case by similar *Hind*II fragments 800 bp larger, the corresponding *Hind*II fragment was seen in each *tonB*$^-$ plasmid. In the cases of f2#17, f2#24 and f2#28, where the 1730 bp *Hind*II fragment 6 was essentially unaltered in size by the *tonB* mutation, a fragment the size of *Hind*II fragment 6 was also present in *Hind*II + III digests of the *tonB*$^-$ plasmid DNAs.

(c) *Minicell analysis of proteins encoded by* tonB$^-$ *plasmids*

The plasmids carrying *tonB* mutations were transformed into χ984 and the [^{35}S]methionine-labelled minicell proteins examined on sodium dodocyl su'fate/gels. The autoradiogram in Figure 3 shows that none of the *tonB*$^-$ plasmids cod·s for the 36,000 M_r protein produced by the *tonB*$^+$ plasmids described above. The other three unique proteins encoded by pRZ526 (compared to pRZ112) are unaltered by the *tonB* mutations. Six of the seven mutant plasmids (pRZ551 to pRZ556) code for new polypeptides smaller than 36,000, which we assume in each case to be a prematurely terminated *tonB*$^-$ protein. Of these, the four polypeptides produced by the insertion

---

FIG. 2. *Hind*II$^+$III restriction patterns of *tonB*$^-$ plasmids compared to the *tonB*$^-$ phage DNAs from which they were derived.

(a) *tonB* point mutations or small deletions: (1) ϕX174 molecular weight standards, (2) pRZ112, the parental plasmid, (3) pRZ557, (4) f2#17, (5) pRZ556, (6) f2#24, (7) pRZ555, (8) f2#28, (9) 13 4/6 5, the f2 phage carrying the entire wild-type *tonB-trp* region. On the right-hand side are indicated the 3 *Hind*II fragments 6, 7 and 8 (1730 bp, 1600 bp, 1570 bp) that comprise the *Hpa*I fragment cloned in pRZ526. They were numbered according to their order in the *Hind*II + III digest of 13 4/6 5 (Jones & Reznikoff, 1977). The size of fragment 6 (*tonB*$^+$) appears to be unaltered in the *tonB*$^-$ phage and plasmid digests.

(b) *tonB* insertion mutations: (1) pRZ112, (2) pRZ553, (3) f2#22, (4) pRZ551, (5) f2#23, (6) pRZ554, (7) f2#29, (8) pRZ552, (9) f2#21, (10) 13 4/6 5. These phages and plasmids lack *Hind*II fragment 6 and have instead a new ~2500 bp fragment which is not present in the 13 4/6 5 digest.

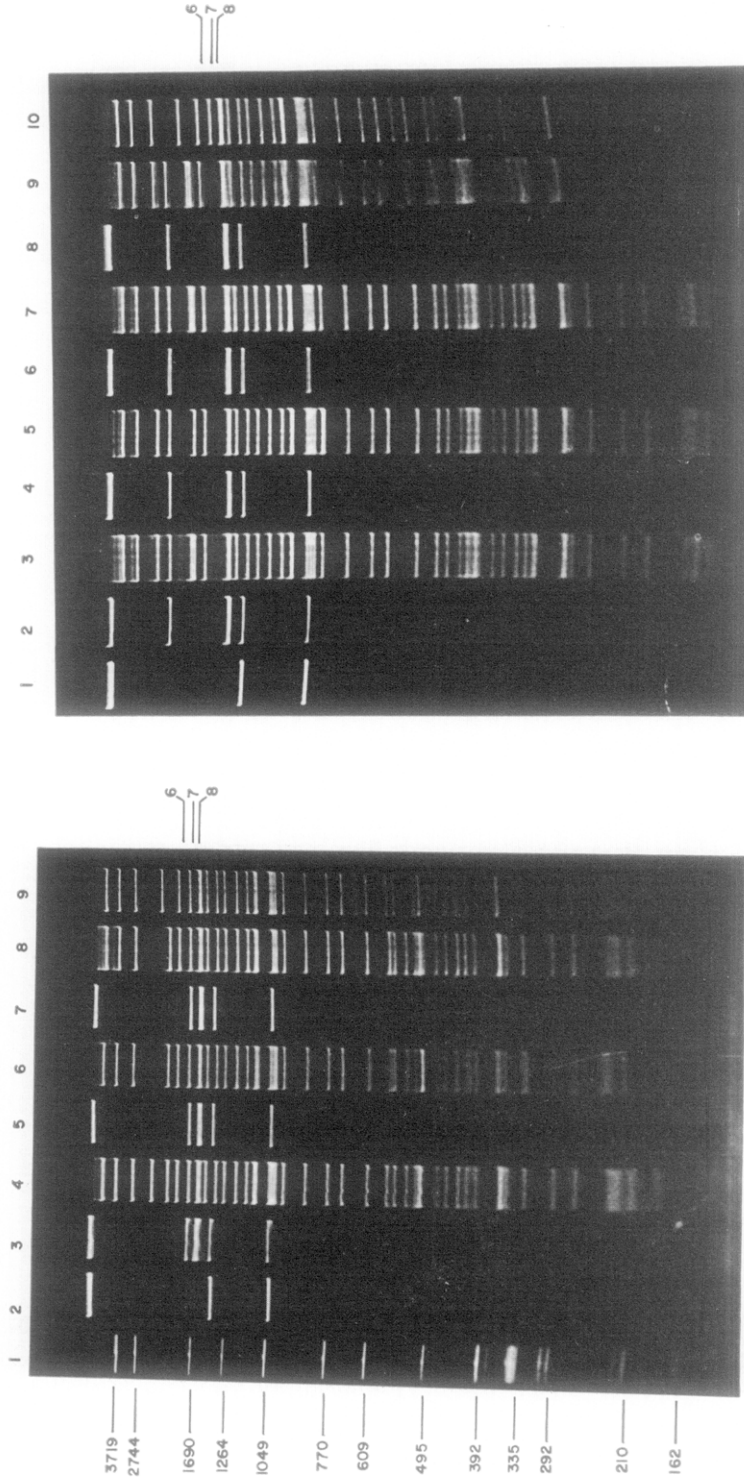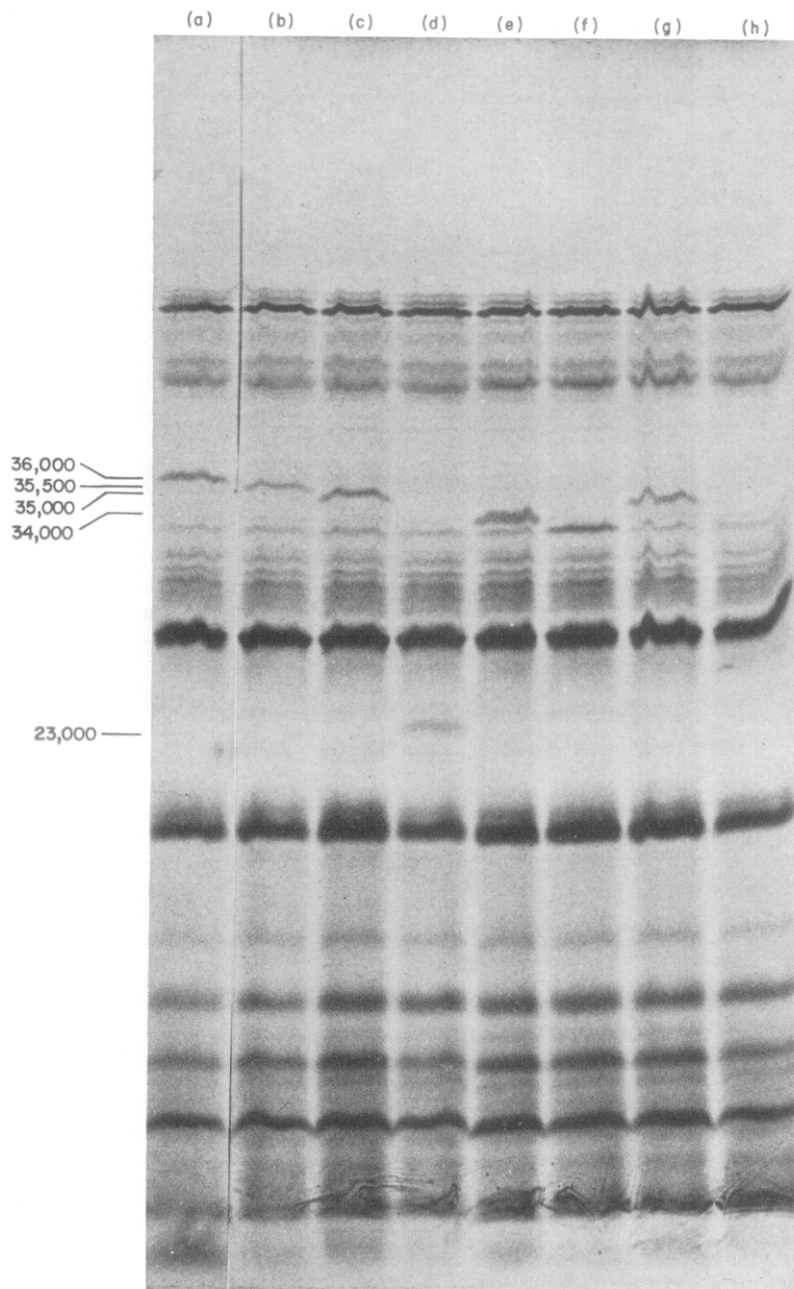Numbers on the far left indicate lengths of double-stranded ϕX174 standards in base-pairs.

FIG. 2

FIG. 3. [^{35}S]methionine-labelled proteins produced by $tonB^-$ plasmids in minicells.

χ984 minicells containing: (a) pRZ526, the plasmid onto which the $tonB^-$ mutations were crossed, (b) pRZ551, (c) pRZ552, (d) pRZ554, (e) pRZ553, (f) pRZ555, (g) pRZ556, (h) pRZ557. On the left are labelled protein bands corresponding to the prematurely terminated $tonB$ polypeptides produced by the insertion mutants (pRZ551, pRZ552, pRZ554, pRZ553). This experiment was run on the same gel shown in Fig. 1.

mutations (pRZ551-35·5K; pRZ552-35K; pRZ553-34K; pRZ554-23K) are of special interest. The physical locations of the insertions can be determined and compared with the sizes of the altered polypeptides. The fact that the 36,000 M_r protein is the only one encoded by the *tonB*$^+$ 1730 bp *Hind*II fragment and that this protein is missing in minicell extracts of all seven *tonB* mutants analyzed strongly suggests that the 36,000 M_r protein is in fact the *tonB* gene product.

(d) *Mapping of the* tonB$^-$ *insertion mutations*

The sizes (~800 bp) of the insertions in the *tonB*$^-$ plasmids pRZ551 to pRZ554 suggested that these *tonB* mutations might be due to IS1 insertions. Restriction enzyme analyses showed that the four insertions do share a common structure that is consistent with known restriction sites in IS1 (Schmidt *et al.*, 1976; Grindley, 1977; Ohtsubo & Ohtsubo, 1978): *Bal*I (1 site); *Pst*I (1 site); *Hae*II (2 sites); *Hind*III, *Bgl*II, *Hind*II and *Hpa*I (no sites). We conclude that these mutations are in all probability IS1 insertions.

The presence of unique *Bal*I and *Pst*I sites in the insertions along with a unique *Bgl*II site in the 1730 bp *tonB*$^+$ *Hind*II fragment itself were exploited to map the insertions. Since the sequence of IS1 is known and the *Bal*I and *Pst*I sites located precisely within it (Ohtsubo & Ohtsubo, 1978), we can map the insertions fairly accurately by restriction enzyme analysis (Fig. 4). Comparisons among the three insertions in the same orientation (pRZ552, pRZ553 and pRZ554) are more reliable than to the one insertion in the opposite orientation (pRZ551) since the three can be directly compared within the same slab gel. The *Bgl*II *Pst*I (or *Bgl*II *Bal*I) band of interest in each mutant was identified as the only unique band present in the double digest and missing in the single digest of that mutant as well as the double digests of the other mutants. Estimates for the *Bgl*II–*Pst*I distances and the *Bgl*II–*Bal*I distances are shown in Table 2.

The precise locations of the insertions together with the molecular weights of the polypeptides which they apparently terminate suggest a direction for *tonB* transcription. The boundaries of the insertions in plasmids pRZ551, pRZ552, pRZ553 and pRZ554 are 805 bp, 770 bp, 750 bp, 455 bp, respectively, from the *Bgl*II site in the 1730 bp fragment (Fig. 5); these insertions result in *tonB*$^-$ polypeptides of approximately 35.500, 35,000, 34,000 and 23,000 M_r, respectively (Fig. 3). The plasmid carrying the insertion farthest from the *Bgl*II site encodes the largest *tonB*$^-$ polypeptide. The plasmid carrying the insertion closest to the *Bgl*II site encodes the smallest *tonB* polypeptide. Within the series of four insertions, the data are most consistent with a direction of transcription which is clockwise on the circular *E. coli* genetic map (from *att*φ80 toward the *trp* operon).

It is worth noting that we do not see any proteins coded by the IS1 sequence in our minicell experiments (Fig. 3); however, small amounts of a protein could be overlooked, or concealed in congested regions of the gel.

(e) *Location of a potential promotor region in the 1730 bp* tonB$^+$ Hind*II fragment*

Indirect evidence concerning the direction of *tonB* transcription suggests the location of a potential promotor region for the *tonB* gene. The *tonB* IS1 insertion in pRZ551 (Fig. 5) encodes a *tonB*$^-$ polypeptide of 35.500 M_r which probably termi-
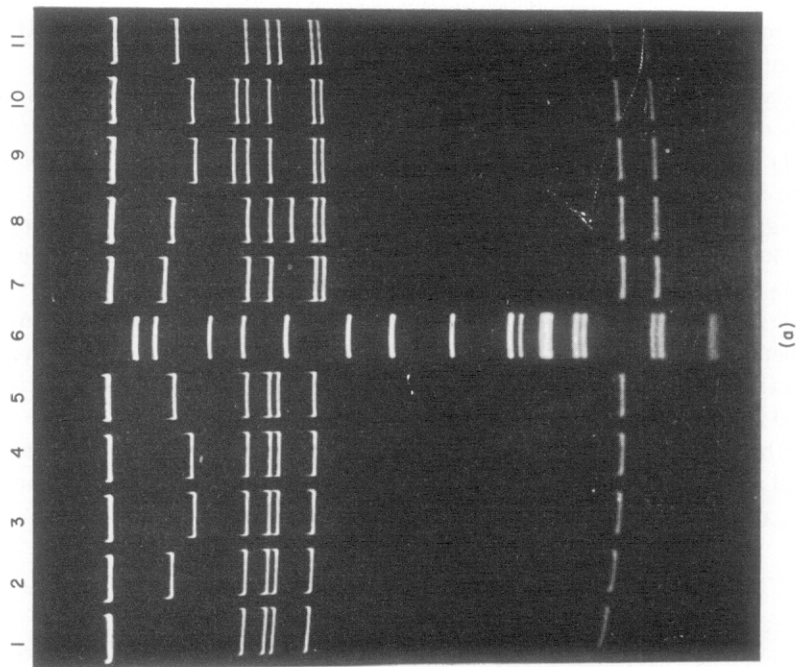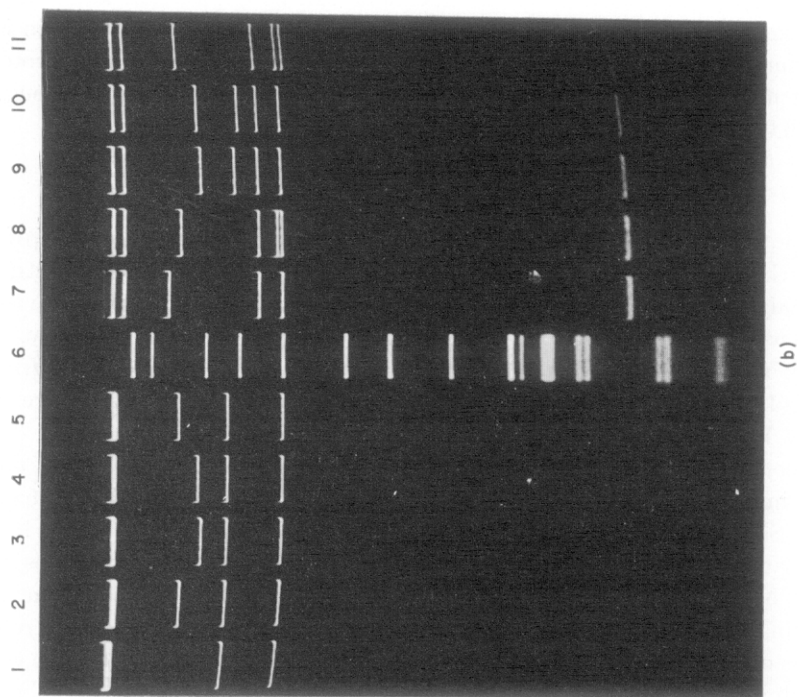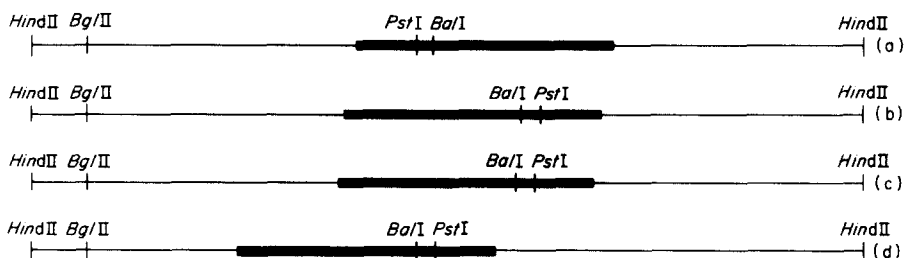
FIG. 4

FIG. 5. Map positions of IS1 insertions in the 1730 bp *tonB*$^+$ *Hind*II fragment.

(a) pRZ551, (b) pRZ552, (c) pRZ553, (d) pRZ554. Unique *Pst*I and *Bal*I sites common to IS1 and each insertion mutant are diagrammed. The insertion in pRZ551 is apparently in an orientation opposite to that of the other insertions. pRZ551, pRZ552, pRZ552 and pRZ554 code for prematurely terminated *tonB*$^-$ polypeptides of 35,500, 35,000, 34,000 and 23,000 M_r, respectively.

nates very close to the IS1 boundary 805 bp downstream from the *Bgl*II site (Table 2). Very little translational readthrough into IS1 would be expected due to the existence of several nonsense codons in all reading frames of the insertion sequence (Ohtsubo & Ohtsubo, 1978). Assuming an average amino acid molecular weight of 130, the 35,500 *tonB*$^-$ polypeptide would be encoded by approximately 820 bp of DNA. This calculation places a potential promotor region for the *tonB* gene at or near the *Bgl*II site (Fig. 5). In order to test this possibility, the ability of *E. coli* RNA polymerase to protect the *Bgl*II site from cleavage was investigated.

TABLE 2

Restriction enzyme mapping of IS1 mutations

tonB$^-$ plasmid	*Bgl*II–*Pst*I distance (bp)	*Bgl*II–*Bal*I distance (bp)	*Bgl*II to boundary of IS1 (bp)
pRZ551	990	1035	805
pRZ552	1360	1300	770
pRZ553	1340	1285	750
pRZ554	1045	985	455

Estimated distances from the unique *Bgl*II site in the 1730 bp *Hind*II fragment to the unique *Pst*I and *Bal*I sites in the 4 *tonB*$^-$ insertion mutants are shown. The distance from the *Bgl*II site to the boundary of the IS1 element has been calculated from the known positions of the *Pst*I and *Bal*I sites within IS1 (Ohtsubo & Ohtsubo, 1978).

FIG. 4. Mapping of the *tonB* IS1 insertions.

(a) *Bgl*II-*Pst*I analysis. (1) pRZ526, the *tonB*$^+$ plasmid onto which the insertion mutants were crossed, (2) pRZ551, (3) pRZ552, (4) pRZ553, (5) pRZ554, (6) ϕX174 molecular weight standards, (7) pRZ526, (8) pRZ551, (9) pRZ552, (10) pRZ553, (11) pRZ554. Lanes 1 to 5 are *Pst*I digests; lanes 7 to 11 are *Pst*I-*Bgl*II double digests.

(b) *Bgl*II-*Bal*I analysis. (1) pRZ526, (2) pRZ551, (3) pRZ552, (4) pRZ553, (5) pRZ554, (6) ϕX174 standards, (7) pRZ526, (8) pRZ551, (9) pRZ552, (10) pRZ553, (11) pRZ554. Lanes 1 to 5 are *Bal*I digests; lanes 7 to 11 are *Bal*I-*Bgl*II double digests.

The *Pst*I and *Bal*I digests indicate that only one *Pst*I and one *Bal*I site exist in each insertion sequence.
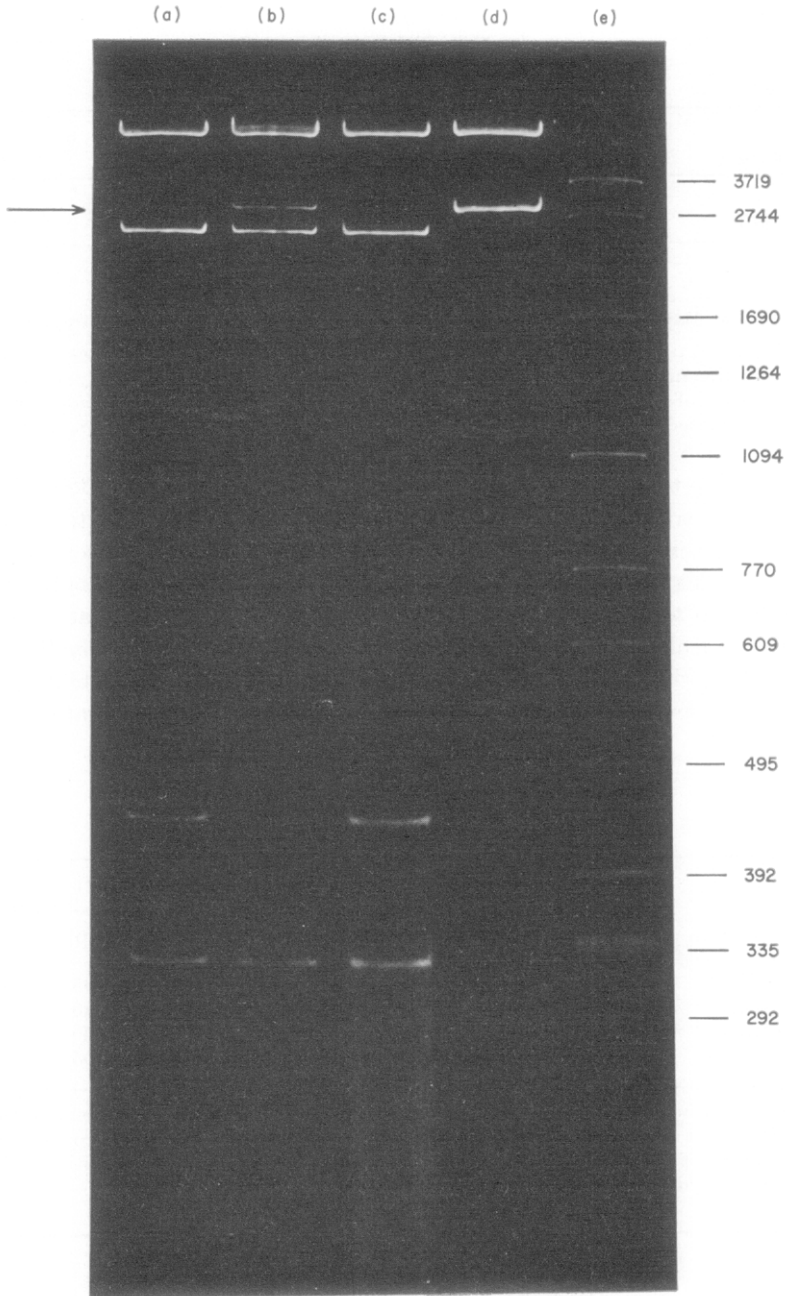
FIG. 6. Inhibition of *Bgl*II cleavage by *E. coli* RNA polymerase.
(a) pRZ546, (b) pRZ546 + *E. coli* RNA polymerase, (c) pRZ546 + *E. coli* RNA polymerase + ribonucleoside triphosphates, (d) pRZ546, (e) φX174 molecular weight standards. (a), (b) and (c) are *Hind*III-*Eco*RI-*Bgl*II triple digests. (d) A *Hind*III-*Eco*RI double digest. Labels on the right indicate molecular weights of φX174 standards in bp. The arrow on the left indicates the position of the 3000 bp fusion fragment resulting from RNA polymerase protection of the *Bgl*II site in or near *tonB*.

The plasmid pRZ546 (carrying the 1730 bp *tonB*+ *Hind*II fragment) is cut into two fragments by a *Hind*III *Eco*RI double digest (Fig. 6(d)). The larger fragment consists of most of the cloning vector (Fig. 8) while the smaller fragment (~3000 bp) encompasses the cloned 1730 bp *tonB*+ fragment. A *Hind*III *Eco*RI *Bgl*II triple digest cleaves ~335 bp off the end of the larger fragment and divides the smaller fragment into ~440 bp and ~2540 bp. (Fig. 6(a)). Incubation of the plasmid DNA with RNA polymerase before addition of restriction enzymes results in a fusion fragment (Fig. 6 (b)) which is the same size as the smaller fragment from the *Hind*III *Eco*RI double digest (Fig. 6(d)). Also in Figure 6(b), the relative intensity of the 440 bp and 2540 bp fragments are detectably diminished. Addition of ribonucleoside triphosphates (Fig. 6(c)) after the preincubation with RNA polymerase but before the addition of restriction enzymes results in loss of the fusion fragment, presumably due to initiation and run-off of the RNA polymerase.

These results suggest that a promotor region may be located at or near the *Bgl*II site. We have no data that directly implicate this region as the promotor for the *tonB*, but it is in the correct location for such a regulatory element.

(f) *Minicell analysis of other proteins from the* tonB *region*

Plasmid pRZ510 carries the 10,000 bp *tonB*+ *Hind*III fragment which defines most of the *tonB* region along with 3700 bp of ϕ80 DNA sequences. Comparison of the proteins encoded by pRZ112 and pRZ510 indicates that the *Hind*III fragment encodes seven unique polypeptides of molecular weights 43,000, 37,000, 36,000, 23,000, 21,000, 13,000 and 10,000 (Fig. 1). Because this fragment carries ϕ80 sequences as well as part of the *tonB* region, it is not possible to assign any proteins to the *tonB* region alone. Since the 10,000 bp *Hind*III fragment has been cloned only in one orientation, it is not possible to determine which, if any, of the proteins are fusion proteins resulting from transcription and translation across the boundary of the cloned fragment. However, this fragment provides a set of proteins against which the proteins encoded by smaller *tonB*+ fragments may be compared. The cloned 4900 bp *Hpa*I fragment encodes four unique proteins (M_r: 36,000, 23,000, 21,000 and 10,000) compared to those encoded by the plasmid pRZ112. Since these proteins are also encoded by the 10,000 bp *Hind*III fragment they are apparently not fusion proteins arising from the plasmid construction. Of these four proteins, only two (36,000 and 21,000), are encoded by the cloned 2760 bp *Pvu*II fragment (pRZ531). And of those two proteins, only one, the *tonB* gene product, is encoded by the cloned 1730 bp *Hind*II fragment.

The cloned *trpA*+ *Hind*III fragment carries the rest of the *tonB* region not covered by the 10,000 bp *Hind*III fragment. Five new proteins are encoded by the *trpA*+ *Hind*III fragment in pRZ429 when compared to the parental plasmid pRZ112. These have molecular weights of 44,000, 38,000, 22,000, 18,000 and 15,000. For unknown reasons, the *trpA* protein (28,700 M_r) was not seen. The above observations are summarized in Figure 7.

These data do not enable us to define the origins of any of these proteins with the exception of the 36,000 *tonB* gene product. They do, however, suggest that there is greater genetic complexity in the *tonB* region than has been previously described (Bachmann *et al.*, 1976).
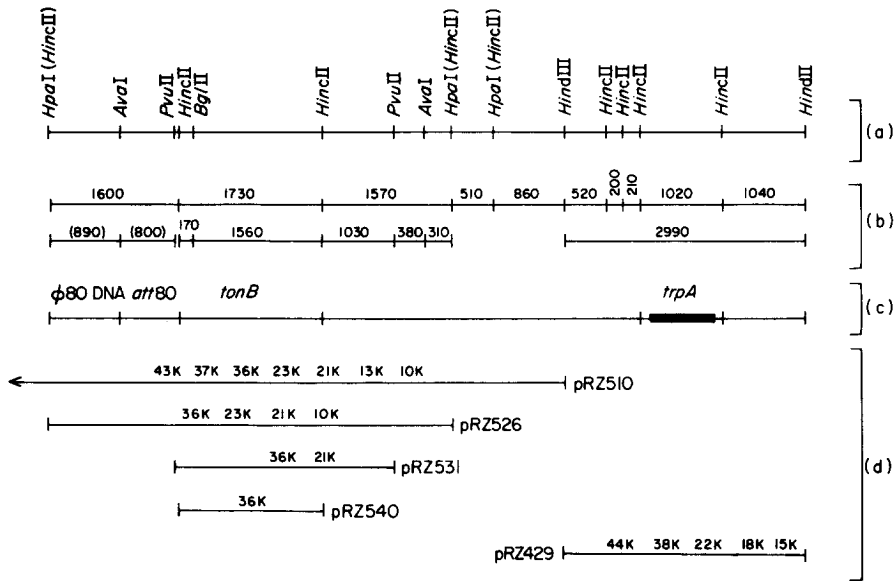
FIG. 7. Correlation of the physical, genetic and translational maps of the *tonB* region.

(a) Partial restriction map of the *attφ80-tonB-trp* region. (b) Distances in base-pairs between sites on the restriction map. (c) Correlation of the genetic map with the physical map. Genetic markers have been located within the fragments they are positioned above. The precise locations of *trpA* and *tonB* are known. (d) Proteins encoded by various cloned *tonB*⁺ restriction fragments. Regions of DNA cloned to generate the indicated hybrid plasmids are diagrammed together with the novel proteins produced from each region. In the cases of the cloned *Hind*III fragments (pRZ510, pRZ429), some of the proteins detected in minicells may be fusion proteins which cross the boundary of the cloned fragment.

43K, 43,000 M_r and so on.

(g) *Identification of a protein from the inverted repeat sequences of Tn5*

A protein with a molecular weight of 48,000 was observed on sodium dodecyl sulfate/gels in the minicell analysis of several of the *tonB*⁺ hybrid plasmids (Fig. 1). This protein is present only in minicell extracts of strains containing plasmids (pRZ112, pRZ531; Fig. 1) where the single Tn5 inverted repeat present in pRZ112 (Fig. 8) is intact. In those cases where the inverted repeat has been interrupted by cloning into the *Hpa*I or *Hind*III sites within the inverted repeat (Jorgensen, R., personal communication), the protein is not detected (pRZ429, pRZ510, pRZ540, pRZ546).

The simplest explanation for this observation is that the 48,000 M_r protein is encoded by the single inverted repeat of Tn5 that is present in pRZ112, since the *Hpa*I site and the *Hind*III site are 980 bp apart in the inverted repeat (Jorgensen, R., personal communication). Alternatively, the protein is the product of translational readthrough into ColE1 or the central region of Tn5. In either case, sequences within the inverted repeat appear to be essential for the expression of the 48,000 protein.
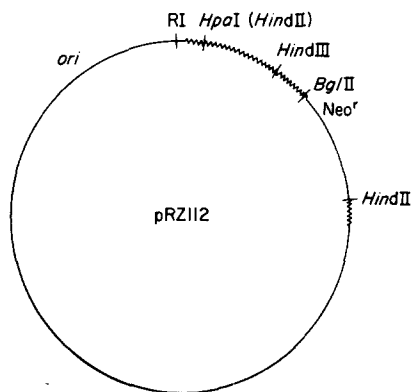
FIG. 8. Partial restriction map of pRZ112.
pRZ112 is a deletion variant of ColE1::Tn5 constructed by R. Jorgensen. In this diagram, the jagged lines represent DNA from the inverted repeat sequences of Tn5. A *Hind*II deletion of the original ColE1::Tn5 plasmid has resulted in almost complete loss of one inverted repeat.

4. Conclusions

We had previously localized the *tonB* gene to a 1730 bp *Hind*II fragment (Postle & Reznikoff, 1978). Here we examine the proteins encoded by *tonB*$^+$ hybrid plasmids described in that study. Our findings suggest that the *tonB* gene product is a protein with an apparent molecular weight of 36,000 on sodium dodecyl sulfate/polyacrylamide gels. (1) Plasmids carrying the 1730 bp *Hind*II fragment encode only one protein, the 36,000 protein, not encoded by the parental plasmid. (2) This protein is produced regardless of the orientation of the 1730 bp fragment in the parental plasmid. (3) This protein is also encoded by *tonB*$^+$ plasmids carrying more extensive portions of the *tonB* region. Synthesis of the 36,000 protein is thus apparently independent of the DNA sequences which flank the 1730 bp fragment. These results suggest that the 36,000 protein is encoded by a sequence that is entirely within this fragment and that it is the *tonB* gene product.

In parallel experiments with plasmids carrying *tonB* mutations, the 36,000 M_r protein is the only protein absent from the set of proteins encoded by each *tonB*$^-$ plasmid. Six of the seven *tonB*$^-$ plasmids encode new polypeptides which are smaller than 36,000 and which we presume to be prematurely terminated derivatives of the *tonB* protein. These results confirm our identification of the 36,000 M_r protein as the *tonB* gene product.

Four of the *tonB* mutations appear to be IS1 insertions into the 1730 bp *Hind*II fragment, which have been precisely mapped relative to the *Bgl*II site in that fragment. There is a strong correlation between the distance of the insertion from the *Bgl*II site and the molecular weight of the *tonB*$^-$ polypeptide encoded by the mutant plasmid; the greater the distance from the insertion to the *Bgl*II site, the larger the *tonB*$^-$ polypeptide seen on polyacrylamide gels. These data suggest that the direction of *tonB* transcription is clockwise (from *attφ*80 toward *trp*) on the circular *E. coli* genetic map. In order to postulate a direction of transcription for *tonB* which is counterclockwise on the *E. coli* genetic map, there would have to be extensive translational readthrough into the IS1 insertions ($>$600 bp) in pRZ551, pRZ552 and pRZ553 in order to generate proteins of the observed molecular weights. Although the minicell-producing strain χ984 carries an amber nonsense suppressor (*supE*), IS1

contains multiple non-amber termination codons in all three reading frames close to either end of the element. Thus it is unlikely that such readthrough could occur.

Three of the four IS1 insertions are clustered within a 60 bp region of *tonB* DNA. This observation, in addition to the fact that the only insertion mutations observed (four out of the 12 *tonB* mutations described previously, Postle & Reznikoff, 1978) were of the IS1 type, may indicate that there are IS1 insertion "hot spots" in *tonB* similar to those found in other regions of the *E. coli* genome (for a review, see Bukhari *et al.*, 1977).

We observed that RNA polymerase inhibits cleavage at the *Bgl*II site in the 1730 bp *tonB*$^+$ *Hind*II fragment. This result suggests that there is a potential promoter region on the 1730 bp *tonB*$^+$ *Hind*II fragment at or near the *Bgl*II site 170 bp from the end of the fragment. This is consistent with the inferred direction of *tonB* transcription. We are investigating the possibility that this *Bgl*II site is within the *tonB* gene promoter.

The minicell analysis of *tonB*$^+$ and *trpA*$^+$ hybrid plasmids indicates the existence of several proteins encoded by different cloned restriction fragments carrying segments of the *tonB* region. Of these proteins we have been able to localize the coding sequences for only the 36,000 M_r *tonB* gene product. However our experiments imply that the *tonB* region possesses a greater genetic complexity than has been previously described. This is consistent with our data regarding fusions of β-galactosidase to proteins within the *tonB* region (Postle & Reznikoff, 1978). Further analysis of cloned fragments from the *tonB* region will be necessary before it is possible to localize the other proteins.

Two interesting observations unrelated to *tonB* have also emerged from these studies. First, the minicell analysis of proteins encoded by the *tonB*$^-$ plasmids carrying IS1 sequences has revealed no obvious protein encoded by that insertion. Such a protein would be $\leqslant 33,000$ molecular weight (assuming an average amino acid molecular weight of 130) since the IS1 element is 768 bp long (Ohtsubo & Ohtsubo, 1978). IS1 has promoter-like sequences near each end where RNA polymerase might bind (Ohtsubo & Ohtsubo, 1978), but it is doubtful that a functional protein would be expressed by IS1 since there are numerous nonsense codons in all three reading frames.

Secondly, the minicell analysis of proteins encoded by Neo^r *tonB*$^+$ plasmids has provided suggestive evidence that at least one inverted repeat sequence from transposable element Tn5 is involved in the synthesis of a 48,000 M_r protein. The inverted repeat sequences of Tn5 are each approximately 1450 bp (Berg *et al.*, 1975) with a calculated protein coding capacity of 62,800 and thus could encode a protein of 48,000. The simplest and most exciting interpretation of the results is that the 48,000 M_r protein is completely encoded by the inverted repeat sequence. The only protein presently known to be encoded by Tn5 is the $\sim$23,000 M_r neomycin phosphotransferase II which specifies neomycin resistance (D. I. Smith, Ph. D. Thesis 1978; D. I. Smith, T. White, J. Davies, unpublished results).

Evidence from studies on the transposable Tn4 (ampicillin resistance) element suggests that other proteins may be encoded by such elements. Sequences within the Tn4 element are known to be necessary for transposition; Amp^r transposition-defective deletions in the central region of Tn4 are complementable in *trans* suggesting the existence of a diffusable product necessary for transposition (Heffron *et al.*, 1977). Likewise, internal deletions of the central region and part of one inverted

repeat in Tn*5* are known to be transposition-defective, but it is not known whether such deletions are complementable (Davies *et al.*, 1977). We are currently investigating the origin of the 48,000 M_r protein and its role in Tn*5* transposition.

There are several qualifications which should be considered in analyzing our data. In two cases where restriction fragments were cloned in only one orientation (pRZ510 and pRZ429) it is possible that some of the proteins we observe may result from transcription and translation across the boundaries of the cloned fragments. In addition the molecular weights which we report are only apparent molecular weights, since any or all of the proteins we detect may reside in the cell membrane; in fact there is evidence that the *tonB* gene product might reside in the inner membrane (Wookey & Rosenberg, 1978). Finally, the use of [^{35}S]methionine to label the proteins will have prevented us from detecting those proteins from the *tonB* region which contain few or no methionine residues.

REFERENCES

Bachmann, B. J., Low, K. B. & Taylor, A. L. (1976). *Bacteriol. Rev.* **40**, 116–167.

Barnes, W. M. (1977). *Science*, **195**, 393–394.

Bassford, P. J. Jr, Bradbeer, C., Kadner, R. J. & Schnaitman, C. A. (1976). *J. Bacteriol.* **128**, 242–247.

Bassford, P. J., Schnaitman, C. A. & Kadner, R. J. (1977). *J. Bacteriol.* **130**, 750–758.

Berg, D., Davies, J., Allet, B. & Rochaix, J. D. (1975). *Proc. Nat. Acad. Sci., U.S.A.* **72**, 3628–3632.

Blair, D. G., Sherratt, D. J., Clewell, D. B. & Helsinki, D. R. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 2518–2522.

Bukhari, A. I., Shapiro, J. A. & Adhya, S. L. (1977). *DNA Insertions, Plasmids and Episomes,* Cold Spring Harbor Press, Cold Spring Harbor, New York.

Casjens, S. R. & Hendrix, R. W. (1974). *J. Mol. Biol.* **88**, 535–545.

Davies, J., Berg, D., Jorgensen, R., Fiandt, M., Huang, T.-S. R., Courvalin, P. & Schloff, J. (1977), In *Topics in Infectious Diseases* (Drews, J. & Hogenauer, G., eds). vol. 2, p. 101, Springer-Verlag, Vienna.

Frazer, A. C. & Curtiss, R., III (1975). *Curr. Topics Microbiol. Immunol.* **69**, 1–84.

Frost, G. E. & Rosenberg, H. (1973). *Biochim. Biophys. Acta,* **330**, 90–101.

Frost, G. E. & Rosenberg, H. (1975). *J. Bacteriol.* **124**, 704–712.

Gottesman, S. & Beckwith, J. R. (1969). *J. Mol. Biol.* **44**, 117–127.

Grindley, N. D. F. (1977). In *DNA Insertion Elements, Plasmids and Episomes* (Bukhari, A. I., Shapiro, J. A. & Adhya, S. L., eds), pp. 115–123, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Hancock, R. E. W. & Braun, V. (1976). *J. Bacteriol.* **125**, 409–415.

Heffron, F., Bedinger, P., Champoux, J. J. & Falkow, S. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 702–706.

Jones, B. B. & Reznikoff, W. S. (1976). *J. Bacteriol.* **128**, 282–289.

Jones, B. B. & Reznikoff, W. S. (1977). *J. Bacteriol.* **132**, 270–281.

Kadner, R. J. & McElhaney, G. (1978). *J. Bacteriol.* **134**, 1020–1029.

Konisky, J., Soncek, S., Frick, K., Davies, J. K. & Hammond, C. (1976). *J. Bacteriol.* **127**, 249–257.

Laemmli, U. (1970). *Nature (London)*, **227**, 680–685.

Landy, A., Abelson, J., Goodman, M. & Smith, J. (1967). *J. Mol. Biol.* **29**, 457–471.

Mandel, M. & Higa, A. (1970). *J. Mol. Biol.* **53**, 159–162.

Matsushiro, A. (1963). *Virology*, **19**, 475–482.

McCray, J. W. Jr & Herrmann, K. M. (1976). *J. Bacteriol.* **125**, 608–615.

Murialdo, H. & Siminovitch, L. (1972). *Virology*, **48**, 785–823.

Neilands, J. B. (1977). In *Bioinorganic Chemistry* (Raymond, K. N., ed.), vol. 2, pp. 30–32, American Chemical Society, Washington, D.C.

Ohtsubo, H. & Ohtsubo, E. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 615–619.

Postle, K. & Reznikoff, W. S. (1978). *J. Bacteriol.* **136**, 1165–1173.

Pugsley, A. P. & Reeves, P. (1976). *J. Bacteriol.* **126**, 1053–1062.

Pugsley, A. P. & Reeves, P. (1977). *J. Bacteriol.* **127**, 218–228:

Roozen, K. J., Fenwick, R. G. & Curtiss, R. III (1971). *J. Bacteriol.* **107**, 21–33.

Rosenberg, A. H. & Gefter, M. L. (1969). *J. Mol. Biol.* **46**, 581–584.

Schmidt, F., Besemer, J. & Starlinger, P. (1976). *Mol. Gen. Genet.* **145**, 145–154.

Sgaramella, V. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 3389–3393.

Wang, C. C. & Newton, A. (1969). *J. Bacteriol.* **98**, 1135–1141.

Wayne, R. & Neilands, J. B. (1975). *J. Bacteriol.* **121**, 497–503.

Wookey, P. & Rosenberg, H. (1978). *J. Bacteriol.* **133**, 661–666.