Proc. Nat. Acad. Sci. USA
Vol. 72, No. 11, pp. 4506–4510, November 1975
Cell Biology

Use of gene 32 protein staining of single-strand polynucleotides for gene mapping by electron microscopy: Application to the $\phi 80 d_3 ilvsu^+7$ system

(visualization of RNA·DNA duplex regions/rRNA genes/tRNA genes/secondary structure mapping)

MADELINE WU AND NORMAN DAVIDSON

Department of Chemistry, California Institute of Technology, Pasadena, Calif. 91125

Contributed by Norman Davidson, September 2, 1975

ABSTRACT **A method for visualizing RNA·DNA duplex regions along a single strand of DNA in the electron microscope is described. A preparation of RNA molecules is hybridized to a long DNA strand containing the coding sequences (genes) for some of the RNAs. T4 gene 32 protein, which binds selectively and cooperatively only to the single-strand regions, is added, followed by glutaraldehyde. The resulting nucleic acid-gene 32 complex is adsorbed to the surface of an electron microscope grid in the presence of ethidium bromide. The single-strand regions are relatively thick (8.5 nm) compared to the duplex (RNA·DNA hybrid) regions (3.5 nm), so that the two kinds of regions are readily recognized by electron microscopy. In favorable cases, tRNA·DNA hybrids of length about 80 nucleotide pairs can be recognized (although with difficulty). The positions of a number of interesting genetic sequences on the DNA of the transducing phage $\phi 80 d_3 ilvsu^+7$ have been mapped. The *r* strand contains 16S, 23S, and 5S rRNA coding sequences in that order. The spacer between 16S and 23S genes has a length of 500 nucleotides and contains the coding sequence for a $tRNA_2^{Glu}$ gene in agreement with previous biochemical observations. The spacer between the 23S and 5S genes has a length of 180 nucleotides. The su^+7 $tRNA^{Trp}$ coding sequence has been mapped on the *l* strand at a position just to the left of the *ilv* genes. Secondary structure loops due to short inverted repeat sequences flanking the 16S, 23S, $tRNA^{Trp}$, and F sequences in the DNA have been observed.**

The position on a DNA strand of a sequence (gene) coding for the synthesis of a particular RNA molecule can be mapped by hybridizing the RNA to the DNA strand *in vitro* and observing the position of this duplex region along the single strand of DNA by electron microscopy. In the standard cytochrome *c* spreading method, duplex regions are thicker and smoother than single strands; the position of a junction between a single strand and a DNA·DNA duplex region can be recognized and mapped fairly confidently provided the respective segments are at least about 500 nucleotides long. The contrast between RNA·DNA hybrids and DNA single strands is not as good; however, in favorable cases mapping is possible but difficult (1–3).

T4 gene 32 protein binds selectively and cooperatively to single-stranded polynucleotides. In formamide cytochrome *c* spreads the single-strand DNA-gene 32 protein complex is thicker than duplex DNA, but discrimination is still difficult (4). Koller, Sogo, and Bujard (5) have found that in the presence of excess ethidium bromide (EtBr), duplex DNA may be mounted on mica or a carbon film without any kind of protein monolayer. In their mounting technique duplex DNA is about 1.8–2.5 nm in diameter.

We have found that using a modified version of this EtBr mounting technique, single strands of DNA complexed with gene 32 protein are relatively thick and easy to see; the discrimination between single-strand and duplex regions is much clearer than in the basic protein methods. In the present communication we describe our application of this method for mapping the positions of the 16S, 23S, and 5S rRNA genes and the several tRNA genes present on the DNA of the transducing phage $\phi 80 d_3 ilvsu^+7$ that was isolated by L. Soll (3).

Abbreviations: C_rt (or C_0t), initial concentration of total RNA (or DNA) (mol of nucleotide/liter) × time (sec); EtBr, ethidium bromide; RF, replicative form.

MATERIALS AND METHODS

The bacteriophages $\phi 80$ and $\phi 80 psu^+3$ were grown from lysogens kindly provided by Dr. J. D. Smith and were handled as described (6). A mixture of $\phi 80 att_{80} i^\lambda (cI857)$ (hy5) and $\phi 80 d_3 ilvsu^+7$ (abbreviated henceforth as $\phi 80 d_3 ilv$) was grown from a thermally inducible lysogen prepared and

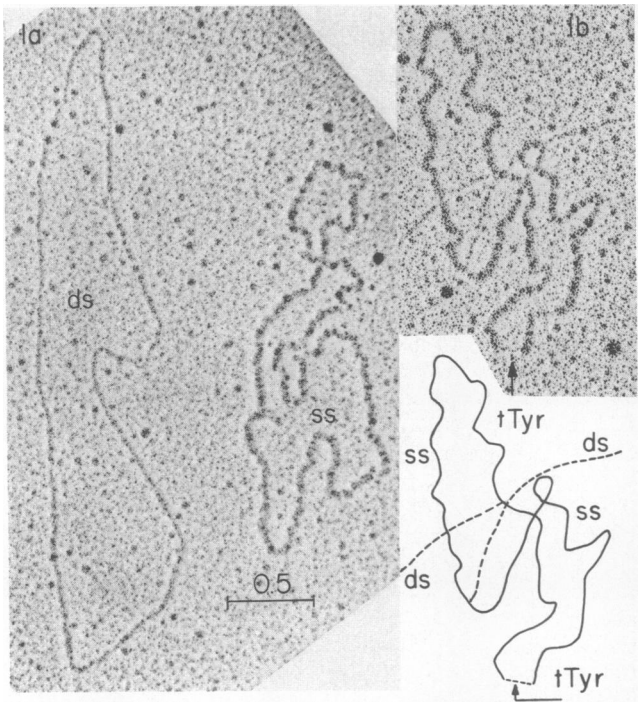

FIG. 1. (a) Electron micrographs of single-strand (ss) ϕX174 and double-strand (ds) ϕX174-RFII DNA molecules mounted by the gene 32-EtBr method. This and all other calibration markers are in kilobase units, with 1 kilobase unit ≈ 0.46 μm. (b) The substitution loop of the $\phi 80 psu^+3/\phi 80$ heteroduplex. The micrograph is interpreted in the tracing below it. Single-strand (ss) and double-strand (ds) DNA regions are identified. The short duplex due to the $tRNA^{Tyr}$ RNA-DNA hybrid is identified as tTyr.

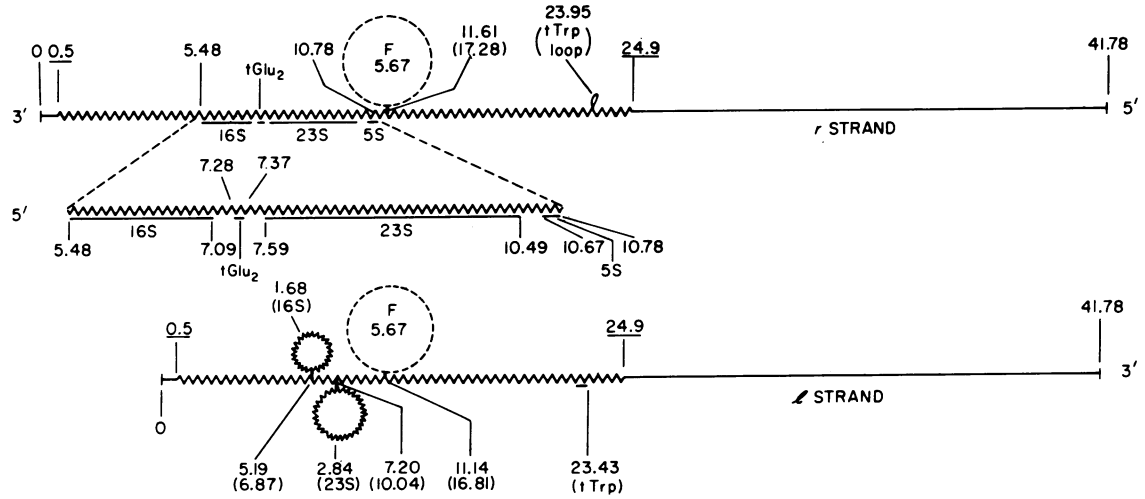

FIG. 2. Map of the positions of the various genetic sequences on the $\phi 80 d_3 ilv su^+7$ DNA molecule as determined in this and previous investigations. Numerical coordinates are in kilobase units starting at the left end of the DNA molecule. The straight solid lines represent phage DNA; ⎓ represents bacterial sequences on the transducing phage DNA; ⎓ represents a DNA-RNA hybrid: The 2.8–8.5 F sequence extends from 11.61 to 17.28 kilobase. The secondary structure loops discussed in the text are shown. $tGlu_2$ and tTrp stand for the respective tRNA genes. A magnified representation of the *r* strand around the rRNA genes is shown. Typical statistical data for our various measurements are as follows. When hybridizations were performed with 16S + 23S + 5S rRNA, 54 out of 100 well extended molecules had both 16S and 23S duplex regions; 30 of these had the F loop, and 28 of these 30 had a clear 5S hybrid. When 5S rRNA was absent, two out of 29 molecules with an F loop seemed to have a duplex region at the 5S gene position. Without $tRNA_2^{Glu}$, 22 of 30 molecules had an intact spacer region between the 16S and 23S duplex regions; seven seemed to have a short duplex at the tRNA position; one was questionable. With added $tRNA_2^{Glu}$, 24 out of 30 molecules had a tRNA-DNA duplex, three had intact spacers, and three were questionable. $tRNA^{Trp}$ did not cause duplex formation in the rRNA spacer region. On the *l* strand (no 16S or 23S hybrids), of 172 molecules, 69, 46, and 57 had the 16S, 23S, and F loops, respectively. On the *r* strand or on the *l* strand in the absence of $tRNA^{Trp}$, the $tRNA^{Trp}$ loop occurred in 39 of 54 molecules. On the *l* strand, with $tRNA^{Trp}$ added, 33 out of 70 appeared to have a $tRNA^{Trp}$-DNA duplex; eight of these had the new larger loop. Standard deviations in length measurements were all about 10%.

provided by Dr. L. Soll. The transducing phage was purified from the helper phage by CsCl banding. *E. coli* 5S rRNA, a mixture of *E. coli* 16S and 23S rRNA, and *E. coli* total tRNA were purchased from the Miles Laboratories. Purified *E. coli* $tRNA_2^{Glu}$ and $tRNA^{Trp}$ were generous gifts from Drs. Steve Eisenberg and M. Yarus.

Single-stranded DNA for electron microscopy was prepared by alkaline lysis of the phage in 0.1 M NaOH after a preincubation with neutral Na_3 EDTA as described (3, 6). $\phi 80 psu^+3/\phi 80$ heteroduplexes were prepared as described (6). RNA-DNA hybrids of various RNA components with $\phi 80 d_3 ilv$ single strands (or to the $\phi 80 psu^+3/\phi 80$ heteroduplex) were prepared by hybridization in a 25-μl volume of solution containing 50% formamide, 0.1 M Tris, pH 8.5, 0.01 M Na_3EDTA, 0.65 M NaCl, 1 μg/ml of $\phi 80 d_3 ilv$ DNA, and RNA components as specified below. Hybridization was at 53° for 5 min. RNA components, when added, were at the following concentrations, corresponding to the C_rt values (in mol·sec/liter) given: 16S + 23S rRNA, 28 μg/ml (C_rt = 0.026); 5S rRNA, 15 μg/ml (C_rt = 0.014); $tRNA_2^{Glu}$ and/or $tRNA^{Trp}$ at 5 μg/ml (C_rt = 0.005 for each); total *E. coli* tRNA, 250 μg/ml (C_rt = 0.23). We estimate $C_rt_{1/2}$ values under these hybridization conditions of about 0.008, 0.0002, and 0.0003 for 16S + 23S rRNA, 5S rRNA, and a purified tRNA, respectively. At the DNA C_0t value for the hybridization (0.0009), not very much DNA·DNA reassociation occurs.

The hybridization reaction mixture or other DNA sample was dialyzed into 0.01 M phosphate buffer, 0.001 M Na_2-EDTA, pH 7.4. A solution of gene 32 protein was added to give a concentration of 20–50 μg/ml, with a DNA concentration of 0.5–1 μg/ml. After 5 min of incubation at 37°, glutaraldehyde was added to 0.2%. The solution was incubated again for 10 min and diluted about 10-fold into a solution containing 50 μg/ml of EtBr, 0.01 M tricine buffer, 0.001 M Na_3EDTA, pH 9.0. A 10-μl drop of this solution was mounted on Parafilm and kept for 10 min at room temperature. A carbon-coated parlodion grid was then touched to the surface of the DNA droplet. The grid was dehydrated in ethanol, stained with 1×10^{-4} M uranyl acetate in 90% ethanol for 30 sec, rinsed in isopentane, and rotary shadowed with platinum at a 1:8 angle. Electron micrographs were usually taken at a magnification of 18,000 with a Philips 301 electron microscope.

RESULTS

ϕX174 and $\phi 80 psu^+3$ DNAs

Fig. 1a is an electron micrograph of a mixture of single-strand (ss) ϕX174 and open circular double-strand (ds) ϕX174 replicative form (RF) DNA that had been treated with gene 32 protein and mounted as described in *Materials and Methods*. The single-strand and duplex DNAs have thicknesses of 8.0 ± 1.5 and 3.5 ± 1.2 nm, respectively; their lengths are 2.38 ± 0.15 and 2.41 ± 0.11 μm. If we take the length of ϕX DNA as 5200 nucleotides, this corresponds to a spacing of 0.0460 nm per base or base pair. The main points are that gene 32 is bound continuously to the single-strand DNA but not at all to the duplex DNA, and that the difference in appearance between duplex and single-strand DNA is very clear. The molecules are well extended and reasonably homogeneous in length.

Fig. 1b shows a part of a $\phi 80 psu^+3/\phi 80$ heteroduplex that has been hybridized with total *E. coli* tRNA. This transducing phage DNA carries a single $tRNA^{Tyr}$ gene. Bacterial and phage single-strand segments of the heteroduplex have measured lengths of 3000 and 2000 nucleotides, respectively, in agreement with our previous study (6). We observe a

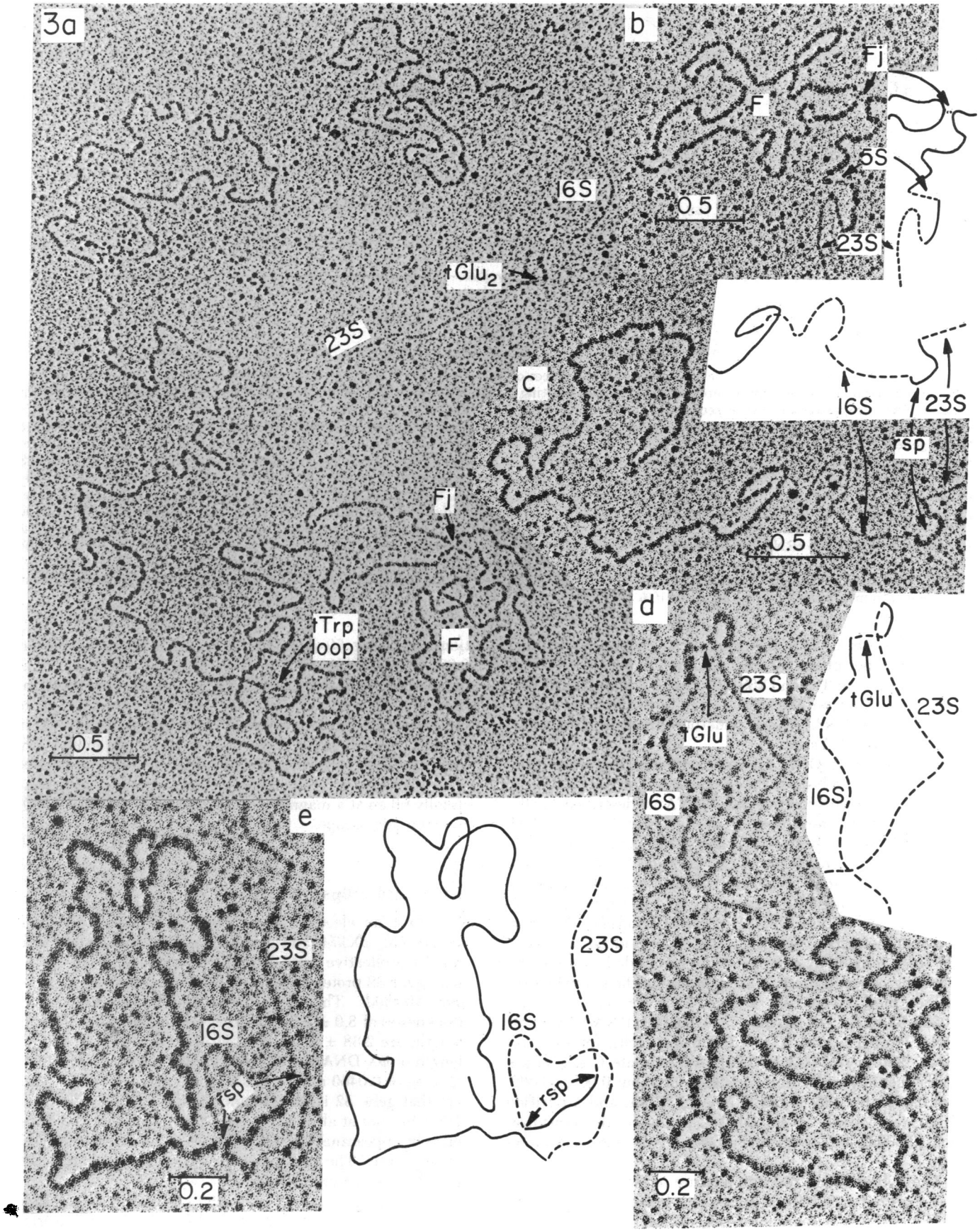

FIG. 3. *(Legend appears at bottom of following page.)*

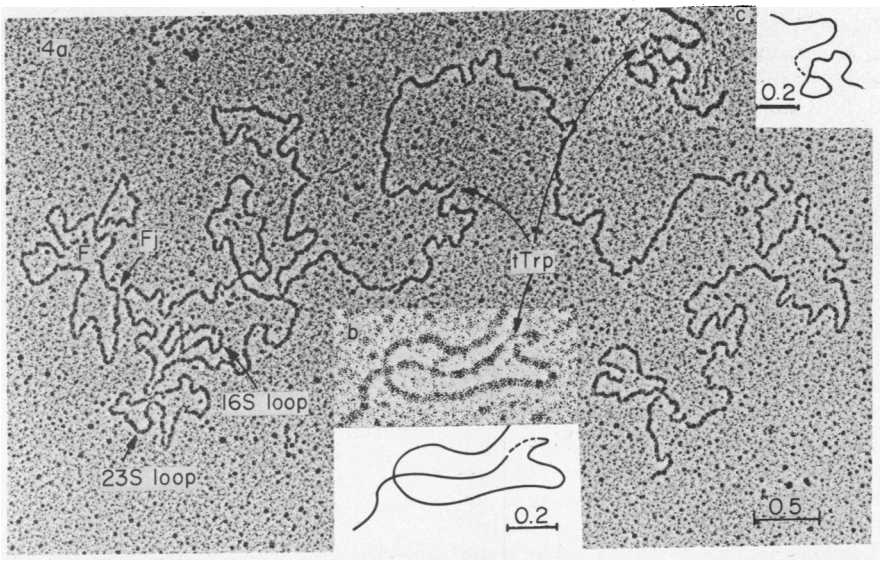

FIG. 4. Micrographs of the *l* strand of $\phi 80 d_3 ilv$. (a) A complete single strand showing the $tRNA^{Trp}$-DNA hybrid, and the F, 16S, and 23S secondary structure loops. (b) A higher magnification view plus a tracing of a $tRNA^{Trp}$-DNA hybrid region. (c) A $tRNA^{Trp}$ hybrid with the adjacent "larger" secondary structure loop.

short duplex segment (indicated by an arrow and tTyr) with a measured length of 85 ± 10 nucleotides, corresponding to the tRNA·DNA hybrid region. This duplex region was seen at the high frequency of 24 out of 50 heteroduplexes (theory for 100% reaction, 25/50) and at the expected position (6) if and only if *E. coli* tRNA was present in the hybridization solution.

Fig. 2 is a map of some of the identified genetic segments on the DNA of the bacteriophage $\phi 80 d_3 ilv$. The figure includes results from previous studies in this laboratory (3), from the work of Lund, Blattner, and Dahlberg (personal communication), and from the present study. The previous electron microscopic mapping study (3) had given the following relevant results: (*i*) There are $\phi 80$, F, and bacterial chromosomal sequences on the DNA as shown. (*ii*) There are 16S and 23S rRNA genes at the positions shown, with a spacer of length 0.57 ± 0.13 kilobase between them. (*iii*) There is an F sequence of length 5.7 kilobase, denoted as the 2.8–8.5 F sequence, slightly to the right of the rRNA genes. (*iv*) There is a segment of bacterial chromosomal DNA of length 7.6 kilobase carrying some *ilv* genes and possibly the su^+7 $tRNA^{Trp}$ gene. Lund, Blattner, and Dahlberg have mapped the several *Eco*RI restriction fragments of $\phi 80 d_3 ilv$; their relevant results are: (*v*) There is one fragment which carries both 16S and 23S sequences, and which also has a $tRNA_2^{Glu}$ gene. Thus this tRNA gene must occur in the spacer region. (*vi*) There is a fragment carrying 23S and 5S rRNA sequences. Thus, as expected from transcription mapping studies, the 5S rRNA gene is to the right of the 23S rRNA gene (7). (*vii*) A larger restriction fragment with sequences extending from somewhere to the right of the rRNA genes over into the $\phi 80$ sequences hybridizes with whole *E. coli* tRNA. The presumption therefore is that this fragment contains the su^+7 tRNA, which is the mutant $tRNA^{Trp}$ (8).

Our results reported below add the following information as to the gene arrangement in $\phi 80 d_3 ilv$. T. Broker and L. Chow, working in this laboratory, had previously observed that the 2.8–8.5 F sequence on the phage DNA sometimes forms a loop structure when spread from weakly denaturing conditions (40–50% formamide) in regular cytochrome *c* spreads (personal communication). The loop is presumably formed by short inverted repeat sequences at its two ends. The F loop structure is also seen at a frequency of about 50–60% under our mounting conditions (Fig. 3a and b). The barely discernible duplex segment (denoted by Fj) separating the F loop from the rest of the $\phi 80 d_3 ilv$ DNA has a length corresponding to 15 ± 7 nucleotides. The F loop is a useful reference feature for mapping other features on a single strand of $\phi 80 d_3 ilv$ DNA.

Electron micrographs of 16S and 23S rRNA·DNA hybrid regions on $\phi 80 d_3 ilv$ DNA are also shown in Fig. 3a. Fig. 3c shows a single-strand spacer (rsp) between the two large rRNA genes at a higher magnification. The length (0.50 ± 0.03 kilobase of the spacer is in agreement with the previous measurements of 0.57 ± 0.13 kilobase (3). If $tRNA_2^{Glu}$ is present in the hybridization mixture a duplex region of length 95 ± 17 nucleotides is seen in the spacer region (Fig. 3a and d), in agreement with the results of Lund *et al.* If 5S rRNA is present in the hybridization medium, a hybrid region of length 110 ± 24 nucleotides due to this gene is found between the 23S gene and the F loop (Fig. 3b). The spacer between the 23S and 5S genes has a length of 0.18 ± 0.02 kilobase.

Since the order of transcription of the rRNA genes is 16S, 23S, and 5S (7), we may identify the strand carrying their coding sequences as the *r* strand of $\phi 80 d_3 ilv$. We can identify the *r* strand and the *l* strand most of the time, because the former has 16S and/or 23S rRNA hybridized to it in all of our grids. We find that the su^+7 $tRNA^{Trp}$ coding sequence

FIG. 3 (*on preceding page*). Micrographs of the *r* strand of $\phi 80 d_3 ilv$. (a) A complete strand showing the 16S and 23S rRNA·DNA hybrid regions, a short duplex due to the $tRNA_2^{Glu}$-DNA hybrid, the F loop (joined to the rest of the DNA by a short inverted repeat sequence Fj), and the $tRNA^{Trp}$ secondary structure loop. (b) A portion of a micrograph and a tracing (to the right) showing the F loop, the 5S rRNA·DNA hybrid, and part of the 23S hybrid region. (c) The 16S rRNA hybrid region, part of the 23S rRNA hybrid, and the single-strand spacer (rsp) between them. (d) The $tRNA_2^{Glu}$-DNA hybrid in the spacer region between the 16S and 23S hybrid regions, with an interpretative tracing. (e) A 16S rRNA secondary structure loop containing a 16SrRNA·DNA hybrid.

is present on the *l* strand at a position 6.79 ± 0.61 kilobase to the right of the F loop, as shown in Fig. 2. The uncertainty in this number and the uncertainty as to the left terminus of the *ilv ADE* operon (3, 9) is such that the $tRNA^{Trp}$ gene is either just to the left of the *ilv* genes or embedded in them (the former being the more plausible alternative). The order of transcription of *ilv ADE* is to the left (10); our physical mapping implies that the $tRNA^{Trp}$ gene is also transcribed leftward. There are several lines of evidence for our assignment of the position of the $tRNA^{Trp}$ gene. There is a small secondary structure loop fairly frequently seen at the position now assigned to $tRNA^{Trp}$ on both the *r* and the *l* strands (the tTrp loop, Fig. 3a). When purified $tRNA^{Trp}$ is added to the hybridization mixture this loop is not seen on the *l* strands. Instead, a short double-strand region (93 ± 26 nucleotides) is seen at the same position (Fig. 4a and b); a larger secondary structure loop to the left of this duplex region is then frequently present (Fig. 4c). Thus, the DNA sequences for $tRNA^{Trp}$ seem to be involved in the small secondary structure loop denoted as tTrp in Figs. 2 and 3a; when they are hybridized to $tRNA^{Trp}$, the flanking sequences can interact with other sequences to form a different secondary structure loop. Secondary structure features associated with unhybridized tRNA genes have been seen previously (11).

We have looked carefully at both strands after hybridizing whole *E. coli* tRNA to the DNA. The $tRNA_2^{Glu}$ hybrid is seen at a frequency of about 80%; the $tRNA^{Trp}$ hybrid is seen at the rather low frequency of about 20%. No other tRNA genes were identified. With the purified $tRNA^{Trp}$, the hybrid region was seen on 50% of the *l* strands.

Interesting secondary structure loops are also seen at the positions of the 16S rRNA and 23S rRNA genes on the *l* strand at frequencies of 65 and 45%, respectively (Fig. 4a). We presume, but do not know, that the same loops would be seen in the *r* strand in the absence of added 16S and 23S rRNA. The 16S rRNA secondary structure loop is sometimes seen on the *r* strands even when a 16S rRNA·DNA duplex region is present (Fig. 3e); this suggests that the inverted repeat sequences that pair to form the loop lie partly or totally outside of the sequences that make up the mature 16S rRNA. These inverted repeat sequences flanking the rRNA genes may be important for the processing by RNase III of pre-rRNA transcripts (12).

FURTHER DISCUSSION

We think it is obvious from the micrographs shown that the method described here is an improvement over the cytochrome *c* method for visualizing and mapping RNA·DNA hybrid regions on a single strand of DNA. The identification of long duplex regions such as those for the 16S and 23S rRNA genes is very easy and certain. The duplex regions for the several tRNA and the 5SrRNA·DNA hybrids are short and rather difficult to see. In the present case, with sufficient background information about the genes present and with the perhaps fortuitous circumstances that some secondary structure features are present as nearby reference marks or as features which disappear when the hybrids are formed, the 5S and tRNA genes could be positively identified and mapped. It may be that for some difficult and complicated problems with less prior information, a technique that combines the present gene 32 protein staining and mounting method with a method for attaching a suitable electron microscope label to the short RNA will be needed for positive identification of short genes.

Note Added in Proof. In a recent paper, Brock *et al.* (13) have published electron micrographs of partially denatured duplex DNA in which the single-strand regions were stained with gene 32 protein, and the molecules then mounted on a thin carbon support in the absence of a protein monolayer. The discrimination between single-strand and duplex regions was excellent.

Drs. Bruce Alberts and Peter von Hippel have made generous gifts of gene 32 protein and have contributed advice and information about its properties. We are indebted to Drs. Steve Eisenberg and Mike Yarus for gifts of purified $tRNA^{Glu}$ and $tRNA^{Trp}$. We have profited from consultations with Drs. Larry Soll, M. Yarus, and Harumi Kasamatsu. We had been informed prior to our studies of the general results of the mapping experiments by Elsebet Lund, James Dahlberg, and Fred Blattner as to the positions of the several short RNA species on $\phi 80 d_3 ilv$ as described in the *text*. We have profited from the previous preliminary studies of this system, including the discovery of the F loop, by Drs. Louise Chow and Thomas Broker. This research has been supported by Grant NIGMS 20927 from the USPHS.

1. Wu, M. & Davidson, N. (1973) *J. Mol. Biol.* **78**, 1–21.
2. Hyman, R. W. (1971) *J. Mol. Biol.* **61**, 369–376.
3. Ohtsubo, E., Soll, L., Deonier, R. C., Lee, H. J. & Davidson, N. (1974) *J. Mol. Biol.* **89**, 631–646.
4. Delius, H., Mantell, N. J. & Alberts, B. (1972) *J. Mol. Biol.* **67**, 341–350.
5. Koller, T., Sogo, J. M. & Bujard, H. (1974) *Biopolymers* **13**, 995–1009.
6. Wu, M. & Davidson, N. (1973) *J. Mol. Biol.* **78**, 35–41.
7. Dolittle, W. F. & Pace, N. R. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 1786–1790.
8. Yaniv, M., Folk, W. R., Berg, T. & Soll, L. (1974) *J. Mol. Biol.* **86**, 245–260.
9. Lee, H. J., Ohtsuho, E., Deonier, R. C. & Davidson, N. (1974) *J. Mol. Biol.* **89**, 585–597.
10. Taylor, A. L. & Trotter, C. D. (1972) *Bacteriol. Rev.* **36**, 504–524.
11. Wu, M., Davidson, N. & Carbon, J. (1973) *J. Mol. Biol.* **78**, 23–34.
12. Nikolaev, N., Silengo, L. & Schlessinger, D. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 3361–3365.
13. Brock, C., Bickle, T. A. & Yuan, R. (1975) *J. Mol. Biol.* **96**, 693–702.