Intermediates in the degradation of mRNA from the lactose operon of Escherichia coli

Joseph R. McCormick*, Janice M. Zengel and Lasse Lindahl*
Department of Biology, University of Rochester, Rochester, NY 14627, USA

Received October 2, 1990; Revised and Accepted April 16, 1991

ABSTRACT

We have analyzed the processing of mRNA from the lac operon in an Escherichia coli strain carrying the lac on a multicopy plasmid. Messenger RNA was analyzed by hybridization and nuclease protection of pulse-labeled RNA and precursor-product relationships were determined by quantitating radioactivity in primary and processed transcripts at various times after induction of the lac promoter or inhibition of transcription with rifampicin. Our results support the existence of two types of processed transcripts with endpoints in the lacZ-lacY intercistronic region. One of these carries lacZ sequences and has a 3' endpoint about 30 bases downstream of this gene. The other carries lacY sequences and has a 5' end in the translation termination region of the lacZ gene. Finally, we have found evidence that transcription is continued at least 268 bases beyond the last gene (lacA) and that this 3' non-translated region is shortened by post-transcriptional processing.

INTRODUCTION

Synthesis and degradation of mRNA molecules are both important for determining the steady state concentration of a given mRNA, and thus for defining the rate of synthesis of protein. Whereas many fundamental aspects of mRNA synthesis are fairly well elucidated, only a rudimentary understanding of the degradation reactions has yet been achieved. Several enzymes have been implicated in posttranscriptional mRNA processing and degradation in Escherichia coli, including RNase III (1–4), RNase II (5), polynucleotide phosphorylase (5), RNase E (6, 7), RNase M (8) and a number of unidentified endonucleases (7, 9–14). (Since some of these enzymes cleave outside the translated portions of the mRNA, it is not clear if they should be classified as processing or degradation enzymes). Additionally, the rate with which mRNA is turned over in E. coli and other bacteria appears, at least in some cases, to be influenced by mRNA secondary structure (e.g. 1, 15).

The lac operon of E. coli has served as a model system for many aspects of gene expression, including turnover of mRNA (16–18). We have used a new approach involving nuclease mapping of pulse-labeled transcripts to improve our understanding of the processing (or degradation) of lac mRNA. Our results demonstrate the existence of both 5' and 3' ends in the intercistronic region between the two proximal genes of the operon, suggesting that the lac transcript is processed into separate fragments carrying either lacZ or lacY sequences. In contrast to previous reports (17, 18), we observed little or no accumulation of RNA molecules with endpoints within lacZ, indicating that this part of the transcript is eliminated very rapidly once degradation has been initiated. Finally, our data suggest that the primary transcript of the lac operon extends at least 268 bases beyond the last gene (lacA) and is then shortened by post-transcriptional processing.

MATERIALS AND METHODS

Bacterial strains and plasmids

NF314 is E. coli B strain AS19 leu which is sensitive to low concentrations of rifampicin (19, 20). LL308 is E. coli K12 F' pro+ lacPF2AM15Y+Δ(pro-lac) nalA recA supE thi. The pSC101-derived plasmid pLF149 (Fig. 1a) has been described previously (21). Plasmid pPZYA (R. Guadino and E.A. Morgan, Roswell Cancer Institute, Buffalo, NY) is a pBR322 derivative and carries a trpA'/lacZ fusion followed by the lacY and A genes. Plasmid pJR19 was constructed (22) from Xplac (23) and pPZYA. It contains the wild-type lac operon, including the promoter-operator region. Both pPZYA and pJR19 carry a promoter for SP6 RNA polymerase upstream of the lac genes. The source of the untranslated region distal to lacA on pJR19 and pPZYA is pMC81 (24) which has been shown to contain at least 1.1 kb of the DNA sequences distal to the lacA gene on the chromosome (25). Thus pJR19 is useful for studying the transcription and processing of mRNA from the entire lac operon, including the distal end.

Media and growth conditions

Cultures were grown exponentially at 30°C in MOPS minimal medium (26) supplemented with 0.2% glucose and 0.18 mM K2HPO4, except that RNA for primer extension experiments was prepared from cultures grown with 1 mM K2HPO4.

* To whom correspondence should be addressed

* Present address: Department of Cellular and Developmental Biology, Harvard University, Cambridge, MA 02138, USA
Cultures of NF314 harboring pLF1Δ49 were also supplemented with 40 μg leucine per ml and 20 μg oxytetracycline per ml. Cultures of LL308 containing pJR19 also contained 1 μg thiamine per ml and 100 μg ampicillin per ml. Cell growth was monitored as optical density (OD) at 450 nM in a Hitachi model 124 spectrophotometer. OD<sup>450</sup> = 1 corresponds to approximately 2×10<sup>6</sup> cells per ml.

**RNA decay experiments**

Cultures of pLF1Δ49/NF314 at OD<sup>450</sup> = 0.15 were supplemented with 10 μCi [3H]uridine per ml and incubation was continued until OD<sup>450</sup> reached 0.6-0.7. <sup>32</sup>PPO<sub>4</sub> was then added to a final concentration of 0.8 mCi per ml and 5 minutes later 1/200 volume of a freshly prepared solution of 20 mg rifampicin per ml in methanol was added. At the indicated times after addition of rifampicin, 3 ml samples of the culture were lysed by mixing with boiling buffer containing 1% sodium dodecylsulfate (27). Yeast RNA (100 μg per sample) was added later 1/200 volume of a freshly prepared solution of 20 mg rifampicin per ml in methanol was added. At the indicated times after addition of rifampicin, 3 ml samples of the culture were lysed and RNA was prepared as described above. The relative recovery of RNA was determined from the trichloroacetic acid precipitable <sup>3</sup>H radioactivities.

**Transcription induction experiments**

Cultures of pJR19/LL308 were supplemented with 4

**Hybridization and T1 mapping of RNA transcripts**

This procedure, adapted from Hansen and Sharp (28), has been described in detail (27). In brief, equal amounts of RNA (determined from the trichloroacetic acid precipitable radioactivities) corresponding to 0.2-0.4 ml of cell culture were hybridized to recombinant single-stranded DNA (1 μg per reaction) carrying the indicated insert (Table 1 and figures). All hybridization probes were derived from M13, except pJR15 for which the vector pD4 was used (29). Non-hybridized RNA was digested with RNase T1 and protected RNA fragments were analyzed by electrophoresis through a 7 M urea-8% polyacrylamide gel (in Fig. 4b, panel A, a 4% polyacrylamide gel was used). Molecular weight markers were prepared by run-off transcription with T7 RNA polymerase of linearized plasmids with known sequence. Radioactive bands were visualized by autoradiography using Cronex III (DuPont) intensifying screens and Kodak XAR film. The radioactivity in bands was quantitated by microdensitometry, using a standard curve relating radioactivity to grain intensity in the film (27).

**Primer extension experiments**

Non-radioactive "in vivo" RNA was prepared from cultures of NF314 containing either pLF1Δ49 or pSC101 at OD<sup>450</sup> = 0.9-1.0 as described for the labeling experiments above. "In vitro" RNA was prepared by transcription of pJR19 with SP6 RNA polymerase in a reaction essentially as described (30). Following RNA synthesis, the pJR19 DNA was degraded with FPLC-purified DNase I and the RNA was purified by extractions with phenol and chloroform:isoamyl alcohol (24:1). The oligonucleotide primer (5'-CGGGAAAGTAGGCTCCC-3') is complementary to mRNA from lacY such that primer extension will be initiated 68 bases from 5' end of lacY. Annealing mixtures (10 μl) contained 2 μl 5× annealing buffer [250 mM Tris-HCl (pH 8.0), 500 mM KCl], 0.6 pmol <sup>32</sup>PPO<sub>4</sub> 5'-end labeled primer and RNA template (either in vivo RNA from 0.3 ml culture or approximately 0.1 μg in vitro RNA). The mixtures were first incubated at 65°C for 10 min then and at 45°C for 2 hrs. Extension reactions contained 3 μl annealing mixture, 1 μl M-MLV reverse transcriptase (200 units/μl; Bethesda Research Laboratories, Inc.), and 1 μl 5× reverse transcription buffer [250 mM Tris-HCl (pH 8.0), 25 mM dithiothreitol, 15 mM MgCl<sub>2</sub>, 5 mM of each of the 4 dNTP's]. The final KCl concentration is thus 60mM. In some cases the 5× reverse transcription buffer also contained 300 mM KCl, resulting in a final KCl concentration of 120 mM. Following incubation for 20 min at 37°C or 45°C, the reaction was terminated by addition of one volume deionized formamide and the reaction products were analyzed by gel electrophoresis as above. Markers were produced from dideoxynucleotide sequencing of single-stranded DNA from M13 phage J41 which carries the lac DNA fragment also present on J40 (Fig. 1a), but in the opposite orientation. The primer was the same used for the primer extension reaction, except that it was not 5' phosphorylated. Reactions were performed in the presence of <sup>35</sup>S-dATP under the conditions recommended by the supplier of the DNA polymerase (United States Biochemicals, Inc.).

**RESULTS**

**Experimental strategy**

Cleavage of primary mRNA transcripts can be characterized either as processing, leading to a shorter but translatable messenger, or as degradation, leading to an untranslatable mRNA. However, this distinction is often not possible, because one does not know if removal of non-translated regions affects the translation, and cleavage of polystronic transcripts may inactivate some, but not all, cistrons. We are therefore not distinguishing between mRNA processing or degradation, i.e. both terms are used synonymously.

<table>
<thead>
<tr>
<th>Table 1.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Probe</td>
<td>DNA derived from</td>
</tr>
<tr>
<td>M65</td>
<td>S10 operon</td>
</tr>
<tr>
<td>pJR15</td>
<td>lac operon</td>
</tr>
<tr>
<td>M60</td>
<td>lac operon</td>
</tr>
<tr>
<td>M605</td>
<td>lac operon</td>
</tr>
<tr>
<td>M563</td>
<td>lac operon</td>
</tr>
<tr>
<td>J32</td>
<td>lac operon</td>
</tr>
<tr>
<td>J39</td>
<td>lac operon</td>
</tr>
<tr>
<td>J40</td>
<td>lac operon</td>
</tr>
<tr>
<td>J44</td>
<td>lac operon</td>
</tr>
<tr>
<td>J46</td>
<td>lac operon</td>
</tr>
<tr>
<td>J47</td>
<td>lac operon</td>
</tr>
</tbody>
</table>

<sup>1</sup> J39 also carries 14 bases of vector sequence distal to lacY on plasmid pLF1Δ49. Transcription from pLF1Δ49 extends into this sequence.

End points refer to the sequences in GenBank (lac) or EMBL (S10 operon) data bases. Accession number for the S10 operon is X02613. Accession numbers for the lac operon are given in the legend to Fig. 2. Endpoints for the genes in the lac operon are: 1284-4358 (lacZ), 4410-5662 (lacY), and 5727-6338 (lacA).
To identify intermediates in the degradation of primary mRNA transcripts we pulse labeled RNA with radioactive phosphate and hybridized it to single-stranded DNA probes carrying fragments from various regions of the lac transcription unit. Unhybridized RNA was degraded with RNase T1; nuclease resistant radioactive RNA fragments were analyzed by gel electrophoresis. If lac mRNA processing or degradation proceeds through intermediate(s) with endpoints within a given hybridization probe (and the intermediates exist long enough to accumulate in detectable quantities), two types of protected fragments are expected: (i) 'Probe-length fragments' resulting from hybridization of unprocessed transcripts. These RNA fragments are slightly longer than the insert in the hybridization probe, because RNase T1 cuts only at G's. (ii) Fragments shorter than the probe length generated by hybridization of molecules with a processed end within the region complementary to the probe.

Ideally, one should be able to distinguish intermediates generated by exonucleases from endonucleolytic products by the number of fragments shorter than the probe length. Exonucleolytic events should lead to a single fragment, whereas products of endonucleolytic cleavage would yield a pair of fragments the sum of whose lengths equals the probe length. However, in practice the distinction between these possibilities is ambiguous, since the cleavage products themselves are subject to further enzymatic attacks. The two products of an endonucleolytic event may be consumed at different rates and only one of the products might accumulate in detectable amounts. A single processing product (fragment shorter than probe length) could thus be due to exonucleolytic degradation, or to endonucleolytic cleavage yielding products with different life-times.

Demonstration of specific mRNA fragments
Since our laboratory is investigating the expression of ribosomal protein genes, the first object of our search for mRNA degradation intermediates was the transcript of a fusion gene, carried by the plasmid pLF1Δ49, consisting of the 5' half of the gene for ribosomal protein S10 fused in frame to the 8th codon of lacZ (Fig. 1a). The fusion gene transcript includes a mutated S10 leader with a 49-base deletion which eliminates the autogenous control of transcription and translation governing the wild-type S10 operon (21). Distal to the S10'/lacZ-lacY intercistronic region is the intact lacZ-lacY intercistronic region and a fragment of the lacY gene (Fig. 1a).

We labeled cells with $^{32}$PO$_4$ for 5 minutes at 30°C before blocking further transcription initiation with rifampicin. RNA was then extracted from samples of the culture at 2, 5 and 7 minutes after rifampicin addition and subjected to RNase T1 mapping.

Figure 1. Identification of endpoints in the lacZ-Y intercistronic region in RNA transcribed from the S10'/lacZ-lacY fusion operon. (a) Map of the S10-lac region of pLF1Δ49. In addition to the S10'/lacZ fusion gene, the plasmid carries the lacZ-Y intercistronic region and the proximal part of the lacY gene. The filled bars below show the fragments (see Table 1 for precise endpoints) harbored in the hybridization probes used in (b). The wavy lines are our interpretation of the mapping of RNA fragments protected by probe J40 [see (b)]. (b) RNase T1 mapping of RNA. A culture of pLF1Δ49 was labeled with $^{32}$PO$_4$ for 5 min prior to rifampicin addition. RNA was isolated at different times after addition of rifampicin and hybridized to J39 or J40 DNA. Hybridized fragments protected from RNase T1 degradation were analyzed by gel electrophoresis and displayed by autoradiography. The three lanes shown for each probe represent hybrids made with RNA prepared 2 (left), 5 (middle), and 7 minutes (right) after addition of rifampicin. Arrows on the left indicate protected fragments; arrows on the right indicate RNA size markers (the Z and Y arrows point to fragments protected by J39; the corresponding fragments protected by J40 migrate slightly faster). (c) Decay kinetics of protected fragments. The amounts of radioactivity in the bands in the gel shown in (b) were quantitated by densitometry of the X-ray film. Each point is the average of the J39 and J40 lanes, normalized to the amounts found in each fragment at 2 min. [The amounts of radioactivity found in each band increase for the first two minutes after rifampicin addition (data not shown) as expected because the time needed for an RNA polymerase to transcribe the region from the promoter through the end of the S10'/lacZ gene is about 2 min at 30°C (31)].
Experiments employing probes containing the intercistronic region between \( \text{lacZ} \) and \( \text{lacY} \) (Fig. 1a) generated not only the probe-length RNA fragments but also several smaller fragments. For example, the probe J39 (Fig. 1a) protected a probe-length fragment about 350 bases long (labeled ZY in Fig. 1b), two fragments about 250 bases long (labeled Y) and two fragments of about 85 bases (labeled Z). The Z and Y fragments are not artifacts of the RNase T1 mapping procedure (see the discussion of the experiment shown in Fig. 5b). These experiments therefore demonstrate the existence of several classes of RNA molecules with one endpoint in the \( \text{lacZ-Y} \) intercistronic region and one endpoint outside this region.

**Mapping of the internal mRNA endpoints in the \( \text{lacZ-Y} \) intercistronic region**

To map the endpoints of the Z and Y fragments we first compared the protected fragments obtained with the two probes, J39 and J40, which have a common promoter-proximal end, but differ by 14 bases at their distal ends (Fig. 1a). As expected the probe-length fragment (ZY) was slightly shorter for J40 than for J39 (Fig. 1b). The protected Y fragments obtained with J40 were also shorter than the fragments protected by J39, but the Z fragments were the same size for both probes. These results indicate that the 3' ends of the Z fragments and the 5' ends of the Y fragments map within the sequences carried by both hybridization probes. The other ends of the protected Z and Y fragments were generated by RNase T1 cleavage at a G residue (presumably the first) outside the lac insert in J39 and J40. The 3' ends of the Z fragments must thus be about 85 bases downstream of the first G upstream of the promoter proximal ends of J39 and J40. Interestingly, this position is at the 3' side of a potential hairpin between \( \text{lacZ} \) and \( \text{lacY} \) (Fig. 2). The 5' ends of the Y fragments must be about 250 bases upstream of the first G distal to the end of the J39 probe, i.e. close to the end of the \( \text{lacZ} \) gene. Our interpretation of the T1 mapping data for J40 are illustrated at the bottom of Fig. 1a.

The 5' endpoints of the Y fragments were mapped more precisely by primer extension. Unlabeled RNA prepared from cells containing the S10'/\( \text{lacZ} \)/lacT plasmid was annealed to a radioactive primer complementary to a sequence within the \( \text{lacY} \) mRNA and the primer was extended with reverse transcriptase at 45°C and 60 mM KCl. The products of this reaction were then analyzed on a urea polyacrylamide gel next to a sequencing ladder (Fig. 3, lane 8). A control reaction was done with RNA prepared from cells carrying the vector pSC101 in order to identify reverse transcription products due to unspecific priming (Fig. 3, lane 7). To determine which of the primer extension products were due to 5' ends of the Y fragments rather than spontaneous pausing or termination of the reverse transcriptase we performed a second control reaction using as templates a mixture of intact \( \text{in vitro} \) synthesized RNA and RNA isolated from the pSC101 containing strain (Fig. 3, lane 9). The results from these experiments exhibited a pair of bands, corresponding to positions at the end of the \( \text{lacZ} \) gene, which were only seen in the reaction using the S10'/\( \text{lacZ} \)/lacT mRNA as template (Fig. 3, lane 8). We therefore conclude that the Y fragments have 5' ends within the last codon and the termination codon of \( \text{lacZ} \) (Fig. 2). [Note that the products from the primer extension reaction run about 'one-half' base faster than the corresponding...

![Figure 2. Summary of RNA endpoints created by processing in the \( \text{lacZ-Y} \) intercistronic region. Shown is the sequence of the intercistronic region (32) with a potential hairpin (33). Termination codons of the \( \text{lacZ} \) gene and the initiation codon of \( \text{lacY} \) are underlined. Numbering is according to the lac operon file in GENBANK (Accession #s J01636, J01637, K01483, and K01793). The 5' endpoints of the processed Y transcripts (Fig. 1a) as determined by primer extension analysis (Fig. 3) are indicated by arrows. The region to which 3' endpoints of the processed Z transcripts have been mapped (see text) is indicated by the bracket under the sequence.](image-url)

![Figure 3. Identification of the 5' endpoints of the processed Y transcript by primer extension. A 5' end-labeled primer complementary to the proximal region of \( \text{lacY} \) transcript was annealed to unlabeled RNA isolated from NF314 containing either pFL1A49 (lanes 2, 5, 8) or the vector plasmid pSC101 (lanes 1, 4, 7). Primer extension was also performed using as template a mixture of RNA from the pSC101-containing strain and \( \text{in vitro} \) transcript from pR19 (lanes 3, 6, 9; for map of pR19, see Fig. 5a). The primer was extended with reverse transcriptase at 45°C (lanes 4–9) or 37°C (lanes 1–3) in the presence of 60 mM KCl (lanes 7–9) or 120 mM KCl (lanes 1–6). The figure shows an autoradiogram of a gel used to separate the reaction products. The solid arrows on the left indicate bands specific to the pFL1A49 lane. The open arrowhead points to the major stop in reverse transcription at high salt and/or low temperature. Markers were made by dideoxynucleotide sequencing using as template single-stranded DNA from the M13 phage J41 which carries the same fragment as J40 (Fig. 1a), but in the opposite orientation. The same primer was used as for the reverse transcriptase reactions, except that it was not phosphorylated at the 5' end. Sequences from the end of \( \text{lacZ} \) and the start of \( \text{lacY} \) are indicated on the right.](image-url)
products in the sequencing ladder, because the primer was 5' phosphorylated only for the former reaction.

Several stops of the reverse transcriptase were found at the promoter distal side of the potential hairpin between lacZ and lacY (Figs. 2 and 3, lane 8). Since these stops were also observed when using intact in vitro synthesized RNA as template (Fig. 3, lane 9) and did not correspond to any of the 5' ends observed in the RNase T1 mapping experiment, we suspected that they were due to the RNA hairpin blocking the migration of the reverse transcriptase as observed previously for other hairpins (34). This interpretation was confirmed when the primer extension reaction was performed at higher salt concentration (120 mM KCl; Fig. 3, lanes 2 and 3) or lower temperature (37°C; Fig. 3, lanes 5 and 6), conditions which are expected to stabilize the hairpin. In these cases, almost all of the reverse transcriptases stopped at the hairpin.

The mRNA fragments ending in the lacZ-Y intercistronic region are due to processing of the primary transcript

As discussed above ('Experimental strategy'), the Z and Y fragments shown in Fig. 1b are most likely due to processing of the primary transcript. To test this interpretation, the

![Figure 4](image-url)

**Figure 4. T1 mapping analysis of mRNA from the S10'/lacZ structural gene.** (a) Hybridization probes used are shown below the map of pLF1Δ49 (see legend to Fig. 1). Alternate shades of grey and black indicates different probes; see Table 1 for precise endpoints of probes. Note that M65 extends 414 base pairs further into the S10 operon than the portion carried by pLF1Δ49. The fragment of the S10'/lacZ fusion gene transcript protected by M65 is therefore shorter than the fragment of the chromosomal S10 operon transcript protected by the same probe. (b) RNase T1 mapping of RNA. RNA was prepared from cells containing pLF1Δ49 which had been labeled with 32P-PO4 for 5 minutes immediately before cell lysis. Radioactive RNA was hybridized to a mixture of probes (panel A) or to probe pJR15 alone (panel B). Protected fragments were analyzed by gel electrophoresis and displayed by autoradiography. In panel A, one hybridization reaction contained M65, M60, and M563 (mix I), the other hybridization reaction contained M65, M565 and J32 (mix II). We have previously shown that identical fragments are protected whether probes are used individually or in mixtures (27). Arrows indicate the probe-length fragments protected by each probe. As explained above, M65 protects a longer fragment (M65Q of the chromosomal S10 operon mRNA than it does from the plasmid derived transcript (M65P). Panel B shows parallel analysis of RNA labeled in vivo as described above and 35S-labeled RNA synthesized in vitro from pPZYA by SP6 RNA polymerase. The arrow labeled P indicates the probe-length fragments. The unlabeled arrows indicate the fragments created by RNase T1 cleavage during the protection assay (see text). Panel C shows an autoradiogram of a formaldehyde-agarose gel used to analyze the in vitro transcript made from pPZYA that was used for the experiment shown in panel B. The arrows labeled 2.9 kb and 1.5 kb indicate the positions of 23S and 16S rRNA, respectively. The open arrow heads point to the positions where in vitro transcription products should have been seen, if the two short fragments observed in the in vitro lane of panel B had been due to premature termination of the SP6 RNA polymerase during the in vitro transcription.
radioactivity in each of the bands in Fig. 1b was quantitated and plotted as a function of time. As seen in Fig. 1c, the radioactivity in probe-length fragments (ZY) decreased before the radioactivity in both the Z and Y fragments. This is consistent with the idea that the Z and Y fragments are due to processing of the primary transcript, since pools of processing products are replenished until the primary transcript has decayed. If the Z and Y fragments were due to transcription termination and/or an internal promoter, we would have expected that the radioactivity in the Z and Y fragments should decrease at least as early as the radioactivity in the probe length fragments. We conclude that the 3' end of the Z fragment and the 5' end of the Y fragment are due to mRNA processing. However, since the fragments overlap by about 30 bases they are not created by a single endonucleolytic break in the original intact transcript. Rather, the ends of the mRNA fragments mapped in these experiments appear to result from two different events.

Analysis of the S10'/lacZ gene

We also used probes from the S10'/lacZ structural gene (Fig. 4a) to search for mRNA fragments with intragenic endpoints. Whether used individually or in combination (Fig. 4b, panel A) probes M565, M563 and J32 never revealed anything but probe-length fragments. Thus no RNA fragments with endpoints in the distal two-thirds of lacZ' accumulated in detectable quantities. Probes M65 and M60 protected almost exclusively probe-length fragments, but in several experiments each probe also revealed two additional bands, barely above the level of detection (not shown). However, since these weak bands did not exhibit kinetics different from the probe-length fragments, we cannot confirm that they are degradation intermediates. In any event, the concentration of these RNA fragments is much lower than that of the Z and Y fragments discussed above.

Probe pJR15 generated mostly probe-length fragments, but also two smaller fragments (Fig. 4b, panel B) whose cumulative length equals the probe length. These smaller fragments were particularly interesting, because their lengths are consistent with an endonucleolytic cleavage at a potential hairpin previously shown to be cut by RNaseH in vitro (35). However, control experiments with in vitro synthesized RNA suggest that they are not the result of in vivo processing, but rather artifacts generated by T1 nuclease digestion. Specifically, T1 mapping of RNA synthesized in vitro from pPZYA also generated two additional bands smaller than the probe-length fragment (Fig. 4b, panel B), even though denaturing gel electrophoresis demonstrated the RNA was structurally intact before hybridization (Fig. 4b, panel C). One of these smaller fragments comigrates with one of the fragments from the in vivo RNA, whereas the other is approximately 15 bases shorter than the corresponding fragment from the in vivo RNA. This difference appears to be due to fortuitous homology between the S10' sequence next to the lacZ' junction in pLF1Δ49 and the vector sequence next to the lacZ insert in the pJR15 probe, so that pJR15 protects a slightly longer fragment from S10'/lacZ fusion mRNA than it does from the lacZ transcript. Consistent with this explanation, there is also a slight difference in the electrophoretic mobility of the probe-length fragments observed with the in vivo and in vitro RNA (Fig. 4b, panel B). These experiments therefore suggest that the shorter fragments of the in vivo RNA are not generated in the cell, but rather result from in vitro cleavage with RNase T1, possibly because a fraction of the RNA forms an imperfect RNA:DNA heteroduplex in which an intramolecular structure in the RNA is vulnerable to RNase T1 cleavage. We have previously observed that tRNA forms an imperfect RNA:DNA heteroduplex with extensive RNase T1 sensitive structure in the hybridized RNA (36).

We conclude from these experiments that degradation intermediates with endpoint within the lacZ gene are much less prevalent than the fragments with endpoints in the lacZ-Y intercistronic region, indicating that once degradation of S10'/lacZ structural message is initiated, complete destruction ensues very rapidly.

Degradation of genuine lac mRNA

To determine if transcripts from the genuine lac operon are subjected to the processing reactions observed with the S10'/lacZ lacY' fusion mRNA, we constructed plasmid pJR19 (Fig. 5a) which contains the intact lac operon inserted into pBR322. Induction of the lac promoter on pJR19 results in high levels

---

Figure 5. Analysis of processing in the lacZ-Y intercistronic region of mRNA from the intact lac operon. (a) Map of pJR19. The positions of the lac and SP6 promoters are indicated. The filled bar indicates the DNA fragment carried on hybridization probe J44. (b) RNase T1 mapping of lac transcripts from pJR19 synthesized in vivo or in vitro. A culture of pJR19/LL308 was labeled with 32P-PO4 for one-half generation prior to induction with IPTG. RNA was isolated immediately before and at the indicated times after the induction and hybridized to probe J44 (see Table 1 for precise endpoints). Hybridized RNA fragments protected from RNase T1 degradation were analyzed by gel electrophoresis and revealed by autoradiography. The right hand lane shows the results of RNase T1 analysis of 35S-labeled transcript synthesized in vitro from pJR19. RNA size markers are indicated on the right. ZY, Z and Y fragments of the lac operon transcript are indicated on the left.
of lac mRNA transcription, facilitating analysis of mRNA processing. To ensure that the lac transcripts were labeled with a constant specific radioactivity, the cells carrying pJR19 were prelabeled with $^{32}$P$_4$ for half a generation time before induction. RNA extracted at different times after induction was analyzed by T1 mapping as above, except that probe J44 (Fig. 5a) was used. As shown in Fig. 5b, we again observed three types of protected fragments from the region between the lacZ and lacY genes: the intact fragments corresponding to the entire length of the hybridization probe (ZY), several fragments with 3' ends about 30 bases downstream of the end of lacZ (Z), and a fragment whose 5' end is at the end of the lacZ gene (Y). Analysis of RNA from cells induced for a long time (15 min) indicate that about 70% of all accumulated transcripts hybridizing to J44 have 3' ends at position Z (data not shown).

To verify that the Z and Y fragments are not artifacts of the T1 mapping, we did a control hybridization with in vitro synthesized RNA. In this case a strong probe-length fragment was observed, but no Z or Y fragments were seen, indicating that the Z and Y bands seen with the in vivo labeled RNA are not artifacts of the nuclease protection protocol (Fig. 5b, right hand lane). We therefore conclude that the same processing intermediates are formed by the S10'/lacZ lacY fusion messenger and the genuine lac transcript.

The precursor-product relationship between the probe-length fragment and the shorter fragments was corroborated by the fact that labeling of the probe-length fragment (ZY) was first visible 2-2.5 minutes after induction, but the shorter fragments could not be observed until 3 or 3.5 minutes after induction (confirmed by longer exposures than the one shown in Fig. 5b). Note that the lack of radioactive bands for the first 2 minutes after induction corresponds to the time required for the RNA polymerase to transcribe at 30°C the approximately 3 kb between the promoter and the lacZ-lacY region (31).

One difference between the experiment with the S10'/lacZ fusion transcript and the intact lac transcript is that relatively smaller amounts of the Y fragments accumulate in the experiment with the intact lac message (compare Fig. 5b with Fig. 1b). This suggests that the relative rates of formation and degradation of the Y fragment differ for the two primary transcripts.

Since the experiments with the S10'/lacZ fusion mRNA suggested the possibility of minute amounts of degradation intermediates with endpoints within the M60 probe, we also investigated if intermediates could be detected in the transcript

![Figure 6](image_url)

**Figure 6.** RNase T1 analysis of lacZ mRNA from the intact lac operon. $^{32}$P$_4$ labeled RNA was prepared from a culture of pJR19/LL308, induced with IPTG for the indicated length of time, as described in the legend to Fig. 5, and subjected to RNase T1 mapping using M60 (Fig. 4a) as probe. Protected RNA fragments were analyzed by gel electrophoresis. On the same gel we also analyzed protected fragments obtained from in vitro synthesized lac RNA (Fig. 4b, panel C) which was added to an unlabeled culture immediately before cell lysis (lane A), to the unlabeled culture immediately after lysis (lane B), or to the unlabeled in vivo RNA after the first phenol extraction (lane C). The arrows indicate the approximate position of the weak bands observed when in vivo RNA from a pLFlA49 containing strain was RNase T1 mapped with M60.

![Figure 7](image_url)

**Figure 7.** Identification of 3' endpoints distal to lacA in RNA transcribed from the intact lac operon. (a) Map of the distal end of the lac operon. The filled bars indicate the DNA fragments used as hybridization probes. The arrows indicate the positions of RNA 3' endpoints determined as shown in (b). (b) RNase T1 mapping of lac transcripts from pJR19. RNA was labeled and hybridized as in the legend of Fig. 5b, except that the lac promoter was induced for 7 min and the J46 and J47 probes were used. Control hybridizations were done with intact $^{35}$S-labeled lac RNA synthesized in vitro. Fragments of lac mRNA with endpoints in the 3' non-translated region as well as probe-length fragments (PL) are indicated with arrows on the right. RNA size markers are indicated on the left. The probe-length fragment generated with J47 is a double band because of incomplete trimming with RNase T1.
from the intact lac operon. Indeed, weak bands running faster than the probe-length fragment were visible after over-exposure of the gel (Fig. 6, left panel). To determine if these weak bands represent in vivo degradation intermediates or artifacts due to limited fragmentation of the RNA during isolation or RNase T1 mapping, we did a control monitoring fragmentation of intact in vitro synthesized lac RNA added to in vivo samples. The radioactive in vitro RNA was added either to a non-radioactive culture just before or immediately after the cells were lysed, or to partially purified non-radioactive in vivo RNA after the first phenol extraction. In all cases, virtually all radioactivity was in the probe-length fragment. However, upon long exposure of the autoradiogram some weak bands smaller than the probe-length fragment were observed (Fig. 6, lanes A, B and C). These weak bands comigrated with fragments observed in the in vivo samples (Fig. 6, left panel). We therefore conclude that even though our RNA isolation procedure yielded mostly intact RNA, a limited amount of fragmentation at specific sites in the RNA might have taken place. Thus the small amount of fragments running faster than probe-length RNA in the in vivo experiment (Fig. 6, left panel) probably represents molecules fragmented during the isolation procedure or are artifacts of the RNase T1 hybridization assay, rather than intermediates of the in vivo degradation process. In any case, we did not observe the two smaller fragments seen in very low quantities with the transcript from the S10' '/lacZ plasmid (discussed above).

3' end formation in the lac operon transcript

There is evidence suggesting that some bacterial mRNAs undergo processing at the 3' end (1, 37). Since little attention has been given to 3' end formation of the lac messenger, we analyzed the distal end of the lac transcript. Cells carrying pJR19 were prelabeled with $^{32}$P-O$_4$ as above, induced with IPTG, and 7 minutes later RNA was extracted. The radioactive RNA was then subjected to RNase T1 mapping using two probes which have a common promoter-proximal end 58 bases upstream of the lacA gene termination codon, but differ at their distal ends. These ends are 324 bases (J46) and 268 bases (J47) downstream of the end of lacA (Fig. 7a). In addition to the probe-length fragments, both probes also protected two shorter fragments labeled A (about 100 bases long) and B (about 160 bases long) in Fig. 7b. Probe J47 also protected a 370 base fragment called C. None of these fragments were generated in a control hybridization with intact in vitro transcribed RNA (Fig. 7b), indicating that A, B, and C fragments are due to in vivo termination and/or processing. Since the same size A and B fragments are protected with both of the hybridization probes (Fig. 7b), the fragments must have 3' endpoints which map about 40 and 110 bases, respectively, downstream of the lacA gene. The C fragment, which is observed only with the J46 probe, must be derived from a transcript with 3' endpoint between the distal ends of the two probes, or about 310 bases from the end of lacA. Fig. 8 summarizes the endpoints found in the 3' untranslated region of the lac operon.

To determine if the fragments were due to transcription termination or to posttranscriptional processing, we hybridized RNA prepared at various times after induction of the transcription. The protected fragments corresponding to the full length of hybridization probe J47 can be observed 4 minutes after induction, but the A and B fragments were not seen until after 5 minutes (Fig. 9). It thus appears that most or all polymerase molecules transcribing the lac operon continue at least to the end of the J47 probe, or 268 bases beyond the end of lacA. This primary transcript is then shortened to create the transcripts with 3' ends corresponding to the A and B fragments. Furthermore, since the probe-length fragment generated with J46 is roughly 25% as intense as the J47 probe-length fragment, a substantial fraction of polymerases proceed more than 324 bases beyond lacA.

The processed transcripts ending at A and B (Fig. 8) quickly become much more intense than the probe-length fragment after
induction of the lac promoter (Fig. 9), suggesting that the shortening of the 3' end occurs very quickly after transcription termination. Analysis of RNA from cells induced for 15 min (data not shown) indicates that about 75% of accumulated transcript ends at the A endpoint about 40 bases downstream from the end of lacA.

**DISCUSSION**

We have used a combination of pulse labeling and fine structure mapping of RNA transcripts to analyze the processing of primary mRNA transcripts from the lac operon of E. coli. This method is extremely sensitive, because mRNA constitutes a large fraction of the radioactivity in pulse-labeled RNA. In contrast, nuclease S1 mapping, the most commonly used procedure for analyzing mRNA processing and degradation, is typically performed with unlabeled RNA in which total mRNA accounts for only a few percent.

We identified major processing intermediates with endpoints in two regions: the lacZ-Y intercistronic region and the untranslated region at the 3' end of the operon. A precursor-product relationship between primary transcripts and these fragments of the lac mRNA was verified by kinetic experiments. No major intermediates with endpoints in structural genes were observed. Clearly, such intermediates must exist, since mRNA is eventually degraded to mononucleotides, but we estimate that they are at least 10-fold less abundant than the Z fragments with a 3' end in the lacZ-Y intercistronic region. This suggests that intermediates ending within reading frames are very short-lived and therefore do not accumulate. Such a mechanism may have evolved to minimize synthesis of incomplete proteins which could both drain the cell's resources and interfere with the functioning of full-size proteins. (Note, however, that two regions of the lac operon were not examined: the extreme 5' end including the leader and the first 8 codons of the lacA gene, and approximately 1.5 kb from the lacY or lacA genes. Thus additional degradation intermediates with endpoints in these regions might exist.) Major mRNA processing sites have been identified within genes II/X and VII of bacteriophage fl (38, 39). However, only the mRNA fragments downstream of these processing sites accumulate in detectable concentrations. Thus even in this case there appears to be no significant amounts of mRNA which could be translated into incomplete proteins.

Processing of genuine lac mRNA and of the artificial S10'/lacZ fusion gene transcript resulted in accumulation of intermediates with identical endpoints within the lacZ-Y intercistronic region. Thus, the processing pathway for this portion of the message is independent of the 5' and 3' ends of the primary transcript. However, the relative concentration of the Y fragments with 5' ends within the translation termination region of lacZ is much lower for the genuine lac transcript than for the S10'/lacZ transcript (compare Figs. 1b and 3b). Apparently, the relative rates of formation and degradation of the Y intermediates differ for the two primary transcripts. This may be due e.g. to differences in translation efficiencies of the lacZ sequences derived from the two plasmids, or to differences in the 3' portion of the mRNAs (transcription on pL1Δ49 continues into the pSCL101 vector sequences).

The processing events generating the internal mRNA endpoints we have observed can only be partially defined at this point. Our finding of the Z and Y fragments at first seem to support the model that degradation of the lac operon transcript involves endonucleolytic cleavage between the two genes (40). However, since they overlap by about 30 bases, the lacZ and lacY mRNA fragments could not result from a single endonucleolytic event. The Y fragments could be the result of degradation of the lacZ part of the primary transcript, since earlier experiments (16) suggested that this process proceeds in an overall 5' to 3' direction. Alternatively, the Y fragments might result from an endonucleolytic cleavage severing the lacZ and lacY portions of the transcript followed by rapid degradation of the lacZ cleavage product. It should be noted that the kinetics of disappearance upon rifampicin arrest of transcription initiation (Fig. 1c) and the labeling kinetics upon induction of the lac promoter (Fig. 5b) both rule out that the Y fragments are due to a weak internal promoter.

Nierlich and coworkers found a 3' end in the lacZ-Y intercistronic region at the location of the end of our Z fragment and, furthermore, showed that the hairpin immediately upstream of the 3' ends of the Z fragments functions as a transcription terminator in vitro, but not in vivo (33, 41). We cannot exclude that some of the Z fragments are created by transcription termination, but two results suggest that RNA processing is a major pathway for their generation. One is that, upon induction of the lac promoter, the Z fragments are not visible until 30 seconds after the primary transcript is observed (Fig. 5b). The other is that the primary (probe-length) transcripts disappear before the Z fragments upon rifampicin arrest of transcription initiation. We have no evidence as to whether the Z fragments are created by an endo- or exonuclease. However, since the 3' ends are distal to a hairpin it is conceivable that the fragments are created when a 3' to 5' exonuclease stalks at the hairpin structure.

We observed three endpoints in the 3' untranslated region of the lac operon mRNA. The existence of one of these (B, Fig. 8) was previously suggested by S1 mapping (unpublished experiments cited in ref. 25). Our labeling kinetics (Fig. 9) show that transcription continues at least 268 bases beyond the end of lacA; the primary transcripts are then trimmed back posttranscriptionally to generate the A and B endpoints (Fig. 8). A similar model has been inferred from genetic and in vitro transcription data for the trp operon transcript (36). The mapping of the A and B endpoints to descending portions of potential hairpins suggest that they could be generated by exonucleases stalling at these hairpins. Similar mRNA stabilizing hairpins have been demonstrated at the ends of several other transcription units (1, 15, 42).

Kennell and coworkers (17, 18) previously reported the existence of multiple 5' and 3' endpoints both within the lacZ and Y genes and in the lacZ-Y intercistronic region. Even though one of our Y fragments corresponds to one of the 5' ends reported by these workers, our pattern of a few distinct processing products is strikingly different from their pattern of numerous endpoints present in similar concentrations. In spite of performing many experiments with different labeling conditions, we never observed most of the endpoints seen by Kennell and coworkers. We have no conclusive explanation for this discrepancy, but several factors should be discussed. First, it is unlikely that our pattern of mRNA degradation intermediates is due to the high rate of lac mRNA synthesis caused by the presence of lac on a multicopy plasmid. We have analyzed RNA from a strain with a single copy lac operon and found the same major endpoints in the lacZ-Y intercistronic region as reported here (not shown). In addition, Murukawa et al. (41) also analyzed lac mRNA from a strain with
ACKNOWLEDGMENTS

We thank R. Guadino and E. A. Morgan (Roswell Cancer Institute, Buffalo, NY) for the pPZYA plasmid. This investigation was supported, in part, by a predoctoral training grant from the National Institute for Allergy and Infectious Diseases. JRM was supported, in part, by the procedures used for RNA preparation should be considered. Extensive controls were presented both by Kennell's group (17,43,44) and in this paper. Nevertheless, we do not believe that it can be excluded that the large collection of bands observed throughout the lacZ-Y intercistronic region (17,18) could be due to limited degradation during RNA purification. Kennell's group proposed that any low level RNA cleavage would happen at random sites and not result in specific bands observed in an S1 analysis (43). However, it is possible that mRNA secondary and tertiary structures result in sites highly exposed to ‘unspecific’ nucleases during RNA purification. Finally, cell physiology might affect the accumulation of partially degraded messengers. Thus it is possible that the different patterns of mRNA fragments seen by us and by Kennell’s group could be due to growth conditions or the strains used.

In summary, our experiments have identified major sites of cleavage in the lac mRNA and proven posttranscriptional processing of the 3’ non-translated end of the lac message.

REFERENCES