In vivo and in vitro detection of the leader RNA of the histidine operon of Escherichia coli K-12

(attenuation/leader RNA/minicells)

RODOLFO FRUNZIO, CARMELO B. BRUNI, AND FRANCESCO BLASI*

Centro di Endocrinologia ed Oncologia Sperimentale del Consiglio Nazionale delle Ricerche and Cattedra di Microbiologia, Instituto di Patologia Generale, University of Naples, 2nd Medical School, Via S. Pansini 5, 80131 Naples, Italy

Communicated by Maxine F. Singer, January 27, 1980

The DNA of the attenuator region of the histidine ABSTRACT operon of Escherichia coli has been transcribed in a purified in vitro system and found to synthesize two major RNA transcripts. The first one, 180 nucleotides long, has been identified as the histidine-specific leader RNA. It contains the coding sequence for the leader peptide [Di Nocera, P. P., Blasi, F., Di Lauro, R., Frunzio, R. & Bruni, C. B. (1978) Proc. Natl. Acad. Sci. USA 75, 4276-4280] and is terminated at the attenuator site. Termination of transcription at this site is extremely efficient in the in vitro system. The leader RNA also has been detected in vivo in a minicell producer strain transformed with plasmids harboring the regulatory region of the histidine operon of E. coli. A second RNA molecule is synthesized in the in vitro system. It has a divergent direction of transcription with respect to the histidine leader RNA, but its role, if any, in the regulation of the histidine operon remains to be ascertained. The existence of the histidine leader RNA lends support to the regulatory mechanism which postulates that regulation of the histidine operon is dependent on the alternative secondary structures that the leader RNA may assume, depending on whether or not the histidine-rich leader peptide is translated.

Expression of the *his* operon in bacteria is regulated at the level of transcription (1-3). The levels of *his* mRNA increase in regulatory mutants and depend on the structure and the levels of charged histidyl-tRNA^{His} (4, 5). Regulation is exerted at a region of DNA upstream of the structural genes at two different sites: a promoter site (6) and a transcriptional barrier called the attenuator (1, 7, 8). Histidine-specific regulation seems to occur mostly at the attenuator site (6).

Transcription termination, first discovered in λ phage (9), is a major regulatory mechanism in prokaryotes. The expression of operons for the biosynthesis of amino acids like *trp* (10), *his* (1, 2, 7, 8, 11), *pheA* (12), *leu* (13), *thr* (14), and *ilv* (15, 16) is regulated by transcription termination at the attenuator barrier. Active synthesis and translation of a small mRNA called leader RNA is an essential part of this regulatory mechanism (2, 11, 17, 18).

Although *his* was the first system in which an attenuator was described (1), no direct proof of a leader RNA has been reported. The transcription termination mechanism is supported by the nucleotide sequence of a region of DNA upstream of the first structural gene *hisG* (7, 8) and of a deletion mutant defective in transcription termination (11).

In this paper we present direct evidence for the synthesis of a leader RNA of the *his* operon of *Escherichia coli* both *in vivo* and *in vitro*. Moreover, a transcription map of the *his* promoter region is presented, which also shows an additional transcription initiation signal about 200 base pairs (bp) upstream of the his promoter. This initiation site is very efficient in vitro and transcribes in an orientation opposite to that of the his operon.

MATERIALS AND METHODS

DNAs and Enzymes. DNA fragments used for *in vitro* transcription (see Fig. 1) have been isolated by digesting plasmid pCB3 DNA (19) with appropriate restriction endonucleases, separating fragments by gel electrophoresis on 1% agarose or 6% (wt/vol) acrylamide gels, and recovering DNA fragments from the gel (20).

E. coli RNA polymerase was from Miles or Boehringer Mannheim. Restriction endonucleases were purified as described (21) or were obtained from commercial sources.

In Vitro Transcription. Individual DNA fragments were transcribed in vitro as described (22) by using RNA polymerase (100–200 μ g/ml), DNA (20–30 μ g/ml), 100 mM KCl, heparin (100 μ g/ml), and 0.2 mM nucleoside triphosphates. [α -³²P]Triphosphates (0.04 mM) were used; specific activity of each was about 300–400 Ci/mmol (1 Ci = 3.7 × 10¹⁰ becquerels). The RNA synthesis mixture was extracted with phenol, precipitated with ethanol, and analyzed by 6% acrylamide/7 M urea gel electrophoresis (20). The gel was exposed to Kodak X-Omat R x-ray film. The RNA from individual bands was recovered as described (20).

Two-Dimensional Chromatography and Sequence Determination of RNA. Ribonuclease T1 treatment was followed by separation of oligonucleotides by high-voltage electrophoresis in the first dimension and by homochromatography on DEAEcellulose plates in the second dimension (23).

The 5' end of RNA was labeled by using $[\gamma^{-32}P]ATP$ (500 Ci/ mmol) as a substrate during *in vitro* transcription (24). The nucleotide sequence of the 5'-labeled RNA was determined as described (25). The enzymatic reactions specific for the four nucleotides with T1, U2, A, and *Physarium* ribonucleases used incubation times, buffer systems, and enzyme/RNA ratios as published (26). Reactions (5 μ l) were stopped with 5 μ l of deionized formamide and dye markers (bromophenol blue and xylene cyanol) and were heated for 2 min at 80°C. The samples were loaded on a 25% acrylamide/8 M urea slab gel (20), electrophoresed for 45 min at 2000 V, and autoradiographed.

Base composition of individual ³²P-labeled oligonucleotides from ribonuclease T1 treatment was determined after ribonuclease T2 digestion; individual nucleotides were separated by high-voltage paper electrophoresis (23), and the corresponding areas were cut and assayed for radioactivity.

In Vivo Labeling of the his Leader RNA. The E. coli K-12 strain DS998, a minicell producer provided by G. Cesareni

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: bp, base pairs.

^{*} Present address: Institute of Genetics and Biophysics of Consiglio Nazionale delle Ricerche, Via Marconi 10, Naples, Italy.

(Cambridge, England), was transformed with different plasmids. Minicells were isolated (27) and resuspended in a phosphate-free medium (28) containing 0.5% glucose and 0.05 mM amino acids. Minicells were washed twice in this medium and incubated 15 min at 37°C to deplete residual phosphate; then 1-2 mCi of $H_3^{32}PO_4$ was added, and the cells were incubated for 30 min at 37°C. RNA was immediately extracted with phenol, precipitated with ethanol, and electrophoresed on a 6% polyacrylamide/7 M urea slab gel. RNA bands were visualized by autoradiography, recovered from the gel (20), treated with ribonuclease T1, and analyzed by two-dimensional chromatography (23).

RESULTS

DNA Fragments Used As Templates for in Vitro Transcription. All DNA fragments were isolated from plasmid pCB3, a recombinant plasmid carrying the first two genes of the *E*. coli K-12 his operon, its regulatory region, and about 1900 bp (to the left) of unknown function (19). Fig. 1 shows the DNA fragments that have been used as templates for in vitro transcription. The HindIII 5300-bp fragment represents the entire DNA insert of pCB3 (19) and was obtained by HindIII digestion of pCB3 DNA. The HinfI 730-bp fragment was obtained by digestion of pCB3 DNA, and the 390-bp fragment Alu I was obtained by digestion of the HinfI fragment. Both the HinfI and Alu I fragments end within the first structural gene hisG and contain the attenuator region of the his operon (7, 19).

In Vitro Transcription. When the HindIII 5300-bp fragment was used as template, a single RNA band of low M_r could be demonstrated on the electrophoresis gel (Fig. 2, lane c). In addition there were high M_r RNA transcripts that were not resolved at the top of the gel. The small-size band (L) was about 180 nucleotides long, as estimated by comparison with the electrophoretic mobility of markers of known length.

In order to locate the promoter for the L RNA, smaller size fragments were transcribed *in vitro*. When the *Hin*fI 730-bp fragment (see Fig. 1) was used as template, two short major transcripts were visible on the gel. One comigrated with the L band, whereas a new one (the A band) had a mobility corresponding to a size of ≈ 230 nucleotides (Fig. 2, lane a). The more prominent of the high M_r transcripts was approximately 400 bases long and was often present in the transcription pattern, its intensity being variable in different experiments. It was not further characterized. When the subfragment of the *Hin*fI 730-bp fragment—the Alu I 390-bp fragment—was used as a template, the L RNA but not the A RNA was made (Fig. 2, lane b). The higher M_r band was extremely heparin sensitive (unpublished data) and must be regarded as being due to a spurious initiation. In the presence of heparin, the L RNA band comprised more than 90% of the incorporated radioactivity.

Several criteria indicated that the L and A transcripts were initiated from different promoters. At very low nucleoside triphosphate concentration (2 μ M), the L band was absent, whereas the A band was still made quite efficiently. At low RNA polymerase:DNA ratios, the L band was made more efficiently, whereas the reverse was true at high salt concentration (200 mM KCl). Finally, Fig. 3 shows two-dimensional chromatographs after ribonuclease T1 digestion of the L and A RNAs. The number of spots and the mobility of the different labeled oligonucleotides indicated that the nucleotide sequence of the two transcripts was different.

Cleavage of the 730-bp *Hin*fI fragment with different restriction endonucleases was used to orient the direction of transcription of the L and A RNAs. Digestion of the *Hin*fI 730-bp DNA with *Apy* I had no effect on the transcription pattern. Both L and A RNAs were still made (compare lanes a and b of Fig. 4). On the other hand, digestion of the *Hin*fI 730-bp DNA with *Hpa* II endonuclease decreased the size of both transcripts (Fig. 4, lane c). *Apy* I cuts in the middle of the *Hin*fI 730-bp DNA, whereas *Hpa* II cuts at the two ends (Fig. 1). *Hpa* II shortens by 20 bp the left end of the *Hin*fI fragment. Another *Hpa* II restriction site is located 30 bp before the end of the attenuator (Fig. 1). The L and A transcripts were shortened by the number of nucleotides predicted from the position of the *Hpa* II restriction sites (Fig. 4). These results indicate that transcription is initiated at the two sides of the *Apy* I site and proceeds in the



FIG. 1. Transcription map of the *his* operon regulatory region. The top line represents the physical map of the *Hin*dIII 5300-bp DNA. The middle line shows the 730-bp *Hin*fl fragment used for most of the *in vitro* transcription experiments with some of the pertinent restriction sites. The *Alu* I 390-bp fragment is shown in heavy black. The promoters, *his* attenuator (Atn), and *hisG* structural gene are represented by empty boxes. Arrows point to the direction of transcription and indicate the length of the two transcripts: the *his* leader RNA and the A RNA. The bottom part of the figure shows the DNA sequence of the antisense strand and the 5' termini of the *his* leader RNA and of the A RNA (unpublished data). The -35 TTG and the Pribnow box of the *his* promoter are underlined. Numbers on top of the DNA sequence are in progression from the left end of the *Hin*fl 730-bp fragment.



FIG. 2. In vitro transcription pattern of different DNA fragments. Lanes: a, RNA transcribed from the Hinfl 730-bp fragment analyzed on a 6% denaturing acrylamide slab gel (L is the his leader RNA, about 180 bases long; the A RNA is about 230 bases long); b, RNA transcribed from the Alu I 390-bp fragment (heparin was omitted from the reaction mixture); c, RNA transcribed from the HindIII 5300-bp fragment. Markers used to estimate the length of the transcripts were: ϕ X174 Hae III DNA fragments, E. coli 5S and 6S RNA, and λ phage 4S and 6S in vitro transcripts.

two opposite directions. A transcription map is in Fig. 1. The L RNA has the right orientation to be the *his* leader RNA. It is transcribed in the direction of the first structural gene. The size decrease caused by digestion of the template with *Hpa* II, about 30 nucleotides, is in agreement with the L RNA being terminated at the attenuator site.

The following data indicate that the A RNA is a runoff transcript. Its size, 230 bases, coincides with the distance between the left end of the *Hin*fl 730-bp fragment and the site where transcription initiation starts (Fig. 1). When other DNA fragments were transcribed, also carrying the initiation site for the A RNA but extending for different lengths downstream, the size of the reulting RNA was increased (unpublished data).

Sequence of 5' and 3' Termini. The 5'-terminal sequence of the L and A RNAs made *in vitro* was analyzed on RNA molecules terminally labeled *in vitro* with $[\gamma^{32}P]ATP$. The bands were excised from the gel and the RNA was eluted and sequenced as described (25). The sequence of the first 14 bases of the L RNA (pppA-U-C-A-G-U-U-G-A-A-U-A-A-A) and of the first 11 bases of the A RNA (pppA-C-U-G-U-A-C-A-A-A) were determined unambiguously (unpublished data). The initiation sites thus identified are shown in Fig. 1 with the DNA sequence of the promoter regions (unpublished data).

The 3' end of the *in vitro*-made L RNA was determined as follows: L RNA labeled either with $[\alpha^{-32}P]$ GTP or with all four nucleoside triphosphates was treated with ribonuclease T1 and analyzed by two-dimensional chromatography (Fig. 5). The oligonucleotide corresponding to the 3' end of the L RNA does not contain any guanosine; therefore, it should be present in the two-dimensional chromatograph of the uniformly labeled RNA and absent in those of the guaosine-labeled RNA. The arrow in Fig. 5B indicates this oligonucleotide. Its base composition (A₁C₁U₆) fitted the sequence of the 3' end of the *his* attenuator (C-T-T-T-T-A-T) (ref. 7; unpublished data). Positioning the 5' and 3' ends of the L RNA on the DNA sequence (ref. 7; unpublished data) allowed the determination of the exact length of the RNA (180 nucleotides).

The his Leader RNA in Vivo. An RNA molecule corresponding to the L RNA made *in vitro* can be isolated from minicells transformed with plasmid pCB3 (19), which carries the entire *his* regulatory region and the first two structural genes. Fig. 6 shows the electrophoretic pattern of the RNA bands obtained from minicells incubated with carrier-free ³²P in a phosphatefree medium. Lane c shows the L and A RNA markers made *in vitro*; lane b shows the RNAs obtained from minicells transformed with plasmid pCB3. A band which comigrates with the L RNA was present. Lane a shows the RNAs obtained from



FIG. 3. Comparison of the twodimensional chromatographs after ribonuclease T1 digestion of the L and A RNA transcripts. (A) The his leader L RNA. (B) The 230-base A RNA. Horizontal arrows point to the direction of the first dimension (electrophoresis); vertical arrows indicate the direction of the second dimension (homochromatography).



FIG. 5. Two-dimensional chromatographs after ribonuclease T1 digestion of the *his* leader RNA synthesized *in vitro* and labeled with $[\alpha^{-32}P]$ GTP (A) or with all four nucleoside $[\alpha^{-32}P]$ triphosphates (B). The arrow in B points to the 3' oligonucleotide that is missing in the guanosine-labeled RNA (A).

transcribed *in vitro* with E. *coli* RNA polymerase. Two observed heparin-resistant transcripts, the A and L RNA, were initiated at different promoters (Fig. 2). These transcripts were oriented

FIG. 4. In vitro transcription of the 730-bp Hinfl fragment digested with different restriction endonucleases. Lanes: a, transcription pattern of the 730-bp Hinfl fragment as such; b, Hinfl fragment redigested with Apy I, purified by phenol extraction and ethanol precipitation, analyzed on a 6% acrylamide gel for completeness of digestion, and subjected to transcription in the purified *in vitro* system; c, Hinfl fragment redigested with Hpa II, treated as described above, and transcribed.

minicells transformed with plasmid pCB5, a plasmid similar to pCB3 except that the region of the *his* promoter is absent and the *his* genes are transcribed under the control of the tetracycline-resistance promoter (7, 19). As expected, no RNA band comigrating with the L RNA was present. In both pCB3- and pCB5-transformed minicells, the A RNA band was not present, in agreement with the data which showed that this transcript was not terminated *in vitro*.

The identity of the band made *in vivo* (Fig. 6, lane b) with the L RNA has been demonstrated by T1 ribonuclease treatment and two-dimensional chromatographic analysis (Fig. 7). The two two-dimensional chromatographs are essentially identical with respect to both the number and position of the oligonucleotides.

DISCUSSION

Previous studies (1, 2) and the nucleotide sequences of the regulatory region of the wild-type (7, 8) and mutant (11) his operon have indicated that in this system specific regulation is exerted at the level of transcription at the attenuator site. The necessary corollary, the existence of a his leader RNA, has not been demonstrated. The synthesis of a leader RNA both in vitro and in vivo has been proven only for the trp operon (17). In this work we identify the his leader RNA and report a transcription map of the regulatory region of the his operon of *E. coli* K-12 (Fig. 1). The region coding for the his leader RNA was identified by in vitro transcription. DNA fragments of several lengths, located upstream of the 5' end of the first structural gene of the his operon and containing the attenuator sequence (7), were



FIG. 6. In vivo RNA transcription in minicells harboring histidine plasmids pCB5 (lane a) or pCB3 (lane b) was analyzed on a 6% denaturing slab gel. Lane c shows RNA transcribed *in vitro* from the *Hinf*I 730-bp fragment.



FIG. 7. Two-dimensional chromatographic analysis after ribonuclease T1 digestion of the *his* leader RNA. (A) The 180-base *his* leader RNA synthesized *in vitro*. (B) The RNA synthesized *in vivo* in minicells and comigrating on 6% acrylamide slab gel with the *his* leader RNA. Horizontal arrows, first dimension; vertical arrows, second dimension.

on the DNA template by transcription of differently restricted DNA fragments (Fig. 4). The nucleotide sequences of the 5' and 3' ends of the L RNA were determined and matched with the DNA sequence of the *his* promoter region (unpublished data; Fig. 1). These data demonstrate that the L RNA is a 180-base-long leader RNA initiated at the *his* promoter (Fig. 1) and terminated at the attenuator, upstream of the first structural gene of the *his* operon. The precise termination site is the run of thy-midine residues located at the descending part of the attenuator stem. *E. coli* minicells containing plasmids carrying the *his* regulatory region also synthesize the *his* leader RNA.

The actual existence of the his leader RNA confirms the regulatory model predicted from in vitro transcription and translation experiments (1, 2) and from the nucleotide sequence of the regulatory region of the his (7, 8, 11) and other biosynthetic operons (10, 29, 30). The mechanism of transcription termination at the his attenuator has not been analyzed before. The his attenuator presents the most extreme features (7, 8) of a typical attenuator (31). In vivo it has been shown that the basal level of his mRNA is only 5% of that of mutants deleted of the attenuator (1). In this paper we show that transcription termination is efficient also in vitro. Alu I 390-bp fragment contains only one transcription initiation site, the his promoter. In the presence of heparin, the his leader RNA represents at least 90% of the total transcription initiation products. Whether any termination factors are involved in vivo is still in doubt. rho mutations have been reported to have either no effect (32) or a slight effect (33) on the expression of the his operon.

We also have demonstrated that a second RNA molecule (the A RNA) is transcribed *in vitro* from the *his* regulatory region in a direction opposite to the *his* leader RNA. This RNA is initiated very efficiently, even at low nucleotide triphosphate concentrations and is not terminated, at least within a few hundred bp from the initiation site. No genetic markers are known in the immediate vicinity of the *his* operon on the *E. coli* K-12 chromosome (34). Therefore, we have no information on the possible

role of this RNA in the physiology of *E*. *coli*. Nor is any evidence available indicating a role of the A RNA on the regulation of the expression of the *his* operon. The presence of two transcription initiation sites, less than 200 bases apart, might be a useful system to study some aspects of the mechanism and control of the RNA chain initiation selection.

We wish to thank G. G. Brownlee, F. Baralle, and P. Arcari for expert advice with the RNA techniques for sequence determination; our colleagues P. P. Di Nocera and P. Verde for performing the DNA sequence determinations; and M. Rosenberg and M. M. Rechler for suggestions and careful reading of the manuscript.

- 1. Kasai, T. (1974) Nature (London) 249, 523-527.
- Artz, S. W. & Broach, J. R. (1975) Proc. Natl. Acad. Sci. USA 72, 3453–3457.
- Bruni, C. B., Colantuoni, V., Sbordone, L., Cortese, R. & Blasi, F. (1977) J. Bacteriol. 130, 4-10.
- Singer, C. E., Smith, G. R., Cortese, R. & Ames, B. N. (1972) Nature (London) New Biol. 238, 72-74.
- 5. Lewis, G. A. & Ames, B. N. (1972) J. Mol. Biol. 66, 131–142.
- Winkler, M. E., Roth, D. J. & Hartman, P. E. (1978) J. Bacteriol. 133, 830–843.
- Di Nocera, P. P., Blasi, F., Di Lauro, R., Frunzio, R. & Bruni, C. B. (1978) Proc. Natl. Acad. Sci. USA 75, 4276–4280.
- 8. Barnes, W. M. (1978) Proc. Natl. Acad. Sci. USA 75, 4281-4285.
- 9. Roberts, J. W. (1969) Nature (London) 224, 1168-1174.
- Oxender, D. L., Zurawsky, G. & Yanofsky, C. (1979) Proc. Natl. Acad. Sci. USA 76, 5524–5528.
- 11. Johnston, M. H., Barnes, W. M., Chumley, F. G., Bossi, L. & Roth, J. R. (1980) Proc. Natl. Acad. Sci. USA 77, 508-512.
- Zurawsky, G., Brown, K., Killingly, D. & Yanofsky, C. (1978) Proc. Natl. Acad. Sci. USA 75, 4271–4275.
- Gemmill, R. M., Wessler, S. R., Keller, E. B. & Calvo, J. M. (1979) Proc. Natl. Acad. Sci. USA 76, 4941–4945.
- 14. Gardner, J. F. (1979) Proc. Natl. Acad. Sci. USA 76, 1706-1710.
- Lawther, R. P. & Hatfield, G. W. (1980) Proc. Natl. Acad. Sci. USA 77, 1862–1866.
- Nargang, F. E., Subrahmanyam, C. S. & Umbarger, H. E. (1980) Proc. Natl. Acad. Sci. USA 77, 1823–1827.
- 17. Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, G. C., Squires, C. & Yanofsky, C. (1975) Science 189, 22-26.
- Zurawsky, G., Elseviers, D., Stauffer, G. V. & Yanofsky, C. (1978) Proc. Natl. Acad. Sci. USA 75, 5988–5992.
- Bruni, C. B., Musti, A. M., Frunzio, R. & Blasi, F. (1980) J. Bacteriol. 142, 32–42.
- Maxam, A. M. & Gilbert, W. (1979) Methods Enzymol. 65, 499-560.
- Bickle, T. A., Pirrotta, V. & Imber, R. (1977) Nucleic Acids Res. 4, 2561-2572.
- Howard, B. H., de Crombrugghe, B. & Rosenberg, M. (1977) Nucleic Acids Res. 4, 827-842.
- Brownlee, G. G. (1972) in Determination of Sequences in RNA, eds. Work, T. S. & Work, E. (North-Holland, Amsterdam), pp. 70-78.
- 24. Miller, J. S. & Burgess, R. R. (1978) Biochemistry 17, 2054-2059.
- Simoncsitis, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. & Guilley, H. (1977) Nature (London) 269, 833–836.
- Horii, T., Ogawa, T. & Ogawa, H. (1980) Proc. Natl. Acad. Sci. USA 77, 313–317.
- 27. Hallewell, R. A. & Sherratt, D. J. (1976) Mol. Gen. Genet. 146, 239–245.
- 28. Bertrand, K., Squires, C. & Yanofsky, C. (1976) J. Mol. Biol. 103, 319-337.
- Keller, E. B. & Calvo, J. M. (1979) Proc. Natl. Acad. Sci. USA 76, 6186–6190.
- Blasi, F. & Bruni, C. B. (1981) in Current Topics in Cellular Regulation, eds. Horecker, B. L. & Stadtman, E. R. (Academic, New York), Vol. 19, in press.
- 31. Rosenberg, M. & Court, D. (1979) Annu. Rev. Genet. 13, 319-353.
- 32. Winkler, M. E. (1978) J. Bacteriol. 135, 721-725.
- Lawther, R. P. & Hatfield, G. W. (1978) J. Bacteriol. 136, 1201-1204.
- Bachman, B. J., Low, K. B. & Taylor, A. L. (1976) Bacteriol. Rev. 40, 116–167.