Nucleotide sequence and biochemical characterization of the *metJ* gene from *Salmonella typhimurium* LT2

Mark L. Urbanowski and George V. Stauffer

Department of Microbiology, University of Iowa, Iowa City, IA 52242, USA

Received 10 December 1984; Accepted 12 December 1984

ABSTRACT

The nucleotide sequence of the *Salmonella typhimurium metJ* gene is presented along with the sequence of the promoter region for the closely linked *metB* gene. The two genes are transcribed in opposite directions, with transcription initiating from a single promoter for *metB*, and from two apparent promoters for *metJ*. RNA polymerase binding sites for *metJ* and *metB*, determined by in vitro protection studies, lie adjacent to each other and may overlap. The two *metJ* promoters, P_{J1} and P_{J2}, are separated by approximately 65 base pairs. Binding of RNA polymerase in vitro could only be observed for P_{J1}, even though transcripts are initiated from both promoters in vivo. The *metJ* gene codes for a polypeptide of 105 amino acids with a calculated Mr of 12,110. The translation start site was determined by N-terminal amino acid sequence analysis of a *metJ-lacZ* fusion protein.

INTRODUCTION

The *metJ* gene in *Salmonella typhimurium* and *Escherichia coli* has been shown to code for a protein involved in the regulation of the methionine biosynthetic pathway, possibly acting as a classical repressor (1-3, 17). We have recently reported the cloning of this gene and the closely linked *metB* gene of *S. typhimurium* (4). Here we present the complete nucleotide sequence of the *metJ* gene and a biochemical characterization of its control region and the control region of the *metB* gene.

MATERIALS AND METHODS**Bacteria and plasmids**

The *S. typhimurium* wild type strain JL781 was used for preparing total cellular RNA. Plasmid pGS107 carries the *S. typhimurium metJ* and *metB* genes and has been described (4).

Plasmid isolation

Plasmid DNA was prepared as described previously (4).

DNA sequence analysis

DNA sequencing methods were those of Maxam and Gilbert (5). Gel elec-

trophoresis was carried out according to Sanger and Coulson (13).

5' S1 nuclease mapping

The S1 nuclease mapping procedure of Weaver and Weissman (6) was used. A 396 base pairs (bp) *Mlu*I fragment, which includes the *metJ* and *metB* transcription initiation regions, was labeled at the 5' ends with γ -³²P-ATP using T₄ polynucleotide kinase (14). The labeled strands were separated (5) and aliquots of each strand precipitated with 50 μ g of total cellular RNA isolated from either JL781 grown in Luria broth, or JL781 carrying pGS107 grown either in glucose minimal medium or glucose minimal medium plus methionine. The mixtures were resuspended in 20 μ l of hybridization buffer (0.4 M NaCl, 0.04 M PIPES, pH 6.4, 1 mM EDTA), heated to 90°C for 2 minutes, and then allowed to hybridize at 55°C for 1 hour. Two hundred μ l of cold S1 nuclease buffer (0.25 M NaCl, 0.06 M Na acetate, pH 4.6, 1 mM ZnSO₄, and 5% glycerol) containing 300 units of S1 nuclease was added and digestion was carried out at 20°C for 1 hour. The S1 nuclease-resistant products were then precipitated and run alongside a sequencing ladder of the same DNA strand.

3' S1 nuclease mapping

The procedure used to map the 3' terminus of the *metJ* transcript was similar to the 5' S1 nuclease mapping given above, with the following modifications. Suitable fragments containing the termination region of the *metJ* gene could not be separated into single strands. Therefore, a double stranded 621 bp *Mlu*I-*Cla*I fragment labeled at the 3' *Mlu*I end using α -³²P-dCTP and the large fragment of *E. coli* polymerase I (14) was used as a probe. The labeled DNA was precipitated along with 50 μ g of total cellular RNA isolated from JL781 carrying plasmid pGS107 and grown in glucose minimal medium. The mixture was resuspended in 20 μ l of an 80% formamide hybridization buffer, heated to 90°C for 2 minutes, and allowed to hybridize 4 hours at 60°C. The sample was then treated as in the 5' mapping procedure above.

RNA polymerase binding sites

The RNA polymerase binding sites for the *metJ* and *metB* genes were determined using the DNase I "footprinting" procedure of Schmitz and Galas (8). The 396 bp *Mlu*I fragment carrying the *metJ* and *metB* control regions was labeled at a single 3' end using α -³²P-dCTP and the large fragment of *E. coli* polymerase I. The fragment was incubated 15 minutes at 37°C in 100 μ l of binding buffer (10 mM Tris-HCl, pH 7.9, 0.1 mM ATP and GTP, 125 mM KCl, 10 mM MgCl₂, 10 mM CaCl₂, 0.1 mM dithiothreitol, and 25% glycerol), and then split into two samples. To one sample 7 μ g of RNA polymerase holoenzyme was added, and the incubation continued for 10 minutes at 37°C.

DNase I was then added to both tubes to give a final concentration of 0.13 µg/ml. The DNase I reaction was stopped after 30 seconds by the addition of 25 µl of a 3 M ammonium acetate and 0.25 M EDTA solution containing 0.15 mg/ml sonicated calf thymus DNA. The DNase I-generated fragments were ethanol precipitated and run alongside a sequencing ladder of the same fragment.

Purification of the *metJ-lacZ* fusion protein

The construction of the *metJ-lacZ* fusion will be described elsewhere. It consists of the first 41 amino acid codons of the *metJ* gene fused to the eighth codon of the β-galactosidase gene carried on plasmid pMC1403, constructed according to the method of Casadaban et al. (15). Purification of the hybrid protein was by affinity chromatography according to Steers (16). Approximately 3 mg of the purified fusion protein was subjected to automated N-terminal amino acid sequencing on a Beckman 890-C sequencer.

Enzymes and chemicals

Restriction enzymes and other DNA modifying enzymes were obtained from BRL (Gaithersburg, MD) or NEB (Beverly, MA) with the exception of T₄ polynucleotide kinase which was obtained from P-L Biochemicals (Milwaukee, WI). The p-aminophenyl-β-D-thio-galactopyranoside-agarose resin used in the affinity chromatography purification of the *metJ-lacZ* fusion protein was purchased from Sigma (St. Louis, MO). All other chemicals were reagent grade and commercially available.

RESULTS

The *metJ* gene was originally isolated on a 19 kb *EcoRI* DNA fragment and was subcloned on a 3100 bp *ClaI* fragment into plasmid pBR322 (designated pGS107) (4). Further analysis localized the gene to a 1000 bp *MluI-ClaI* fragment within pGS107. A physical map of this 1000 bp *MluI-ClaI* fragment is presented in Fig. 1 along with the DNA sequencing strategy employed. The DNA sequence was determined for both strands by the Maxam-Gilbert technique (5), and all restriction enzyme sites were overlapped. The nucleotide sequence for the *S. typhimurium metJ* gene and the deduced amino acid sequence are shown in Fig. 2. Also shown is the nucleotide sequence for the control region of the *metB* gene.

Location of the 5' ends of *metJ* and *metB* mRNA

Our previous work using Tn5 insertion mutagenesis indicated that the *metJ* promoter region would most likely be situated on a 396 bp *MluI* fragment (4). The S1-nuclease mapping procedure of Weaver and Weissman (6) was used

Figure 1. Restriction endonuclease recognition sites and nucleotide sequence determinations used to establish the *metJ* sequence. Arrows indicate the extent of each sequence determination.

to map the 5' ends of the *metJ* and *metB* mRNA transcripts using the *MluI* fragment as a probe. As shown in Figure 3, two apparent 5' termini exist for the *metJ* transcripts, separated by approximately 65 bp. Although the S1 nuclease mapping procedure is not strictly quantitative, a comparison of the band intensities of the protected fragments within a particular lane is valid. Note in Fig. 3A an increase in the fraction of transcripts initiating from P_{J2} relative to P_{J1} when the RNA for hybridization was isolated from cells grown with methionine (lane c) compared to those grown without methionine (lane b). Thus, transcription initiation at these two sites appears to be regulated in the opposite manner by the presence of methionine in the growth medium. A single 5' terminus was observed for the *metB* transcript. The most likely transcription initiation sites deduced from these results are shown in Figure 2. Sequences that resemble the consensus sequences for the -10 Pribnow box and -35 regions (7) are underlined for both *metJ* transcripts (P_{J1} and P_{J2}) and the *metB* transcript (P_B).

In vitro RNA polymerase protection of the *MluI* fragment

Interactions of RNA polymerase with the proposed promoter regions were examined by binding RNA polymerase to the 396 bp *MluI* fragment and observing protection from subsequent DNase I digestion according to the methods of Schmitz and Galas (8). The protected "footprints" of RNA polymerase bound to the *metJ* and *metB* promoter regions are shown in Figure 4. The corresponding protected regions of the sequence are shown in Figure 2. Under the

Figure 2. The nucleotide and deduced amino acid sequence of the *S. typhimurium metJ* gene. Heavy brackets above the sequence indicate the regions protected by RNA polymerase. Arrows indicate the direction and most likely transcription initiation sites for the two *metJ* promoters (P_{J1} and P_{J2}) and the *metB* promoter (P_B). The most likely Pribnow box and -35 regions for P_{J1} , P_{J2} , and P_B are underlined, as are the possible ribosome binding sites (Shine-Dalgarno sequences) for both genes. A possible operator sequence (10) is boxed. The two translation termination codons for *metJ* are indicated by asterisks, and the proposed transcription termination region is indicated by the vertical arrows.

in vitro binding conditions used in this study (see Materials and Methods), RNA polymerase appears to bind and protect two regions on the *MluI* fragment: one including the 5' terminus of the longest *metJ* transcript (P_{J1}) and its associated -10 and -35 regions, and the other including the 5' terminus of the *metB* transcript and its associated -10 and -35 regions.

Location of the 3' terminus of *metJ* mRNA

The 3' end of *metJ* mRNA was determined by the S1 mapping procedure using a 621 bp *MluI*-*ClaI* fragment ³²P-labeled at the *MluI* 3' terminus as a probe (Materials and Methods). The results of the S1 mapping gel and the proposed termination region are shown in Fig. 5. This proposed termination region is

located about 40 bases distal to the two translation termination codons ending the *metJ* coding sequence (Fig. 2).

Amino acid sequence and composition

The deduced amino acid sequence of the *metJ* gene product shown in Fig. 2 specifies a polypeptide containing 105 amino acid residues. A likely Shine-Dalgarno ribosome binding sequence (9) precedes the AUG initiator codon located at bases 126-128. The choice of this codon as the correct initiation codon is based on N-terminal amino acid sequence studies of a fusion protein containing the amino-terminal 41 amino acids of the *metJ* protein fused to the 8th amino acid of β -galactosidase. Construction of this fusion will be presented elsewhere. These studies show a perfect match of the first 12 amino acids of the fusion protein with amino acids 2 through 13 of the deduced amino acid sequence shown in Fig. 2. The formyl-methionine apparently was removed.

The amino acid composition of the *metJ* gene product, as deduced from the DNA sequence, is shown in Table 1. The calculated molecular weight of Mr 12,110 is in good agreement with the Mr 12,000 value estimated from SDS-polyacrylamide gel electrophoresis of the *metJ* gene product (4). The codon usage frequency is given in Table 2.

DISCUSSION

The promoter regions of the *metJ* and *metB* genes of *Salmonella typhimurium* appear to be complex. S1 nuclease mapping experiments show that the *metJ* gene has two promoters active in vivo (P_{J1} and P_{J2}), separated by approximately 65 bp. However, when these regions are examined by the "foot-printing" technique to determine the binding sites for *E. coli* RNA polymerase, only one of the *metJ* promoters (P_{J1}) shows significant binding of polymerase

Figure 3. Location of the 5' termini of the *metJ* and *metB* gene transcripts. A 396 bp *MluI* fragment carrying the control regions for both the *metJ* and *metB* genes was labeled at the 5' termini with ^{32}P , the strands separated, and each strand hybridized to total cellular RNA isolated from JL781 grown in Luria broth (lane a), JL781 transformed with plasmid pGS107 and grown in glucose minimal medium (lane b), or JL781 carrying pGS107 and grown in glucose minimal medium plus methionine (lane c). Hybridization mixtures were then treated with S1 nuclease and the S1 nuclease-resistant DNA products electrophoresed alongside a sequencing ladder of the original DNA strand. Base numbering of nucleotides is in reference to Fig. 2. A, location of the 3' ends of the protected *metJ* DNA probe corresponding to the 5' termini of *metJ* mRNA at nucleotides +1 (P_{J1}) and +61 to +65 (P_{J2}); B, location of the 3' end of the protected *metB* DNA probe corresponding to the 5' terminus of *metB* mRNA at nucleotide -95 (P_B).

Figure 4. In vitro binding of RNA polymerase to *metJ* and *metB* control regions. A ^{32}P -labeled double strand DNA fragment containing the *metJ* and *metB* control regions was incubated either without (lane a) or with (lane b) *E. coli* RNA polymerase in the presence of 0.1 mM GTP and ATP, and subsequently partially digested with DNase I. The resulting mixture of digestion products was denatured and run alongside of a sequencing ladder of the original fragment. The brackets show the region of the fragment protected from DNase I digestion when RNA polymerase was present in the initial incubation reaction. Asterisks indicate positions which appear to have enhanced susceptibility to DNase I digestion. The corresponding protected sequence is given in Figure 2.

Figure 5. Location of the 3' terminus of *metJ* mRNA. A 621 bp *MluI*-*Clal* fragment carrying the distal part of the *metJ* gene and ^{32}P -labelled at the 3' *MluI* terminus was hybridized to total cellular RNA isolated from JL781 bearing plasmid pGS107. The hybridization mixture was then treated with S1 nuclease and the S1 nuclease-resistant products run alongside a sequencing ladder of the original DNA fragment. The 3' end of *metJ* mRNA (lane a) is indicated by brackets and occurs at nucleotides 482-484 as numbered in Fig. 2. It cannot be determined from this experiment whether transcription actually terminated at this region or whether *metJ* mRNA was processed to this site from a longer transcript. The proposed *metJ* termination region is also shown. The major 3' termini determined by the S1 mapping are indicated by the arrows.

Table 1. Amino acid composition of the *metJ* gene product.

Ala	8	Gln	3	Leu	10	Ser	6
Arg	7	Glu	14	Lys	8	Thr	6
Asn	3	Gly	4	Met	2	Trp	2
Asp	6	His	3	Phe	2	Tyr	3
Cys	1	Ile	7	Pro	6	Val	4

Total number of residues = 105; calculated MW = 12,110.

under the *in vitro* conditions used. It is possible that the failure to see RNA polymerase protection of P_{J2} is due to improper conditions in the *in vitro* binding reaction, e.g., the absence of some binding factor. Alternatively, RNA polymerase may have a much higher affinity for P_{J1} and, when bound at P_{J1} , excludes further RNA polymerase binding at P_{J2} . However, we have not ruled out the possibility that the transcript from P_{J2} is a processed form of the transcript initiated from P_{J1} .

The region of DNA between P_{J1} and the *metB* promoter (P_B) includes a short palindromic sequence (bases -36 to -59, Fig. 2) which in *E. coli* has been proposed as a possible operator sequence (10). This sequence is fully conserved in *S. typhimurium*. Interestingly, in *S. typhimurium* it

Table 2. Codon usage frequency in *metJ*.

Phe UUU	2	Ser UCU	0	Tyr UAU	1	Cys UGU	1
Phe UUC	0	Ser UCC	1	Tyr UAC	2	Cys UGC	0
Leu UUA	1	Ser UCA	0	END UAA	0	END UGA	1
Leu UUG	3	Ser UCG	0	END UAG	1	Trp UGG	2
Leu CUU	0	Pro CCU	1	His CAU	0	Arg CGU	4
Leu CUC	1	Pro CCC	1	His CAC	3	Arg CGC	2
Leu CUA	0	Pro CCA	1	Gln CAA	2	Arg CGA	1
Leu CUG	5	Pro CCG	3	Gln CAG	1	Arg CGG	0
Ile AUU	3	Thr ACU	0	Asn AAU	0	Ser AGU	1
Ile AUC	4	Thr ACC	3	Asn AAC	3	Ser AGC	4
Ile AUA	0	Thr ACA	0	Lys AAA	5	Arg AGA	0
Met AUG	2	Thr ACG	3	Lys AAG	3	Arg AGG	0
Val GUU	1	Ala GCU	2	Asp GAU	6	Gly GGU	1
Val GUC	1	Ala GCC	3	Asp GAC	0	Gly GGC	1
Val GUA	0	Ala GCA	2	Glu GAA	8	Gly GGA	0
Val GUG	2	Ala GCG	1	Glu GAG	6	Gly GGG	2

lies right at the junction between the RNA polymerase binding sites for P_{J1} and P_B . A repressor molecule bound at this site might prevent RNA polymerase binding at both P_{J1} and P_B . The absence of RNA polymerase at P_{J1} could allow RNA polymerase binding at the lower affinity site P_{J2} . Indeed, bound repressor might even be necessary to facilitate polymerase binding at P_{J2} , with P_{J2} acting as a low level maintenance promoter. Consistent with this hypothesis is the observation that the relative amount of mRNA initiated by P_{J1} decreases when the RNA for hybridization was isolated from cells grown in the presence of methionine compared to when cells are grown without methionine (Fig. 3A, lanes b and c). The opposite is observed for mRNA initiated at P_{J2} (Fig. 3A, lanes b and c). The two sets of multiple bands seen for P_{J2} are possibly the result of "nibbling" of the DNA fragments by S1 nuclease (11), or may in fact be two separate initiation sites for P_{J2} .

A single promoter was observed for the *metB* gene, and appears to be responsive to methionine addition to the medium during cell growth (Fig. 3B, lanes b and c). The location of this promoter, determined by S1 nuclease mapping, is also consistent with the RNA polymerase protection experiment shown in Fig. 4 and summarized in Fig. 2. This location is not consistent, however, with the promoter location reported for the *metB* gene in *E. coli* (10), even though the sequences of the two organisms are very homologous in these control regions. A major difference resulting from this discrepancy is that the sequence proposed as the possible operator site discussed above would be contained within the RNA polymerase binding site for the *metB* gene of *S. typhimurium*, whereas in *E. coli* this sequence would be contained within the *metB* mRNA transcript itself, which begins 97 bases further upstream at base -8 in Fig. 2. It is interesting that no S1 nuclease resistant DNA was observed for the P_B promoter when the mRNA for hybridization was isolated from *S. typhimurium* not carrying plasmid pGS107 (Fig. 3B, lane a). Similar results were reported for the *E. coli metB* promoter (10). Using the same mRNA preparation, we did observe S1 nuclease resistant DNA for both P_{J1} and P_{J2} (Fig. 3A, lane a). These results suggest that the *metJ* gene may be more efficiently transcribed than the *metB* gene or that *metJ* mRNA may have a longer half life. Since in vivo studies using *metB-lacZ* and *metJ-lacZ* fusions indicate higher levels of the *metB* gene product are produced than the *metJ* gene product (Urbanowski and Stauffer, unpublished), translation of *metJ* mRNA could be an important point of control. However, the amino acid codon usage given in Table 2 shows little preference for rare codon usage, although some similarities of codon usage by other weakly expressed proteins (18)

do exist, e.g., the preferential use of the glycine codon GGG and the phenylalanine codon UUU.

The translation initiation site and the assignment of a ribosome binding site for the *metB* gene was based on a comparison of the deduced amino acid sequence shown in Fig. 2 with the published amino acid sequence of cystathionine γ -synthase purified from *E. coli* (12). The translation initiation site and the assignment of a ribosome binding site for the *metJ* gene was based on a comparison of the deduced amino acid sequence shown in Fig. 2 with the first 12 N-terminal amino acids determined for the *metJ* protein fused to β -galactosidase. It is possible that the protein was initiated at one of the two other in-phase methionines upstream (positions 69-71 and 102-104), and subsequently processed to our amino acid number 2. However, neither of these other two methionine codons are preceded by a good ribosome binding site.

The experimental results obtained in this study are consistent with the models previously proposed for regulation of the methionine genes (17), wherein the product of the *metJ* gene functions as a trans-acting repressor. One possible version of this model, consistent with the results presented here, is that transcription of the *metB* gene occurs from one promoter, P_B , and that transcription of the *metJ* gene occurs from two distinct promoters P_{J1} and P_{J2} . When cells are grown in methionine-free medium, RNA polymerase can bind to and transcribe from both P_B and P_{J1} efficiently, while P_{J2} is transcribed somewhat less efficiently. In cells grown in medium containing repressing levels of methionine, however, repressor competes with RNA polymerase for binding sites in the DNA region between P_B and P_{J1} , reducing the efficiency of transcription initiation at these promoters. In addition, the less frequent binding of RNA polymerase at P_{J1} would allow more frequent binding at P_{J2} , thus stimulating increased transcription from P_{J2} and maintaining repressor mRNA at relatively constant levels. In support of this model, we have constructed a hybrid *metJ-lacZ* fusion protein whose synthesis is under the control of the *metJ* control region and have measured β -galactosidase activity under various growth conditions. Although β -galactosidase levels in this system respond to methionine in the growth medium, the difference between repressed and derepressed states is only 1.5 to 3 fold (manuscript in preparation). We intend to quantitate more carefully the amounts of mRNA produced by P_{J1} and P_{J2} under various growth conditions in order to show, first, if the regulation of enzyme activity levels seen in the

metJ-lacZ fusion protein studies is at the level of transcription, and second, if this regulation is due to reciprocal accessibility of P_{J1} and P_{J2} .

ACKNOWLEDGEMENTS

We thank Glen Wilson for performing the automated amino acid sequence analysis. This investigation was supported by Public Health Service grant GM 26878 from the National Institute of General Medical Sciences.

REFERENCES

1. Lawrence, D. A., Smith, D. A. and Rowbury, R. J. (1968) *Genetics* **58**, 473-492.
2. Su, C.-H. and Greene, R. C. (1971) *Proc. Natl. Acad. Sci.* **68**, 367-371.
3. Ahmed, A. (1973) *Mol. Gen. Genet.* **123**, 299-324.
4. Urbanowski, M. L. and Stauffer, G. V. (submitted) *Gene*.
5. Maxam, A. M. and Gilbert, W. (1980) *Methods in Enzymol.* **65**, 499-560.
6. Weaver, R. F. and Weissman, C. (1979) *Nucleic Acids Res.* **7**, 1175-1193.
7. Rosenberg, M. and Court, D. (1979) *Ann. Rev. Genet.* **13**, 319-353.
8. Schmitz, A. and Galas, D. J. (1979) *Nucleic Acids Res.* **6**, 111-117.
9. Shine, J. and Dalgarno, L. (1975) *Nature* **254**, 34-38.
10. Duchange, N., Zakin, M. M., Ferrara, P., Saint-Girons, T., Park, I., Tran, S. V., Py, M.-C. and Cohen, G. N. (1983) *J. Biol. Chem.* **258**, 14868-14871.
11. Plamann, M. D. and Stauffer, G. V. (1983) *Gene* **22**, 9-18.
12. Tran, S. V., Schaeffer, E., Bertrand, O., Mariuzza, R. and Ferrara, P. (1983) *J. Biol. Chem.* **258**, 14872-14873.
13. Sanger, F. and Coulson, A. R. (1978) *FEBS Letters* **87**, 107-110.
14. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. Casadaban, M. J., Chou, J. and Cohen, S. N. (1980) *J. Bacteriol.* **143**, 971-980.
16. Steers, E., Jr., Cuatrecasas, P. and Pollard, H. B. (1971) *J. Biol. Chem.* **246**, 196-200.
17. Rowbury, R. J. (1983) In: *Amino Acids: Biosynthesis and Genetic Regulation*, Herrman, K. M. and Somerville, R. L. (eds.). Addison-Wesley, Reading, MA.
18. Grosjean, H. and Fiers, W. (1982). *Gene* **18**, 199-209.