Free Rad. Res. Comms., Vols. 12-13, pp. 443-449
Reprints available directly from the publisher
Photocopying permitted by license only

HALOBACTERIUM HALOBIUM Mn-SOD GENE: ARCHAEBACTERIAL AND EUBACTERIAL FEATURES

MARVIN L. SALIN*, MARY V. DUKE, DIN-POW MA, and JOHN A. BOYLE

Department of Biochemistry and Molecular Biology, P.O. Drawer BB, Mississippi State University, Mississippi State, MS 39762, USA

A 1.8 kb *Pst*I fragment from *Halobacterium halobium* DNA was found to hybridize to synthetic oligonucleotide probes constructed by using the sequence of the N-terminus of a Mn-containing superoxide dismutase purified from *H.halobium*. The entire insert containing a 600-bp coding sequence for Mn-SOD and its 5' and 3' flanking regions was sequenced. The derived amino acid sequence of the structural gene showed a similarity to other manganic and iron-containing superoxide dismutases in normally conserved regions. Primer extension analysis of the *H. halobium* Mn-SOD mRNA showed that gene transcription begins 14 bases upstream of the translational start. A Shine-Dalgarno sequence and archaebacterial consensus promoter sequences were observed. Several other promoter and terminator nucleotide sequences homologous to prokaryotic and eukaryotic organisms were found.

KEY WORDS: Archaebacteria, halophilic protein, transcript mapping, superoxide dismutase, nucleotide sequence, amino acid sequence.

INTRODUCTION

Archaebacteria are microoganisms representing a third and unique line of evolution. The group consists of halophiles, methanogens and thermoacidophiles, which share features with eubacteria as well as eukaryotes.[1] Eubacterial features include prokaryotic cell structure and organization, whereas eukaryotic features include the presence of introns in stable RNA genes, an RNA polymerase that shows similarities to eukaryotic RNA polymerases, and a translation apparatus that is sensitive to antibiotics known to interact with eukaryotic ribosomes. [2]

Halophiles are charcterized by growth in salt concentrations in excess of 1 M and can usually be found growing in salt brines or salty bodies of water such as the Dead Sea.[3] These organisms generate ATP under anaerobic conditions by using a light-driven proton pump, bacteriorhodopsin, in coordination with an ATP synthase. Under aerobic conditions, they maintain a standard respiratory metabolism.[3] Therefore, halophilic organisms would be exposed transiently to high levels of oxygen and superoxide dismutase should be present in the organsim at levels to insure sufficient oxyradical scavenging.

The enzyme from *Halobacterium halobium* is a manganic protein of about 40,000 Mr and consists of two equally sized subunits.[4] The gene has been cloned and the nucleotide sequence determined.[5] Here we report on the identification of the

* To whom inquiries should be addressed.

transcriptional start site. The 5′ -flanking region shows features in common with other archaebacteria. In addition, some regions have been found which show homology to eubacteria while other regions show homology to eukaryotes.

MATERIALS AND METHODS

A clone containing the Mn-SOD gene was isolated from a *Halobacterial halobium* cosmid genomic library.[6] A description of the Mn-SOD gene isolation procedure as well as the sequencing strategies and protocols have been published previously.[5] The transcriptional start site was determined by primer extension as described by Golden *et al.*[7] A synthesized 17-mer (TCGTGGTGCCACGTGAC) complementary to the 5′ region of SOD mRNA was ^{32}P-labeled by polynucleotide kinase and used as the primer in the extension reaction.

RESULTS

The complete nucleotide sequence for the 1.8 kb *Halobacterium halobium* DNA fragment is shown in Figure 1. The open reading frame comprises 600 nucleotides coding for 200 amino acids. The results of the primer extension (Figure 2) indicated that the transcript starts 14 nucleotides upstream from the initiation codon. The first base in the transcript is *G*. Unfortunately, this region is highly compressed on the gel and the DNA sequence is consequently difficult to unambiguously read. The problem of band compression on gels was also apparent in the original gene sequencing.[5] The final sequence determination was achieved by multiple analyses of the region and its complement.

The amino acid sequence of the open reading frame was compared and aligned to other manganic superoxide dismutases, and the results are shown in Figure 3. Major regions of homology are found interspered throughout the molecule.

Based upon X-ray structural analysis and sequence alignment, His^{26}, His^{81}, Asp^{175}, and His^{179} have been tentatively assigned as ligands to the manganese.[8,9] These four amino acids have been aligned and labeled with asterisks. As can be seen in Figure 3, the positions of these critical amino acids, with the possible exception of His^{81} in *E. coli*, have been conserved.

DISCUSSION

The expression of archaebacterial genes is less than completely understood. Fewer than 30 transcriptional start sites have been mapped for genes in this entire kingdom.[10] Of this number, greater than half represent either genes coding for stable RNA transcripts or genes found in viral-like particles. Therefore, relatively few genes for archaebacterial proteins have been studied in detail. A close examination of the sequence depicted in Figure 1 and Figure 3 shows that the gene for Mn-SOD exhibits both eubacterial as well as eukaryotic features.

Previous work on archaebacterial promoters has suggested the presence of conserved sequences in an A and B box arrangement.[10,11] The B box is a weakly conserved region located in the immediate vicinity of the transcriptional start site. The consensus

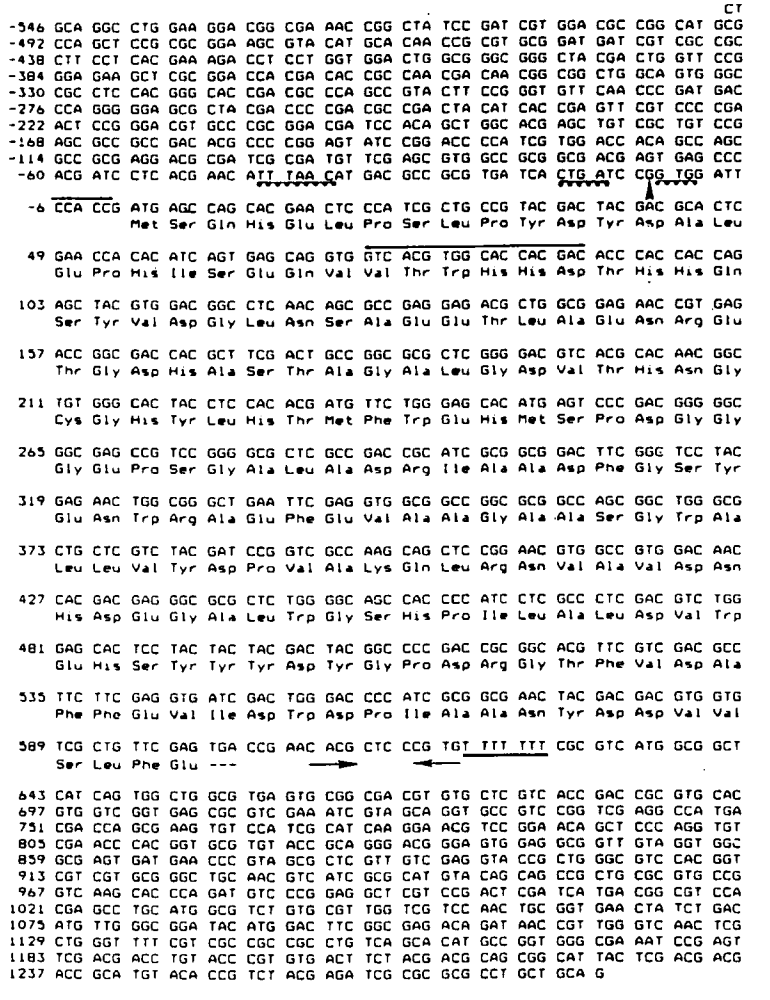

FIGURE 1 Nucleotide sequence of *Halobacterium halobium* Mn-SOD gene and its flanking regions. Presumed consensus archaebacterial transcriptional promoters at the 5′ end are noted by underlines containing dots. The arrowhead marks the transcriptional start site. Dashes above the nucleotides immediately preceeding the open reading frame denote a putative modulator of eukaryotic translation. A double line at the 3′ end denotes the poly T transcriptional termination signal; the arrows represent an inverted repeat required for formation of a loop structure. A 17-mer complementary to the overlined sequence was used in primer extension analysis.

sequence is (A/T)TG(A/C). There are two regions in the gene transcript that could be this box (Figure 1). The sequence CTGA ends three nucleotides upstream from the start site (arrow in Figure 1). The sequence GTGG is located immediately after the start site. These sequences are underlined with dots in Figure 1. The A box is a more highly conserved sequence ordinarily centered about 25 nucleotides upstream form box B and from the start of transcription. Its consensus sequence is usually rendered as TTA(T/A)ATA. The sequence TTTAACAT is found at positions -23 to -30. This appears to be an excellent candidate for the A box both in sequence and location.

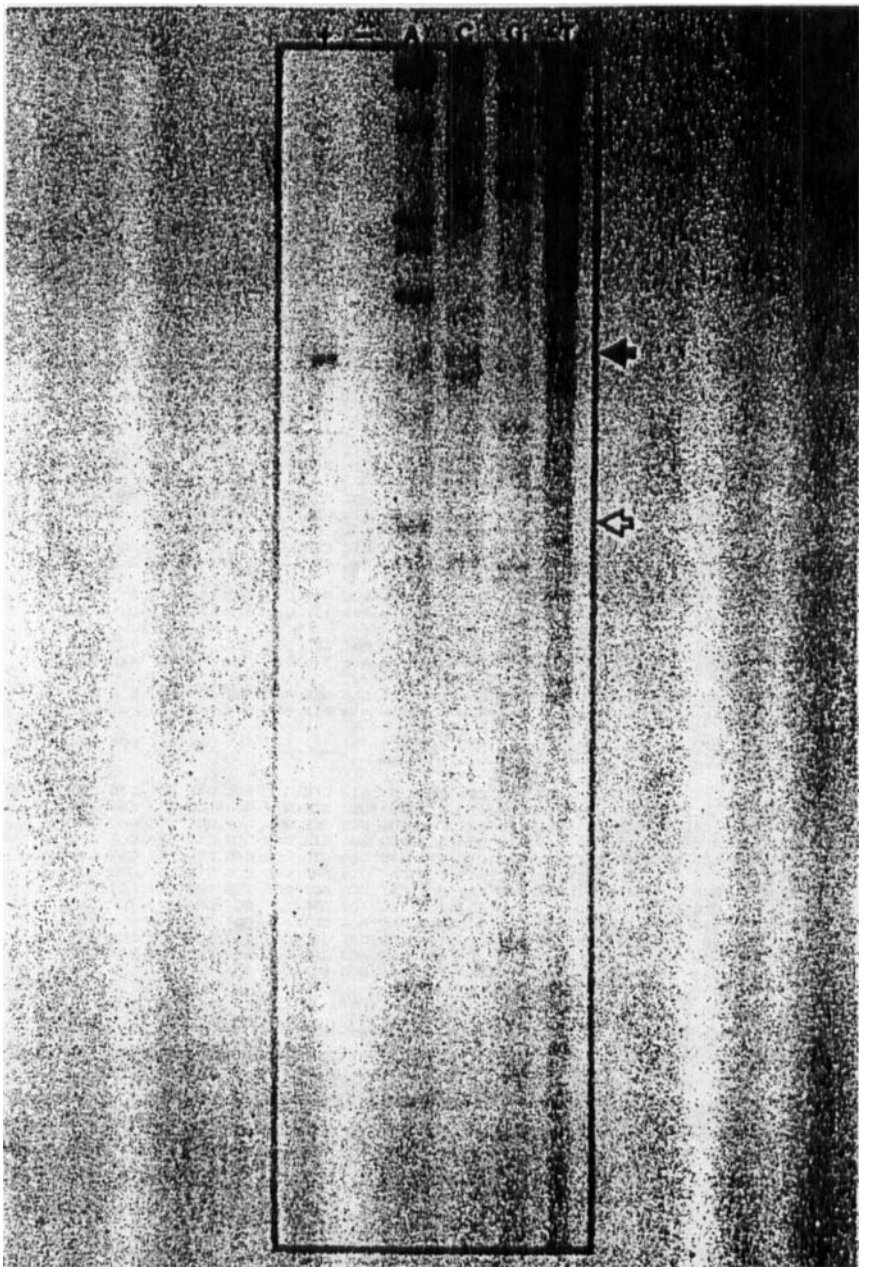

FIGURE 2 Determination of the transcriptional start site by the primer extension method. The plus (+) and minus (−) symbols are the extension reaction with and without mRNAs, respectively. Four lanes A,C,G,T represent dideoxy sequence reactions of the M13 single-stranded DNA template using a 17-mer oligonucleotide as a sequencing primer. The extended product is marked by a closed arrowhead and the translational initiation site is denoted by a open arrowhead.

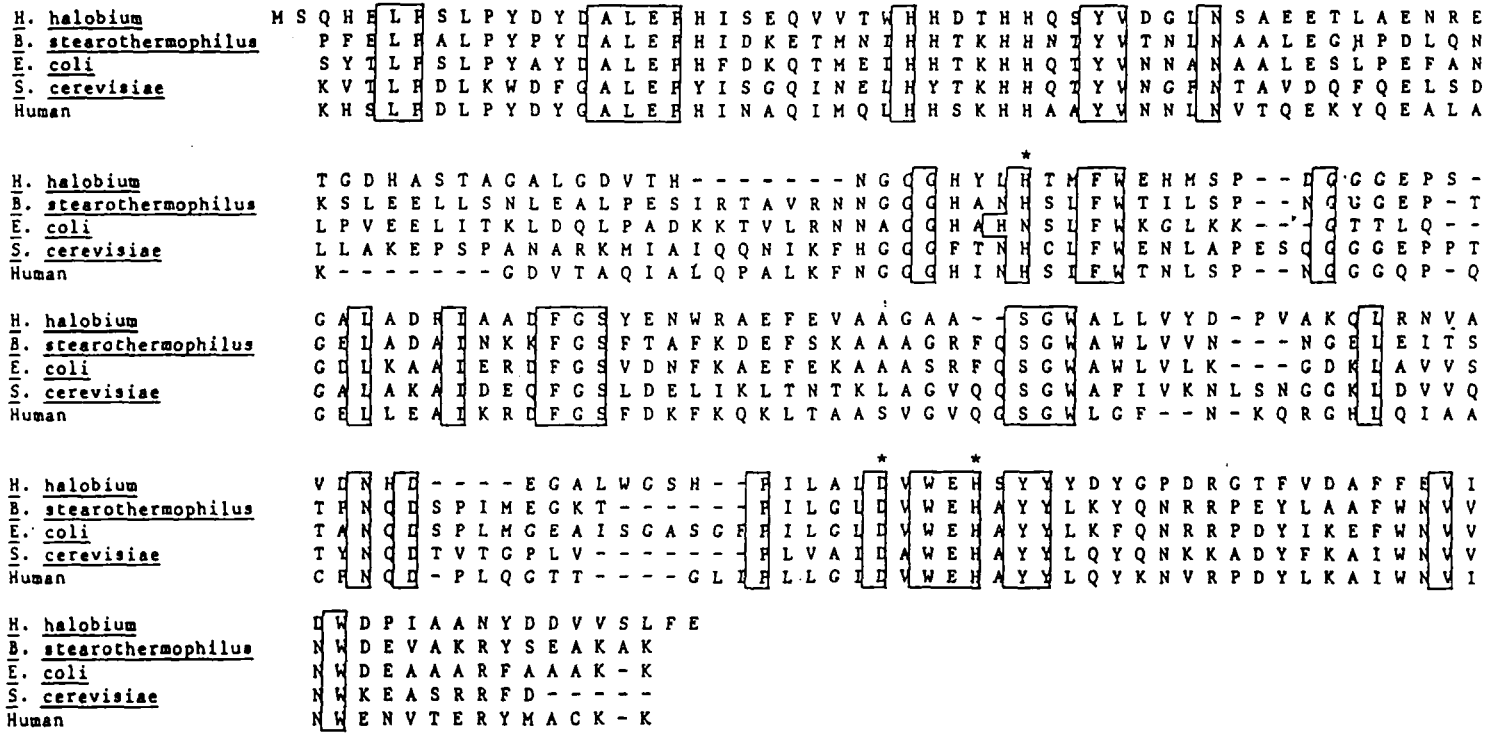

FIGURE 3 Amino acid sequence of *Halobacterium halobium* Mn-SOD in comparison to other sequenced Mn-SODs. Data from *Bacillus stearothermophilus, Escherichia coli, Saccharomyces cerevisiae*, and human liver Mn-SODs were obtained from a compilation in Bannister *et al.*[9] Amino acids with greatest homology are in boxes. Asterisks are placed over amino acids postulated to be critical in metal liganding. Gaps have been introduced to achieve greatest sequence homology.

The 3′ end of the 16S rRNA of *Halobacterium halobium* is 3′ UCCUCCACU.[12] This sequence is remarkably similar to the sequences seen in eubacterial organisms.[12] The 5′ end of the gene transcript, GGUGG, is a good Shine-Dalgarno sequence which is capable of binding to the 16S rRNA. Several halobacterial and other archaebacterial gene transcripts for proteins have start sites within one, two, or three nucleotides of the translation initiation codon.[10,11] These genes appear to have sequences complementary to the 3′ end or the 16S RNA within the gene rather than in the untranslated leader. Therefore, there seems to be no obvious positional rules directing translational initiation beyond the necessity for rRNA-mRNA complementary sequences. However, the fact that the archaebacterial sequences are essentially the same as the eubacterial sequences may indicate that this method of assisting translation initiation predates the split between the two kingdoms. Eukaryotes do not use such complementary sequences and therefore have either lost the necessity or developed independently of the other two urkingdoms prior to the appearance of Shine-Dalgarno-type interactions. The sequence, CCACCT, (shown by dots over the nucleotides) directly adjacent to the start codon is analogous to the consensus sequence used in eukaryotic translation.[13]

At the 3′ end of the gene a series of T's are observed (double underline). This string of seven T's is preceded by an inverted repeat of 3 nucleotides. If one counts two of the seven T's as part of the inverted repeat, then the repeat becomes 5 nucleotides long and surrounds a potential loop of 4 bases. Potential stem and loop regions preceeding a string of T's are characteristic of rho independent transcriptional termination sequences associated with eubacteria.

In conclusion, it would appear therefore, that this *H. halobium* gene possesses certain eukaryotic as well as eubacterial features in its promoter, translation initiation and terminator sequences.

Acknowledgements

This work was supported by funds made available from the Mississippi Agricultural and Forestry Experiment Station.

References

1. C.R. Woese, L.J. Magrum, and G.E. Fox, (1978) Archaebacteria. *Journal Molecular Evolution*, **11**; 245–252.
2. P.P. Dennis, (1986) Molecular biology of archaebacteria. *Journal of Bacteriology*; **168**, 471–478.
3. D.J. Kushner, (1985) The halobacteriaceae. In Woese, C.R., and Wolfe, R.S. (eds). The Bacteria- A Treatise on Structure and Function. Academic Press, NY, pp. 171–214.
4. M.L. Salin, and D. Oesterhelt, (1988) Purification of a manganese-containing superoxide dismutase from *Halobacterium halobium*. *Archives in Biochemisty and Biophysics*, **260**; 806–810.
5. M.L. Salin, M.V. Duke, D. Oesterhelt and D-P. Ma, (1988) Cloning and determination of the nucleotide sequence of the Mn-containing superoxide dismutase gene from *Halobacterium halobium*. *Gene*, **70**; 153–159.
6. P. Hegemann, A. Blanck, H. Vogelsang-Wenke, F. Lottspeich and D. Oesterhelt, (1987) The halopsin gene. I. Identification and isolation. *EMBO Journal*, **6**; 259–264.
7. S.S. Golden, J. Brusslan and R. Haselkorn, (1986) Expression of a family of psbA genes encoding a photosystem II polypeptide in the cyanobacterium *Anacystis nidulans*R2. *EMBO Journal*, **5**; 2789–2798.
8. W.C., Stallings, K.A. Pattridge, R.K. Strong and M.L. Ludwig, (1985) The structure of manganese superoxide dismutase from *Thermus thermophilus* HB8 at 2.4Å resolution. *Journal of Biological Chemistry*, **260**; 16424–16432.

9. J.V. Bannister, W.H. Bannister and G. Rotilio, (1987) Aspects of the structure, function, and applications of superoxide dismutase. CRC Critical Reviews in Biochemistry. CRC Press, Boca Raton, FL. pp. 111–180.
10. W. Zilling, P. Palm, W-D. Reiter, F. Gropp, G. Pühler and H-P. Klenk, (1988) Comparative evaluation of gene expression in archaebactria. *European Journal of Biochemistry*, 173; 474–482.
11. J.W. Brown, C.J. Daniels and J.N. Reeve (1989) Gene structure, organization, and expression in archaebacteria. *CRC Critical Reviews in Microbiology*16; 287–338.
12. Shine, J. and L. Dalgarno. (1974) The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA:complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National* Academy of Sciences USA, 71; 1342–1346.
13. M. Kozak (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44; 283–292.

Accepted by Prof. G. Czapski