Volume 244, number 1, 34-38 FEB 06772 February 1989

Characterization of the structural genes for the DNA-binding protein H-NS in Enterobacteriaceae

A. La Teana*+, M. Falconi*, V. Scarlato°, M. Lammi and C.L. Pon+

*Laboratory of Genetics, DBC, University of Camerino, 62032 Camerino (MC), Italy, +Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Berlin 33, Germany, °International Institute of Genetics and Biophysics, CNR Naples, Italy and Dept. of Cell Biology, University of Calabria, 87036 Arcavacata di Rende (CS), Italy

Received 2 December 1988

The promoter region of *Escherichia coli hns*, the structural gene for the DNA-binding protein H-NS, has been identified by use of a promoter search vector and the in vivo transcriptional start point by primer extension analysis. The homologous *hns* genes of two other Enterobacteriaceae, *Proteus vulgaris* and *Serratia marcescens*, were identified by heterologous hybridization with a DNA probe derived from *E. coli hns*, cloned and sequenced. Taking into account only the invariant nucleotides and amino acids, the homology of H-NS among the three organisms was found to be >70% at the DNA level and >75% at the protein level. The three *hns* genes were also found to have nearly identical transcriptional and translational signals.

Chromatin; DNA-binding protein; Promoter structure; Nucleotide sequence; (Enterobacteriaceae)

1. INTRODUCTION

In bacteria, several proteins with DNA-binding properties have been implicated in folding and condensing the long circular DNA molecule [1-3]. Among these proteins, NS (HU) is the most abundant and best characterized. The primary structure of NS has been determined and found to be highly conserved in many prokaryotic species which have proteins sharing a number of identical residues, especially in the C-terminal half, in the hydrophobic regions and in the DNA-binding arms [2,4].

Another protein believed to be implicated in the organization of bacterial chromatin is the 15.4 kDa protein H-NS (probably homologous to B1 [5] and to one of the H1 variants reported by Spassky et al. [6]), which has recently been characterized and whose *E. coli* structural gene *hns* has been cloned and sequenced [1,7-9]. Here, we report on the identification of the promoter region and of the point of transcriptional initiation of this gene. Furthermore, the degree of conservation of *hns* has been evaluated by cloning and sequencing the homologous genes from two other enteric bacteria, *Proteus vulgaris* and *Serratia marcescens*.

Correspondence address: C.L. Pon, Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Berlin 33, Germany

The nucleotide sequences presented here have been submitted to the EMBL/Gen Bank database under the accession no. Y00976

2. MATERIALS AND METHODS

2.1. *Promoter search*

DNA fragments derived from the upstream region of *E. coli hns* were obtained by restriction endonuclease digestions (fig.1) and tested for promoter activity by cloning in pKK232-8 (Pharmacia) upstream from the chloramphenicol resistance (CAT) gene which lacks a promoter.

2.2. *Primer extension analysis*

Total RNA was extracted from cell pellets of *E. coli* MRE6600 obtained from 50 ml of a culture grown to $0.6A_{550}$ in complete medium. The cells were resuspended in 3.7 ml of 100 mM Tris-HCl buffer, (pH 8.3) containing 2 mM EDTA and 1% SDS and incubated for 5 min in a boiling water bath. After the addition of 75 μl of 2 M KCl, samples were kept on ice for 10 min, then centrifuged at 18000 rpm for 15 min. To 3.5 ml of the supernatant, 4.65 g CsCl and 34 μl of 98% β-mercaptoethanol

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

were added. The samples were then centrifuged for 18 h at 33 000 rpm in the Spinco SW50.1 at room temperature. The resulting RNA pellets were resuspended in 400 μl of 10 mM Tris-HCl (pH 8.0), containing 1 mM EDTA and the RNA precipitated by addition of 0.1 vol. of 2 M LiCl and 2 vols ethanol. Aliquots (1–10 μg) of the RNA preparation were mixed with 1 pmol of a 5′ ^{32}P-labelled primer consisting of the 17-mer with the sequence 5′-GCGCACGAAGAGTACGG-3′ which is complementary to the sequence underlined in fig.1. Following precipitation with ethanol/LiCl, the samples were resuspended in 7 μl of 50 mM Tris-HCl buffer (pH 8.3), containing 6 mM $MgCl_2$, 10 mM NaCl and 10 mM DTT. After 3 min at 65°C and 1 min cooling at −80°C, the primer was allowed to anneal for 40 min on ice. After the addition of 1 μl (1.7 U) AMV reverse transcriptase (Pharmacia) and 2 μl of a solution containing 2 mM each of the four dNTPs, the reaction was allowed to proceed for 30 min at 48°C before being stopped by addition of formamide gel sample buffer. The samples were then analyzed by 6% polyacrylamide gel electrophoresis in urea.

2.3. *Gene identification, cloning and sequencing*

The 200 bp *Hin*cII-*Hin*cII and 270 bp *Eco*RI-*Pst*I fragments derived from *E. coli hns* [8] were nick-translated and used as probes in cross-hybridization [10] to total digests of *P. vulgaris* and *S. marcescens* chromosomal DNA. A 3 kb *Hin*dIII fragment from *P. vulgaris* and a 4 kb *Eco*RI fragment from *S. marcescens* which gave positive hybridization bands were cloned into pACYC177 and pACYC184, respectively [11]. After restriction mapping of the cloned inserts and further hybridization analysis, the appropriate fragments were subcloned into M13mp18 or mp19 or pTZ18 or 19 for sequencing by the dideoxy chain-termination method [12] using 2′-deoxy-7-deazaguanosine triphosphate in place of dGTP.

3. RESULTS AND DISCUSSION

3.1. *Identification of the in vivo promoter region of* hns *of* E. coli

Several DNA restriction fragments derived from

```
                                                                  ▼1
  1  TCGTCTCATTCAAAAAACCTCCGCAACCCCATGTTTTCACATAACTGTTGCGTTGACCAA  60

 61  TTGAATCTACAGTAGCCTTTTTTAATATTTCATTCTCCATTTCAATGCGTTGTAGCTTTT 120

121  TCCTCAGCTCACGTATTTCGATTTGTTCTGGTGTTATCGGAGAGGATTTTGGTGTTTTGC 180

181  CCTGACGCTCATCACGCAGTTGTTTGACCCATCTTGTCATTGTGGAAAGGCCAACATCCA 240
                                                  ▼1
241  TAGCTTTGGCGGCATCTGCCACCGTGTAGTTCTGGTCAACAACCAGTTGAGCGGATTCGC 300
     ▼2
301  GTTTAAACTCTGCGCTGAAATTTCTTTTTTTCATTGGAGCACCTGTGTTGTTCTGAGGTG 360

361  AGCATATCACCTCTGTTCAGGTGGCCAAATTCAGTGTGCCACTTCACCTCGCTTGTTATA 420

421  AGCGGGTAAATGACTGCTGGTAACTATTCACAATCTTTAACCTGTTGCGCAAGTAATAGC 480

481  CCTCTGTTGACCTCCAGGAGATAGTGCAATACTAAGTCCATGCTCTTATTGCGACTTGTT 540
                                        ▼3
541  CTACTTTTCATCATTCGCTTAATAGGGAATTCTCGTAAACACAACTAATACAGAAGACTG 600

601  AAAGGTCGTCAGCCTACGATAATCTCCCCATAAAATGTGACATGAATCAGGAAGTTTTAA 660

661  CCTCACGTGCTGCGAAATCATCGGTGTAAATAGGGCTATATGCCGCGTCTTTTCTGGCTA 720
                                                          ▼3
721  ATTTTATGAAAAGATATTTATTGGCGGCACAAAATAAAGAACAATTTTGAATTCCTTACA 780
                                                     *
781  TTCCTGGCTATTGCACAACTGAATTTAAGGCTCTATTACTACCCCAACAAACCACCCCAA 840
                                              ▼4
841  TATAAGTTTGAGATTACTACAATGAGCGAAGCACTTAAAATTCTGAACAACATCCGTACT 900

                     MetSerGluAlaLeuLysIleLeuAsnAsnIleArgThr

901  CTTCGTGCGCAGGCAAGAGAATGTACACTTGAAACGCTGGAAGAAATGCTGGAA..... 960

     LeuArgAlaGlnAlaArgGluCysThrLeuGluThrLeuGluGluMetLeuGlu.....
```

Fig.1. Upstream and proximal region of *E. coli hns* gene. The four restriction fragments tested for promoter activity in pKK232-8 comprised between the inverted triangles are: (1-1) *Hin*cII-*Hin*cII; (2-3) *Dra*I-*Eco*RI; (3-3) *Eco*RI-*Eco*RI; (3-4) *Eco*RI-*Fok*I. The transcription and translation initiation signals are indicated in boldface and underlined, the sequence complementary to the primer used for the primer extension analysis is underlined, and the start point of transcription, identified by this method (see fig.2) is indicated by an asterisk.

the upstream region of *E. coli hns* (fig.1) were tested for promoter activity. The only fragment found to have promoter activity was the 108 bp *Eco*RI-*Fok*I segment immediately preceding and overlapping the beginning of the coding region of the gene.

In agreement with this finding, primer extension analysis demonstrated that the point of transcriptional start is the adenine at position 826 and, to a much lesser extent, 827 (fig.2). Since the coding region of *hns* is followed by a typical rho-independent termination signal, these data indicate that *hns* is transcribed as a monocistronic unit.

3.2. *Homology between the* hns *genes of* E. coli, P. vulgaris *and* S. marcescens

A large number of gram-negative and gram-positive bacteria as well as archaebacteria were searched for the presence of genes homologous to *E. coli hns*. This search was carried out by Southern hybridization of total chromosomal DNA digests with probes derived from *E. coli hns*. These experiments suggested that *hns*-like genes exist in all gram-negative bacteria examined, but, in spite of the presence of hybridization bands, we could not establish with certainty the presence of this gene in gram-positive bacteria. Finally, in the case of the archaebacteria, the experiments failed to reveal any positive cross-hybridization band, even when hybridization conditions of low stringency were used.

To establish the degree of evolutionary conservation of the *hns* gene and of its product and to determine which are the most strictly conserved parts of the molecule, we cloned and sequenced the *hns* genes from *P. vulgaris* and *S. marcescens*. The DNA sequences of the two genes and the deduced amino acid sequences are presented in fig.3; in fig.4, we present a comparison of the primary structures of the H-NS proteins known so far. As seen from the above figures, and from the quantitative data concerning the number of identical nucleotides and amino acids (table 1), it appears that *hns* is highly conserved both at the DNA and at the protein level, although the *E. coli* protein is two amino acids longer than the other two; it is noteworthy that the homology at the DNA level is restricted to the structural gene and to its transcriptional and translational signals (indicated in bold letters in fig.3) but that hardly any DNA homology

Fig.2. Identification of the transcriptional start point of *E. coli hns*. Primer extension analysis was carried out as described in section 2. Lanes: (1–4) G, A, T and C, lanes of the DNA sequence gel; (5–7) primer extension reaction with 6, 3 and 9 μg total RNA from *E. coli* MRE600.

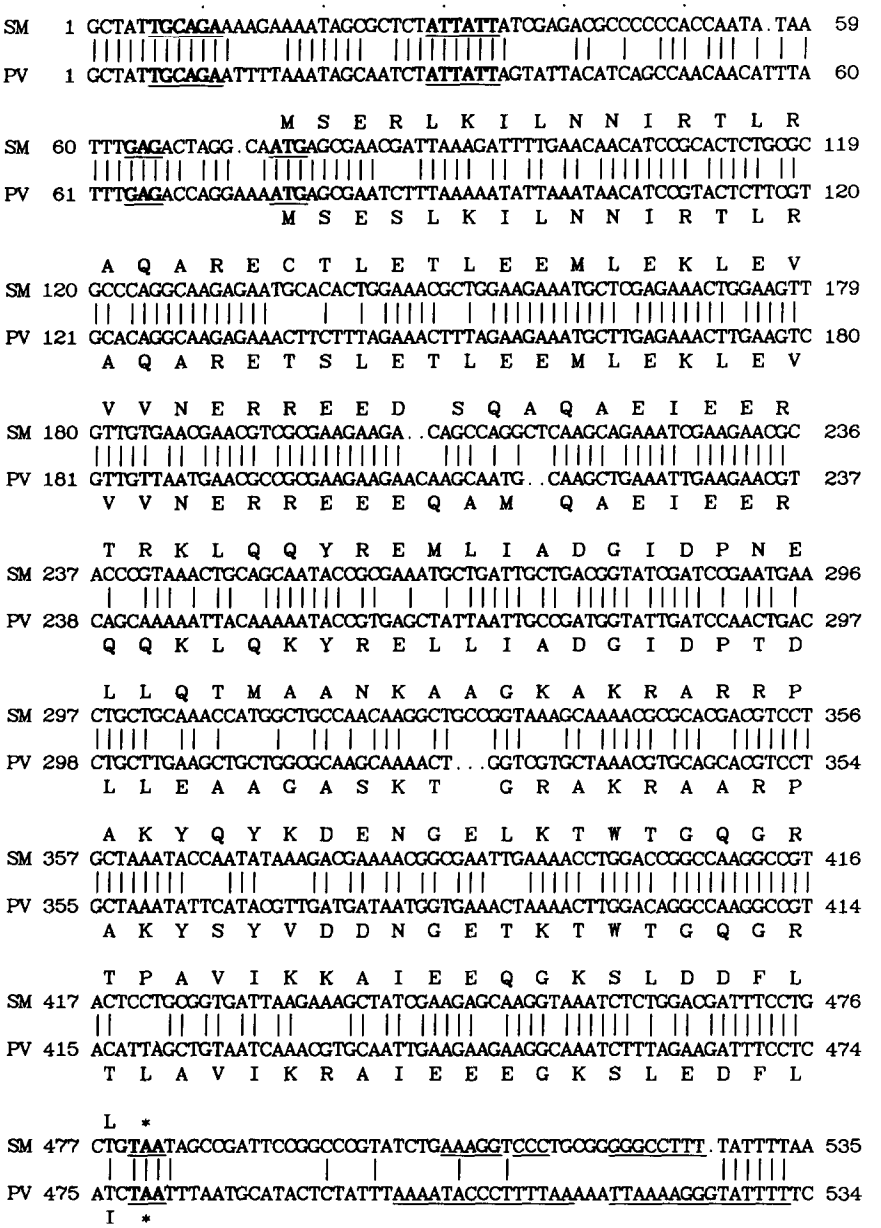

Fig.3. Comparison of the *hns* genes of *Proteus vulgaris* (PV) and *Serratia marcescens* (SM) and their corresponding H-NS proteins. The putative transcriptional operator and translational signals are underlined and printed in boldface. The regions of self-complementarity downstream of the genes are underlined.

can be found in the flanking regions (not shown). The transcriptional and translational signals are also identical to those of *E. coli* but for the presence of a C in the fifth position of the −10 box in the place of a T which is found in both *Proteus* and *Serratia.* At the protein level, if one also takes into account the conservative amino acid replacements, the extent of homology among the three H-NS proteins approaches 90%. Among the conserved residues are most of the basic amino acids and all the aromatic residues (i.e. three tyrosines, one phenylalanine and the single tryptophan residue, which in *E. coli* H-NS was found to be buried within the protein structure and sen-

```
SM      1 SERLKILNNIRTLRAQARECTLETLEEMLEKLEVVVNERREEDSQAQAEI  50
EC      1 SEALKILNNIRTLRAQARECTLETLEEMLEKLEVVVNERREEESAAAAEV  50
PV      1 SESLKILNNIRTLRAQARETSLETLEEMLEKLEVVVNERREEEQAMQAEI  50

SM     51 EERTRKLQQYREMLIADGIDPNELLQTMAANKAAGKAKRARRPAKYQYKD 100
EC     51 EERTRKLQQYREMLIADGIDPNELLNSLAAVKSGTKAKRAQRPAKYSYVD 100
PV     51 EERQQKLQKYRELLIADGIDPTDLLEAAGASKTG.RAKRAARPAKYSYVD  99

SM    101 ENGELKTWTGQGRTPAVIKKAIEEQGKSLDDFLL   134
EC    101 ENGETKTWTGQGRTPAVIKKAMDEQGKSLDDFLIKQ 136
PV    100 DNGETKTWTGQGRTLAVIKRAIEEEGKSLEDFLI   134
```

Fig.4. Comparison of the primary structures of H-NS from *Escherichia coli* (EC), *Serratia marcescens* (SM) and *Proteus vulgaris* (PV).

Table 1

Homology among *hns* genes and H-NS proteins in Enterobacteriaceae

	DNA	Protein
EC/PV	70.4	78.3
EC/SM	78.9	88.0
PV/SM	71.8	77.8

Values represent the percentage of identical nucleotides and amino acids. EC, *E. coli;* PV, *P. vulgaris;* SM, *S. marcescens*

sitive to DNA binding [9]). The single Cys residue found in *E. coli* H-NS, on the other hand, is found in the protein from *Serratia,* but not in that from *Proteus* and must, therefore, be considered to be functionally dispensable. Overall, the rather large degree of conservation of H-NS, even when one takes into account that all three bacteria examined belong to the Enterobacteriaceae, suggests the existence of rather strict structural constraints for this DNA-binding protein.

Acknowledgements: This work was aided by a grant of the Italian Ministry of Public Education to Dr C.O. Gualerzi (University of Camerino) in whose laboratory part of the work was carried out. The help of Mrs Marisa Bartoli in some experiments is gratefully acknowledged.

REFERENCES

[1] Gualerzi, C.O., Losso, M.A., Lammi, M., Friedrich, K., Pawlik, R.T., Canonaco, M.A., Gianfranceschi, G.L., Pingoud, A. and Pon, C.L. (1986) in: Bacterial Chromatin (Gualerzi, C.O. and Pon, C.L. eds) pp. 101–134, Springer, Heidelberg.

[2] Drlica, K. and Rouviere-Yaniv, J. (1987) Microbiol. Rev. 51, 301–319.

[3] Pettijohn, D.E. (1988) J. Biol. Chem. 263, 12793–12796.

[4] Djik, J. and Reinhardt, R. (1986) in: Bacterial Chromatin (Gualerzi, C.O. and Pon, C.L. eds) pp. 185–218, Springer, Heidelberg.

[5] Varshavsky, A.J., Nedospasov, S.A., Bakayev, V.V., Bakayeva, T.G. and Georgiev, G.P. (1977) Nucleic Acids Res. 4, 2725–2745.

[6] Spassky, A., Rimsky, S., Garreau, H. and Buc, H. (1984) Nucleic Acids Res. 12, 5321–5340.

[7] Falconi, M., Gualtieri, M.T., La Teana, A., Losso, M.A. and Pon, C.L. (1988) Mol. Microbiol. 2, 323–329.

[8] Pon, C.L., Calogero, R.A. and Gualerzi, C.O. (1988) Mol. Gen. Genet. 212, 199–202.

[9] Friedrich, K., Gualerzi, C.O., Lammi, M., Losso, M.A., Pawlik, R.T. and Pon, C.L. (1988) FEBS Lett. 229, 197–202.

[10] Mosely, S.L. and Falkow, S. (1980) J. Bacteriol. 144, 444–446.

[11] Chang, A.C.Y. and Cohen, S.N. (1978) J. Bacteriol. 134, 1141–1156.

[12] Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143, 161–178.