A SINGLE DNA-BINDING PROTEIN FROM *PSEUDOMONAS AERUGINOSA* HOMOLOGOUS TO PROTEINS NS1 AND NS2 (HU PROTEINS) OF *ESCHERICHIA COLI* AND OTHER BACTERIA

Alastair R. HAWKINS and John C. WOOTTON
Department of Genetics, University of Leeds, Leeds LS2 9JT, England

Received 8 June 1981

1. Introduction

NS1 and NS2 are two homologous, 90-residue polypeptides from *Escherichia coli* that have been completely sequenced [1,2]. They are basic, heat-stable DNA-binding proteins, present in approximately equimolar amounts, that are also found associated with native 30 S ribosomal subunits [2]. The same pair of proteins, purified as a mixture by DNA-binding properties or as part of nucleoprotein complexes, has been isolated from *E. coli* in different laboratories and given different names [3–6]. Here we adopt the prior name HU [3] for this family of prokaryotic DNA-binding proteins, and the names NS1 and NS2 [1,2] to denote specifically the *E. coli* sequences. A recent discussion of the possible relationships of HU proteins to other prokaryotic DNA-binding proteins and to histones is given in [6].

We report here the purification and partial sequence of an HU protein from *Pseudomonas aeruginosa* strain PA01607. Only a single HU protein, which is very closely homologous to NS1 and less closely homologous to NS2 of *E. coli* was found in this strain. This protein resembled the pair of *E. coli* proteins in showing DNA-condensing properties and (in contrast to the stimulatory effect reported [3]) in inhibiting transcription in vitro of bacteriophage λ DNA. We also report, on the basis of limited N-terminal sequence data, that *Bacillus subtilis*, like *E. coli*, has two closely related HU proteins.

2. Materials and methods

Single-stranded DNA-cellulose was prepared from Munktell 410 cellulose [7]. Materials and methods for amino acid analysis, protease digestion, peptide purification, manual Dansyl-Edman sequencing and amide assignments were as previously described [8–10].

HU proteins were purified from *P. aeruginosa* strain PA01607, *E. coli* strain J6-2 and *B. subtilis* strain IT$^+$ (a derivative of strain Marburg) by the following modification of the method in [3]. Pellets of late exponential phase cells, grown in shaken culture in Oxoid nutrient broth, were washed and resuspended to 0.5 g ml^{-1} in buffer 'A' (10 mM Tris–HCl, 10 mM $MgCl_2$, 2 mM $CaCl_2$, 1 mM EDTA, 50 mM NaCl, pH 8.1) and sonicated in the presence of 10 μg ml^{-1} DNase I. After digestion of the DNA at 8°C for 4 h the suspension was cleared by a 30-min low speed (75 000 × *g*) and a 4-h high speed (195 000 × *g*) centrifugation. The supernatant was dialysed into buffer 'B' (20 mM Tris–HCl, 10 mM EDTA, 1 mM 2-mercaptoethanol, 10% (v/v) glycerol, pH 8.1) and loaded onto a single-stranded DNA-cellulose column. After washing, the column was eluted with a linear 250 ml 0.0 M to 2.0 M NaCl gradient in buffer 'B'. Fractions were screened for HU proteins by sodium dodecyl sulphate gel electrophoresis and amino acid analysis, HU proteins ($M_r \sim 9500$) eluting at approximately 0.4 M NaCl. Fractions containing HU were pooled, dialysed and run through a column of Whatman DE-52 DEAE-cellulose equilibrated with buffer 'B'. HU proteins were eluted in the unretarded fractions, contaminated being retarded. In some cases a final heat treatment (10 min, 97°C) was used to precipitate small quantities of thermolabile contaminants.

The effect of HU proteins on transcription in vitro was assayed using bacteriophage λ DNA and *E. coli* RNA polymerase containing 0.7 equivalents of sigma subunit (both Boehringer, Mannheim) by the same methods as in previous work [3].

3. Results

Yields of purified HU proteins were equivalent to ~13 000 polypeptide chains per cell from *P. aeruginosa* and *E. coli*, similar to reported yields from *E. coli* [3,4]. Lower yields of ~3300 molecules per cell were obtained from *B. subtilis*, losses probably being due to proteolysis. HU proteins from all three sources were indistinguishable in size by sodium dodecyl sulphate gel electrophoresis, were heat stable, and showed very similar amino acid analyses resembling the characteristic compositions published for the *E. coli* [1–4] and cyanobacterial [11] proteins.

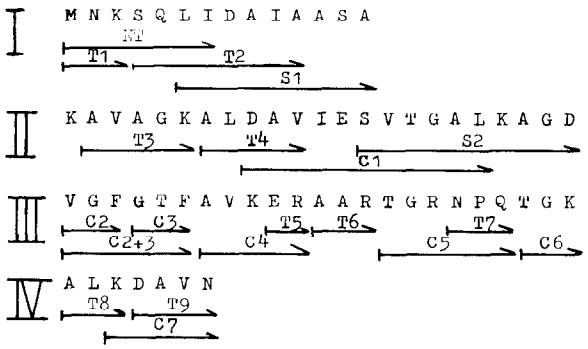

Fig.1. Partial amino acid sequences of *P. aeruginosa* HU protein. Fragments labelled I, II, III and IV correspond respectively to residues 1–14, 18–40, 45–67 and 84–90 of *E. coli* NS1 (see fig.2). Peptides from digestion with trypsin (T), chymotrypsin (C) and *Staphylococcus aureus* V8 protease (S) were fractionated by successive steps of column chromatography on AG50X4, Biogel P4 and AG1X2 and sequenced by manual methods, as previously described [8–10]. Rightward pointing arrows indicate the extent of sequences determined that were unambiguous and showed good consistency with amino acid analyses according to the criteria described in [8]. NT: manual N-terminal sequence of the intact protein. Acid/amide assignments were from TLC of anilinothiazolinone derivatives produced in the Edman degradation [10] and were also consistent with peptide net electrostatic charge for T1, C1, T5, C7 and T9. Gln-5 (fragment I) was a clear assignment and the staphylococcal protease split at this residue, giving S1, is presumed to reflect partial deamidation. The first Lys of fragment II is the presumed C-terminus of T2, assigned from the specificity of trypsin and the presence of Lys but not Arg in the amino acid analyses of T2 and S1. Peptide pairs T3/T4, C3/C4, T5/T6, T6/C5, C5/C6 and T8/T9 are aligned only on the basis of homology with *E. coli* NS1. All other pairs (T1/T2, T2/S1, T4/C1, C1/S2 and C2/C3) are aligned by independently derived overlapping sequences. Most of the sequences have been determined at least twice in independent digests, but those represented by T3, C4, T6 and C6 are from only single peptide sequences and must be regarded as unconfirmed.

```
     1                   10                  20        25
 HU                    A       S   ? ? ?     V     K
NS1  M N K S Q L I D K I A A G A D I S K A A A G R A L
NS2        T         V     E K   E L     T Q   K A

     26      30                  40                  50
 HU      V   E       G A     A     ? ? ? ?
NS1  D A I I A S V T E S L K E G D D V A L V G F G T F
NS2  E S T L   A I                 A   Q

     51                  60                  70        75
 HU                                      ? ? ? ? ? ? ? ?
NS1  A V K E R A A R T G R N P Q T G K E I T I A A A K
NS2  K   N H     E                          K         N

     76      80                  90
 HU  ? ? ? ? ? ? ? ?
NS1  V P S F R A G K A L K D A V N
NS2      A   V S                 K
```

Fig.2. Sequence comparisons of *P. aeruginosa* HU protein and *E. coli* proteins NS1 and NS2. The complete sequence of NS1 [1] is shown together with the residue that differ in *P. aeruginosa* HU (fig.1) and NS2 [1] (above and below the NS1 sequence respectively). ?, indicates positions for which corresponding *P. aeruginosa* residues have not been reliably determined.

The partial sequence of *P. aeruginosa* HU protein (fig.1) was deduced from sequences of peptides purified after digestion with trypsin, chymotrypsin or staphylococcal protease V8, aligned by overlaps where indicated and by homology with *E. coli* NS sequences. For three parts of the 90 residue *E. coli* sequences (residues 15 to 17, 41 to 44 and 68 to 83) no reliable corresponding *P. aeruginosa* sequences have been determined. Possible candidates for these missing sequences as judged by amino acid composition, were obtained in low yield peptides, thus bringing the total length of the *P. aeruginosa* polypeptide to approximately 90 residues. Considering just the 67 residues aligned (fig.2), there are only 9 differences between *P. aeruginosa* HU and *E. coli* NS1, compared with 19 between *E. coli* NS1 and NS2 and 23 between NS2 and *P. aeruginosa* HU. Most differences are compatible with single nucleotide substitutions.

In the preparations of *P. aeruginosa* HU peptides used for sequencing, obtained from several separate protein purifications, there was no evidence of sequence microheterogeneity (although this cannot be ruled out for the parts of the protein not definitively sequenced). In contrast the protein purified from *E. coli*, which we subjected to a similar protocol of digestion with trypsin or chymotrypsin followed by purification and sequencing of peptides, was clearly a mixture of NS1 and NS2 in approximately similar amounts, all peptide sequences obtained being compatible with those published [1]. *B. subtilis* HU

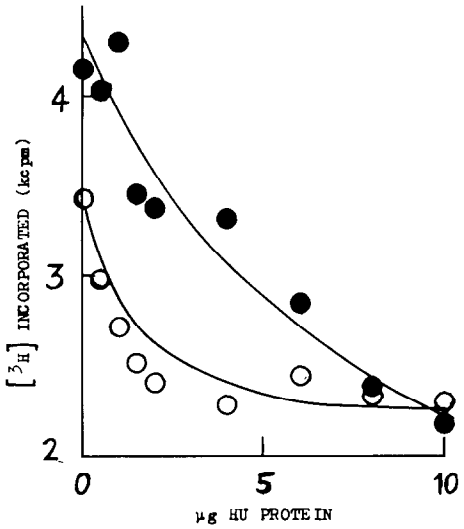

Fig.3. The effect of *P. aeruginosa* and *E. coli* HU proteins on the transcription in vitro of bacteriophage λ DNA. The incorporation of [^{3}H]UTP (Amersham, 40 Ci mol^{-1}) into acid-insoluble polyribonucleotides was measured using for each assay (final volume 120 μl) 6 μg of λ DNA (this amount in preliminary experiments gave optimal incorporation in the absence of HU proteins) and 1 unit of RNA polymerase (supplier's definition). The DNA was preincubated for 5 min at 37°C with the indicated amounts of *P. aeruginosa* (•) or *E. coli* (○) HU proteins in the assay buffer before addition of enzyme. Buffers and other assay conditions were as described in [3]. Each point is the average of assays from two experiments carried out on separate days. The *E. coli* HU protein used was shown by partial sequencing to be an approximately equimolar mixture of NS1 and NS2.

protein also appeared, from manual N-terminal sequencing, to be a mixture of two homologous polypeptides in approximately equal amounts. Its sequence Met–Asn–Lys–Thr/Asn–(Pro)–Leu– (Pro-5 is a tentative identification) showed clear microheterogeneity, with approximately equal yields of Thr and Asn at position 4. The other residues correspond closely to the *P. aeruginosa* and *E. coli* sequences.

E. coli HU proteins have been reported to stimulate transcription of bacteriophage λ DNA markedly [3]. This effect was reinvestigated using complexes of λ DNA with either the mixture of *E. coli* proteins or *P. aeruginosa* HU (fig.3). In contrast to the earlier report HU proteins reduced the amount of transcription significantly. Electron microscopic and light scattering studies showed that *P. aeruginosa* HU and the mixture of *E. coli* proteins are very similar in their DNA-condensing properties (A. R. H., M. J. McPherson, V. Virrankoski-Castrodeza and J. C. W., to be published separately). In the range of DNA and protein concentrations used in the transcription assays (fig.3) tightly compact DNA–HU complexes were observed.

4. Discussion

The single HU protein found in our extracts of *P. aeruginosa* is strongly homologous to NS1 of *E. coli* and is similar in amount extracted, in purification properties, in affinity for DNA and in DNA-condensing properties to the pair of proteins found in all other bacteria examined (which include *Salmonella* in addition to *E. coli*, *B. subtilis* and cyanobacteria, J. Rouvière-Yaniv, personal communication). Might we have overlooked a second *P. aeruginosa* protein? No other HU protein was present in the extracts used: other small DNA-binding proteins occurred in the extracts of all three bacterial species but were clearly distinct from HU proteins in amino acid compositions and in their elution properties from DNA-cellulose and DE-52 columns. Therefore the most likely explanation is that *P. aeruginosa* has only a single HU type. Possibly this bacterial strain might have a second protein that differs from the first, and from all other bacterial HU proteins studied, in remaining completely bound to intracellular complexes under the extraction conditions used (or in being completely degraded in extracts, which is unlikely in view of the stability of other HU proteins). In the low salt extraction buffers used, together with DNase I, in our work and elsewhere [3–6], relatively little of the total HU is released from DNA–protein complexes [5,12] or ribosomes [1,2]. The total intracellular content may be as high as 80 000 to 100 000 polypeptide chains per cell [6,15] compared with the 13 000 typically purified by this method. Since the function of HU proteins in vivo is not established, further work on this and other *P. aeruginosa* DNA-binding proteins would be valuable, particularly in view of the suggestion [12,15] that in *E. coli* NS1 and NS2 might function as a heterodimer in the compaction of DNA.

Amino acid sequences (fig.2) and other properties [11] of HU proteins are very strongly conserved during evolution. More amino acid substitutions have been fixed in the N-terminal half of the polypeptide than in the C-terminal half. Interestingly, *P. aeruginosa* HU is closer in sequence to *E. coli* NS1 than are the two *E. coli* sequences to each other. Caution is

required in applying a simple model of evolutionary divergence to bacteria [13] because gene duplication, gene loss or intergeneric gene transfer might have occurred several times on different evolutionary lines. There is considerable phylogenetic distance, shown by DNA–DNA and DNA–mRNA hybridization methods [14], between *P. aeruginosa* and *E. coli* and both these gram-negative bacteria are remote from cyanobacteria and the gram positive *B. subtilis*.

The discrepancy between our transcription assays (fig.3) and those previously published [3] might be due to undefined differences in the RNA polymerase or λ DNA preparations used. Our data show considerably more incorporation of [^{3}H]UTP in the absence of HU protein than do the corresponding control assays in [3].

Acknowledgements

We thank Helen Gilmore for operating the amino acid analyser and the SRC for a studentship (A. R. H.) and a research grant (J. C. W.).

References

[1] Mende, L., Timm, B. and Subramanian, A. R. (1978) FEBS Lett. 96, 395–398.
[2] Syryanarayana, T. and Subramanian, A. R. (1978) Biochim. Biophys. Acta 520, 342–357.
[3] Rouvière-Yaniv, J. and Gros, F. (1975) Proc. Natl. Acad. Sci. USA 9, 3428–3432.
[4] Berthold, V. and Geider, K. (1976) Eur. J. Biochem. 71, 443–449.
[5] Varshavsky, A. J., Nedospasov, S. A., Bakayev, V. V., Bakeyeva, T. B. and Georgiev, G. P. (1977) Nucleic Acids Res. 4, 2725–2745.
[6] Lathe, R., Buc, H., Lecocq, J.-P. and Bautz, E. F. K. (1980) Proc. Natl. Acad. Sci. USA 77, 3548–3552.
[7] Alberts, B. and Herrick, G. (1972) Methods Enzymol. 21, 198–217.
[8] Wootton, J. C., Taylor, J. G., Jackson, A. A., Chambers, G. K. and Fincham, J. R. S. (1975) Biochem. J. 149, 739–748.
[9] Wootton, J. C., Baron, A. J. and Fincham, J. R. S. (1975) Biochem. J. 149, 749–755.
[10] Holder, A. A., Wootton, J. C., Baron, A. J., Chambers, G. K. and Fincham, J. R. S. (1975) Biochem. J. 149, 757–773.
[11] Haselkorn, R. and Rouvière-Yaniv, J. (1976) Proc. Natl. Acad. Sci. USA 73, 1917–1920.
[12] Rouvière-Yaniv, J., Yaniv, M. and Germond, J.-E. (1979) Cell 17, 265–274.
[13] Sneath, P. H. A. (1974) Symp. Soc. Microbiol. 24, 1–39.
[14] McCarthy, B. J. and Bolton, E. T. (1963) Proc. Natl. Acad. Sci. USA 50, 156–164.
[15] Rouvière-Yaniv, J. (1977) Cold Spring Harbor Symp. Quant. Biol. 42, 439–447.