The amino acid sequence of a small DNA binding protein from the archaebacterium *Sulfolobus solfataricus*

Makoto Kimura, Junko Kimura, Philippa Davie, Richard Reinhardt and Jan Dijk*

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Ihnestrasse 63-73, D-1000 Berlin 33, Germany

Received 23 August 1984

The thermoacidophilic archaebacterium *Sulfolobus solfataricus* possesses several DNA binding proteins which may have a histone-like function. Two particularly dominant species have molecular masses of 7 and 10 kDa, respectively. We have purified one of the small proteins which occurs in a relatively large amount and have determined its amino acid sequence. The protein is characterized by a high lysine content; in the N-terminal region the lysine residues occur in an alternating order: X-K-X-K-X-K-X-K. The amino acid sequence does not indicate any obvious homology to those DNA binding proteins whose sequences have been determined.

Archaebacteria Sulfolobus solfataricus *DNA binding protein* *Amino acid sequence*

1. INTRODUCTION

In comparison with the detailed knowledge of the structure of eukaryotic chromatin, very little is known of the ultrastructure of the bacterial chromosome. Although it is obvious that the 1–2 mm long circular genome has to be compactly folded in order to fit into the 1–2 μm large cell, the mechanism of folding is not understood. Histone-like proteins (9.5 kDa), called HU, HD, NS1/2 or DNA binding protein II [1], have been isolated from several bacteria and their amino acid sequences have been determined [2–6]. The protein from the thermophile *B. stearothermophilus* has been crystallised [7] and its three-dimensional structure solved [8,9].

A similar protein (HTa) has been purified from the thermoacidophilic archaebacterium *Thermoplasma acidophilum* [10]. Its amino acid sequence [11] shows a clear homology to the group of eubacterial proteins [4]. Authors in [12] have reported the occurrence in several *Sulfolobus* species of at least two groups of DNA binding proteins with estimated molecular masses of 9 and 6 kDa, respectively.

We have studied these proteins in the archaebacterium *Sulfolobus solfataricus* which is able to grow at temperatures as high as 87°C [13]. Several DNA binding proteins with molecular masses of either 10 or 7 kDa were found. A protein of the smaller category which occurred in a relatively large amount was purified to homogeneity and its amino acid sequence was determined.

2. EXPERIMENTAL

S. solfataricus cells (DSM1616) were grown at 80°C and pH 4.8. Cells were harvested by centrifugation at a cell density of 2–3 g/l (wet wt, 2–3 A_{650} units). Ribosomes and DNA binding proteins were prepared as in [7]. Purified proteins were stored in small aliquots at −80°C in 0.35 M NaCl, 20 mM sodium phosphate (pH 7.0).

Filter binding assays were performed as in [14], the KCl concentration was increased to 260 mM. ^{14}C-labelled *E. coli* DNA was obtained from Dr C.O. Gualerzi and ^{3}H-labelled RSF1010 plasmid DNA from Dr E. Lanka.

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V.

DNA binding experiments using exclusion chromatography on Sephadex G-50 were carried out as in [7].

Procedures for enzymic digestion and isolation of peptides by fingerprinting have been described in [4]. Cleavage with CNBr (6 mg) was performed on a 1-mg aliquot of the protein in 300 μl 70% formic acid for 20 h in the dark at room temperature. The peptides were separated on a Sephadex G-75 superfine column (1 × 210 cm) in 10% acetic acid. Sequence determination was performed as in [4].

3. RESULTS AND DISCUSSION

S. solfataricus cells were grown in rich medium [13] in a double-walled glass container, equipped with a reflux condenser. The temperature was kept at 80°C by a circulation bath and the pH at 4.8 (measured at 80°C) with 1 M H_2SO_4, using a titrator. Details of the fermentation will be published elsewhere.

The DNA binding proteins were isolated according to the procedure used for the purification of the DNA binding protein II from *B. stearothermophilus* [7]. The proteins were separated by gel filtration on Sephacryl S-300 in 1 M NH_4Cl. Two groups of proteins, with apparent molecular masses of 10 and 7 kDa, respectively, were obtained. The smaller proteins were purified further by chromatography on CM-Sepharose CL-6B. The column (1.6 × 25 cm) was eluted with a linear gradient of 0.1–0.5 M NaCl. Two components were detected on SDS gels. From 60 g of cells, 10 mg of component I and 3 mg of II were obtained. These proteins probably correspond to the B1 and B2 proteins described in [12].

The proteins were identified as DNA binding proteins by filter binding assays [14], using *E. coli* and RSF1010 plasmid DNA. The conditions given in [14] were modified in order to exclude spurious binding of small basic proteins, such as lysozyme, cytochrome *c* and ribosomal proteins.

These results were confirmed by binding experiments on a Sephadex G-50 column where comigration of the protein with DNA in the void volume peak was demonstrated by analysis on SDS gels.

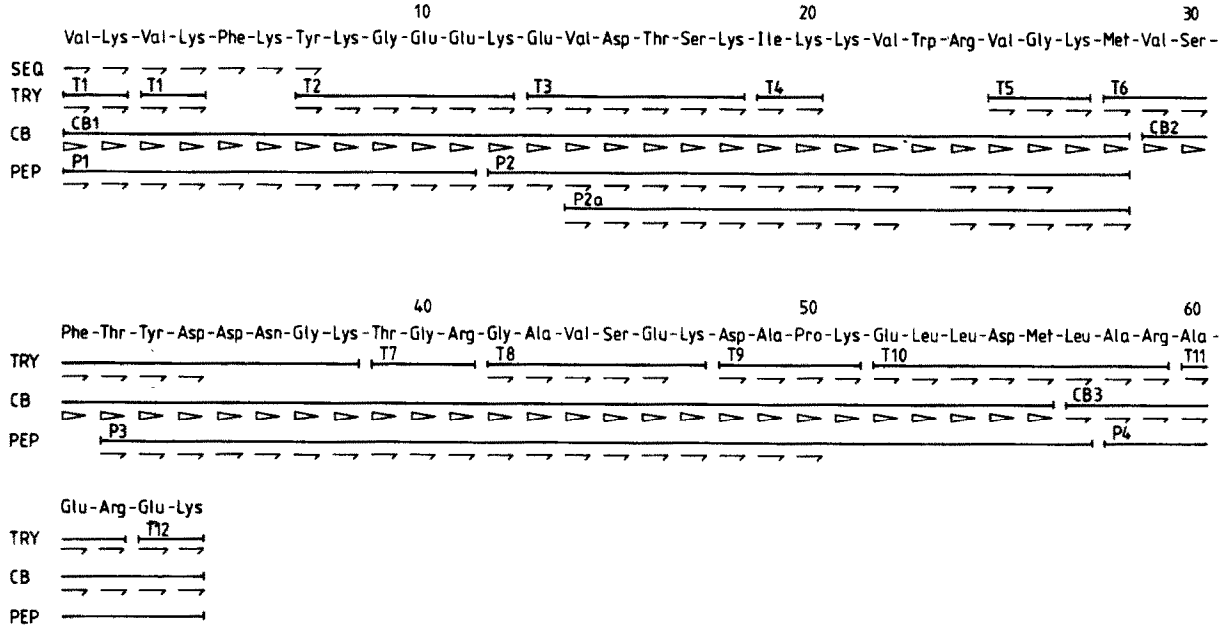

Fig.1. The amino acid sequence of the DNA binding protein from *Sulfolobus solfataricus*. Sequence data on individual peptides are indicated as follows: (—) sequenced by the DABITC/PITC double coupling method; (▷) sequenced by the manual solid-phase procedure. TRY and PEP indicate peptides derived from digestion with trypsin and pepsin, respectively; CB, cleavage with CNBr.

	5				10				15	
S. solfataricus	V	K	V	K	F	K	Y	K	G	E
H6	P	K	R	K	S	A	T	K	G	D
HMG-1	G	K	G	D	P	K	K	P	G	G
HMG-2	G	K	G	D	P	D	K	P	R	G
HMG-14	P	K	R	K	V	S	S	A	E	G
HMG-17	P	K	R	K	A	E	G	D	A	K

S. solfataricus	E	E	K	G	V	D	-			
H6	E	P	A	R	R	-				
HMG-1	K	M	S	S	Y	-				
HMG-2	K	A	A	Y	-					
HMG-14	A	A	K	E	E	-				
HMG-17	G	D	A	K	G	D	G	A	K	-

Fig.2. Comparison of the N-terminal region of the DNA binding protein from *Sulfolobus solfataricus* with that of several high-mobility-group proteins. The sequence information on the latter is derived from [16,17].

The amino acid sequence of component I was determined using the same techniques as for the DNA binding protein from *B. stearothermophilus* [4]. The complete sequence was obtained from two large and one small CNBr peptide which were sequenced by the manual solid-phase sequencing method [15]. Confirmation of this sequence was obtained from tryptic peptides and, in the N-terminal region, from peptic cleavage products. The results are summarised in fig.1.

From a first inspection of the sequence it is clear that this protein is not homologous to any of the prokaryotic DNA binding proteins from which it also differs in size. There is a remote resemblance to a group of eukaryotic DNA binding proteins, the high-mobility group (HMG) of proteins. These proteins possess a similar but not identical alternation of lysine residues (fig.2) at the N-terminus [16,17] which is believed to be important for the interaction with DNA [18].

Further studies on the mode of binding of these small proteins to DNA will be required in order to understand the relevance of apparent structural homologies to their function. Preliminary studies indicate that these small proteins exist in several forms, all of which have a very similar amino acid sequence.

ACKNOWLEDGEMENTS

Drs E. Lanka and C.O. Gualerzi kindly provided us with radioactive DNA for the filter binding assays. Mathias Grote is thanked for providing his unpublished results on the DNA binding assays and Dr H.G. Wittmann for his continuous interest and support.

REFERENCES

[1] Geider, K. and Hoffmann-Berling, H. (1981) Annu. Rev. Biochem. 50, 233–260.
[2] Mende, L., Timm, B. and Subramanian, A. (1978) FEBS Lett. 96, 395–398.
[3] Laine, B., Kmiecik, D., Sautiere, P., Biserte, G. and Cohen-Solal, M. (1980) Eur. J. Biochem. 103, 447–461.
[4] Kimura, M. and Wilson, K.S. (1983) J. Biol. Chem. 258, 4007–4011.
[5] Hawkins, A.R. and Wootton, J.C. (1981) FEBS Lett. 130, 275–278.
[6] Laine, B., Belaiche, D., Khanaka, H. and Sautiere, P. (1983) Eur. J. Biochem. 131, 325–331.
[7] Dijk, J., White, S.W., Wilson, K.S. and Appelt, K. (1983) J. Biol. Chem. 258, 4003–4006.
[8] Tanaka, I., White, S.W., Appelt, K., Wilson, K.S. and Dijk, J. (1984) FEBS Lett. 165, 39–42.
[9] Tanaka, I., White, S.W., Appelt, K., Dijk, J. and Wilson, K.S. (1984) Nature, in press.
[10] DeLange, R.J., Green, G.R. and Searcy, D. (1981) J. Biol. Chem. 256, 900–904.
[11] DeLange, R.J., Williams, L.C. and Searcy, D.G. (1981) J. Biol. Chem. 256, 905–911.
[12] Thomm, M., Stetter, K.O. and Zillig, W. (1982) Zbl. Bakt. Hyg. C3, 128–139.
[13] Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H. and Scholz, I. (1980) Arch. Microbiol. 125, 259–269.
[14] Lammi, M., Paci, M. and Gualerzi, C.O. (1984) FEBS Lett. 170, 99–104.
[15] Wittmann-Liebold, B. and Lehmann, A. (1980) in: Methods in Peptide and Protein Sequence Analysis (Birr, C. ed.) pp. 49–72, Elsevier, Amsterdam.
[16] Walker, J.M., Parker, B.P. and Johns, E.W. (1978) Int. J. Pept. Prot. Res. 12, 269–276.
[17] Walker, J.M., Goodwin, G.H. and Johns, E.W. (1979) FEBS Lett. 100, 394–398.
[18] Cary, P.D., King, D.S., Crane-Robinson, C., Bradbury, E.M., Rabbani, A., Roodwin, G.H. and Johns, E.W. (1980) Eur. J. Biochem. 112, 577–580.