Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 5788–5791, August 1990
Biochemistry

HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus fervidus*, is most closely related to histones

(DNA–protein complex/thermal stability/molecular evolution/histone sequence conservation)

K. SANDMAN*, J. A. KRZYCKI*, B. DOBRINSKI†, R. LURZ†, AND J. N. REEVE*‡

*Department of Microbiology, Ohio State University, Columbus, OH 43210; and †Max-Planck-Institute, Abteilung Trautner, Ihnestrasse 63-73, D-1000 Berlin 33, Federal Republic of Germany

Communicated by Carl R. Woese, May 22, 1990

ABSTRACT **_Methanothermus fervidus_ grows optimally at 83°C. A protein designated HMf (histone _M. fervidus_) has been isolated from this archaeal hyperthermophile that binds to double-stranded DNA molecules and increases their resistance to thermal denaturation. HMf binding to linear double-stranded DNA molecules of >2 kilobase pairs also increases their electrophoretic mobilities through agarose gels. Visualization of this compaction process by electron microscopy has demonstrated the formation of quasispherical, macromolecular HMf–DNA complexes. HMf is a mixture of approximately equal amounts of two very similar polypeptides designated HMf-1 and HMf-2. Determination of the DNA sequence of the gene encoding HMf-2 (_hmfB_) has revealed that over 30% of the amino acid residues in HMf-2 are conserved in the consensus sequences derived for eucaryal histones H2A, H2B, H3, and H4. These archaeal polypeptides and eucaryal histones appear therefore to have evolved from a common ancestor and are likely to have related structures and functions.**

The discovery of hyperthermophilic microorganisms, species that grow above 70°C and in some cases as high as 115°C, has challenged us to determine how these living systems function and can withstand such very high temperatures. *Methanothermus fervidus* is a hyperthermophilic, methanogenic archaeal species[§] that grows well under laboratory conditions, optimally at 83°C, and is therefore an attractive candidate for such studies. Several *M. fervidus* genes encoding proteins and stable RNAs have already been sequenced and they, and their encoded products, compared with the functionally homologous macromolecules from moderate thermophiles and mesophiles (2–5). As the genome of *M. fervidus* is overall only 33 mol % G+C, its thermal denaturation *in vivo* at 83°C must be prevented by extrinsic factors such as a high intracellular salt concentration and DNA-binding proteins. The cytoplasm of *M. fervidus* does, in fact, contain ≈950 mM K^+ and ≈300 mM 2,3-(cyclic)diphosphoglycerate (6). We now report the isolation and characterization of a DNA-binding protein (HMf) from *M. fervidus* that, based on primary sequences, is most closely related to eucaryal histones.[¶] Binding of HMf to double-stranded DNA (dsDNA) molecules increases their resistance to thermal denaturation, increases the electrophoretic mobility through agarose gels of DNA molecules >2 kilobase pairs (kbp), and results in the formation of large macromolecular complexes reminiscent of nucleosomes.

MATERIALS AND METHODS

Protein Isolation and Purification. *M. fervidus* cultures were grown as described (3). Cells were ruptured in 3 M KCl/50 mM Tris·HCl, pH 8.0, by passage through a French pressure cell at 20,000 psi (≈138 MPa). Following centrifugation (30,000 × *g*, 4°C, 30 min, then 100,000 × *g*, 20°C, 90 min) the cleared supernatant was passed through a Bio-Gel P200 column. DNA-binding activity was detected in fractions of the column effluent by its ability to *increase* the electrophoretic mobility of *Hin*dIII-linearized pUC19 DNA during subsequent electrophoresis through 0.8% (wt/vol) agarose gels. Fractions with DNA-binding activity were combined and dialyzed overnight against 100 mM NaCl/50 mM Tris·HCl/2 mM Na_3PO_4, pH 8.0. Adsorption to dsDNA-cellulose and elution with 1 M NaCl/50 mM Tris·HCl/2 mM Na_3PO_4, pH 8.0, resulted in the release of material (designated HMf) that formed only one band, with an apparent molecular weight of 7200, when analyzed by SDS/PAGE. Gel filtration under nondenaturing conditions indicated a molecular weight of 17,500. Application of HMf to a Vydac C_{18} reverse-phase HPLC column (2.1 × 150 cm) and elution with a linear 5–80% gradient of acetonitrile at a flow rate of 200 μl/min resulted in its resolution into approximately equal amounts of two polypeptides designated HMf-1 and HMf-2. Microsequencing determined that 32 of the 35 N-terminal residues (91%) of HMf-1 and HMf-2 were identical (see Fig. 3).

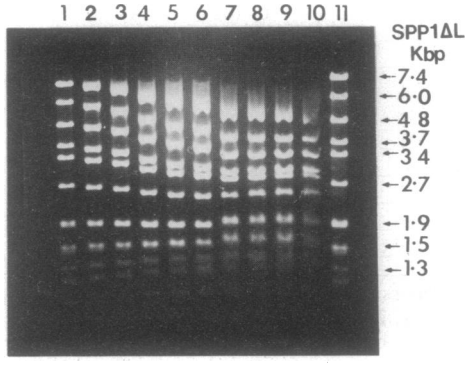

FIG. 1. Electrophoretic separation (0.8% agarose) of *Eco*RI fragments of phage SPP1ΔL DNA bound by various amounts of HMf. Fragments were visualized by fluorescence after staining with ethidium bromide. The ratio (wt/wt) of HMf to SPP1ΔL DNA (8) in lanes 2–10 was 1:8, 1:4, 1:2, 3:4, 1:1, 5:4, 3:2, 7:4, and 5:2, respectively. Separation of *Eco*RI-generated restriction fragments of SPP1ΔL DNA in the absence of HMf is shown in lanes 1 and 11.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: dsDNA, double-stranded DNA.

‡To whom reprint requests should be addressed.

§This paper uses the taxonomic classification proposed by Woese *et al.* (1), in which the highest taxa are the domains Archaea (archaebacteria), Bacteria (eubacteria), and Eucarya (eukaryotes).

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. M34778).

DNA Cloning and Sequencing. A ^{32}P-labeled oligodeoxynucleotide probe, synthesized based on the amino acid sequence obtained for HMf-2 (see Fig. 3), was used in a colony hybridization screen to identify *Escherichia coli* clones containing pUC19-based recombinant plasmids containing HMf-2-encoding *M. fervidus* DNA. Plasmid DNA was isolated, and the HMf-2-encoding region was subcloned and sequenced.

Electron Microscopy (EM). Various amounts of HMf were mixed with 100 ng of pUC18 DNA in 50 μl of 100 mM KCl/50 mM triethanolamine/2 mM K_3PO_4 (7) at pH 8 and the mixtures were incubated at 37°C for 10 min. The DNA–protein complexes that formed were fixed for 15 min with 0.2% glutaraldehyde before passage through a Sepharose 4B column equilibrated with 10 mM Tris·HCl/7 mM $MgCl_2$ at pH 7.5, adsorption to freshly cleaved mica, and visualization by EM as previously described (7). The glutaraldehyde fixation step was omitted, as indicated, in some experiments.

RESULTS

HMf Compacts and Stabilizes dsDNA. HMf, which constitutes ≈1% of the *M. fervidus* soluble protein, was isolated

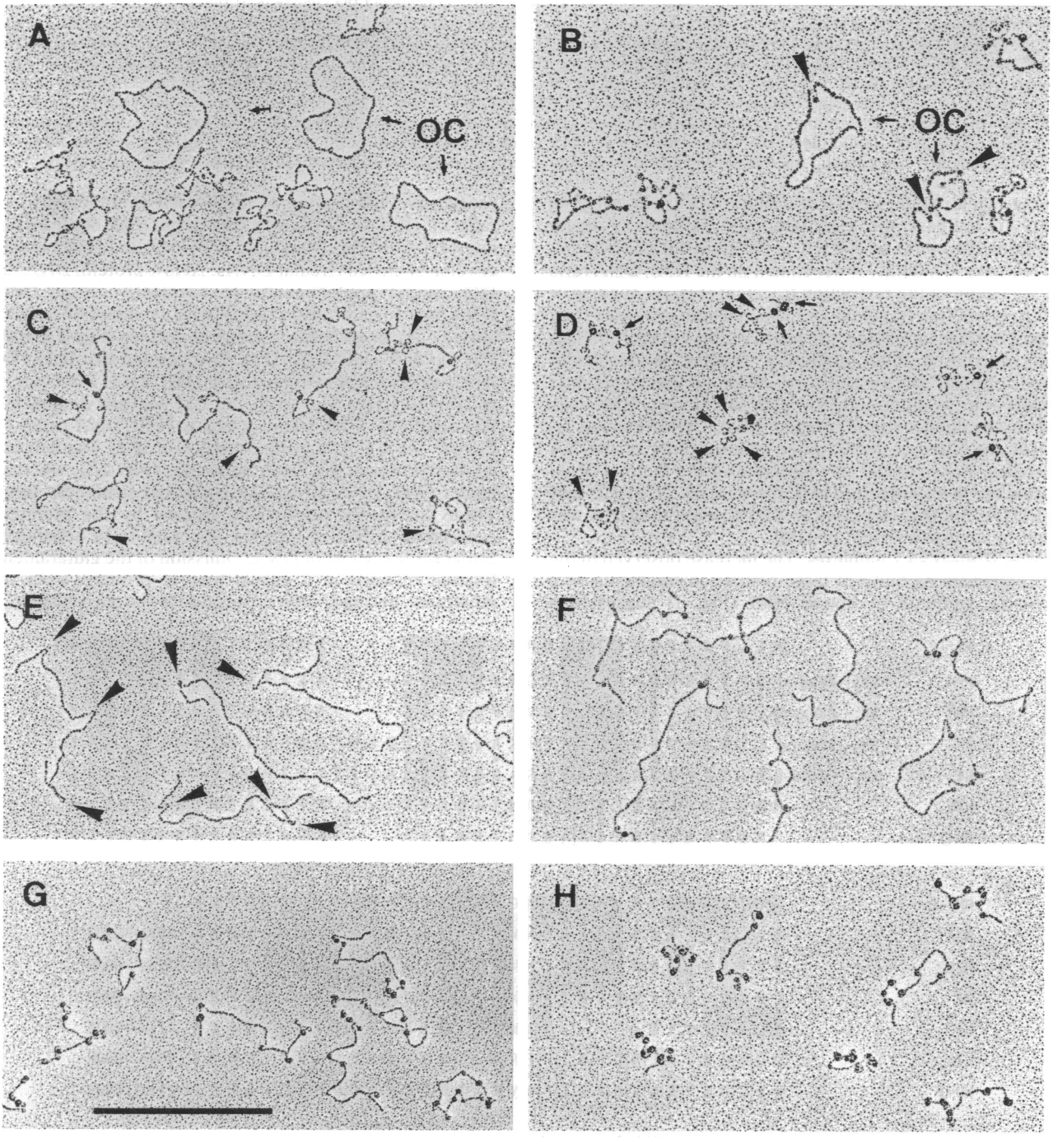

FIG. 2. Electron micrographs of HMf–pUC18 complexes. (*A* and *B*) HMf bound to a mixture of open circular (OC) and covalently closed circular pUC18 DNAs at HMf/DNA ratios (wt/wt) of 1:2 and 1:1, respectively. (*C* and *D*) HMf bound to *Sma* I-linearized pUC18 DNA, *without* glutaraldehyde fixation, at ratios of 1:2 and 2:1, respectively. (*E–H*) HMf bound, and fixed, to *Sma* I-linearized pUC18 DNA at HMf/DNA ratios of 1:5, 1:2, 1:1, and 2:1, respectively. The arrows in *C* and *D* point to compact structures similar to those seen in fixed samples; the arrowheads indicate the small DNA loops revealed by the loss of HMf. The arrowheads in *E* indicate the kinks that form at an early stage in complex formation. (Bar = 0.5 μm.)

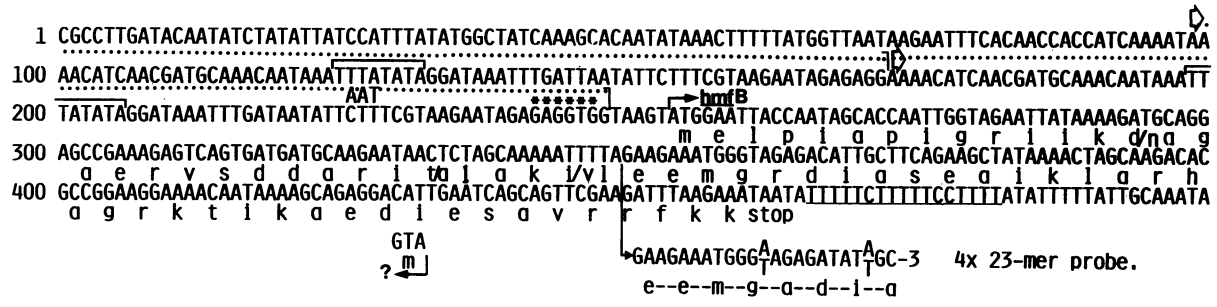

FIG. 3. DNA sequence of the gene encoding HMf-2 (*hmfB*). Microsequencing provided the sequence of the N-terminal 35 amino acids of HMf-1 and 50 amino acids of HMf-2. Amino acids are shown by lowercase one-letter symbols. The differences detected in the HMf-1 sequence from the HMf-2 sequence are shown at positions 14 (Asn for Asp), 27 (Ala for Thr), and 31 (Val for Ile). The amino acid sequence from HMf-2 used to design the 4-fold degenerate (4×) oligonucleotide probe is listed below the probe. The starts of a 74-bp direct duplication are indicated by open arrows and the dotted line above the sequence indicates the duplicated regions. Two consensus BoxA archaeal promoter sequences (9) are bracketed and a ribosome binding sequence for *hmfB* is indicated by asterisks. The oligo(dT) sequences that conform to the motif established for transcription terminators in thermophilic archaea (10) are underlined. An open reading frame located on the opposite strand from *hmfB* starts at the ATG codon indicated ? and terminates at the TAA codon indicated at positions 128–126. This open reading frame is not obviously related to sequences in the Protein Identification Resource data base (National Biomedical Research Foundation, release 17).

(average yield, ≈500 μg/g wet weight of cells) by using its ability to increase the electrophoretic mobility of linear restriction fragments of dsDNA as the assay for purification. The increase in mobility is dependent on the HMf/DNA ratio and the length of the restriction fragment (Fig. 1). Binding of HMf to restriction fragments >2 kbp in length increased their electrophoretic mobilities whereas binding to shorter fragments decreased their mobilities. Binding to covalently closed, circular DNAs had little effect on electrophoretic mobility but, as determined by EM (see below), HMf does bind to covalently closed circular DNAs. HMf binding increases the thermal denaturation temperatures of dsDNA molecules. For example, the midpoint of the thermal denaturation curve for linearized pUC18 was at 60.5°C in 0.25 mM Na_2EDTA, but this increased to 84.5°C when the DNA was bound by an equal weight of HMf.

EM of HMf–DNA Complexes. The increase observed in the electrophoretic mobility of large restriction fragments with increasing amounts of bound HMf (Fig. 1) suggested a progressive compacting of these dsDNA molecules as the ratio of HMf to DNA was increased. EM confirmed this interpretation (Fig. 2) and also demonstrated HMf binding to covalently closed circular DNAs. At low protein/DNA ratios HMf bound predominantly to covalently closed circular DNA molecules when presented with a mixture of covalently closed circular and open circular molecules (Fig. 2 *A* and *B*). The mobility-shift assay (Fig. 1) suggested that linear DNA molecules became saturated with HMf at a 1:1 (wt/wt) ratio of HMf to DNA. However, the EM analysis demonstrated that additional HMf bound at higher HMf/DNA ratios (Fig. 2 *E–H*). At a 1:5 ratio the presence of HMf was detectable as sharp kinks in the DNA molecules (Fig. 2*E*), whereas at higher HMf/DNA ratios quasispherical structures formed. These structures, presumably formed by the preferential association of HMf molecules, appeared to be separated by protein-free regions of DNA. Omission of the glutaraldehyde fixation resulted in a partial loss of proteins from these structures, during or after adsorption to the mica surface, and

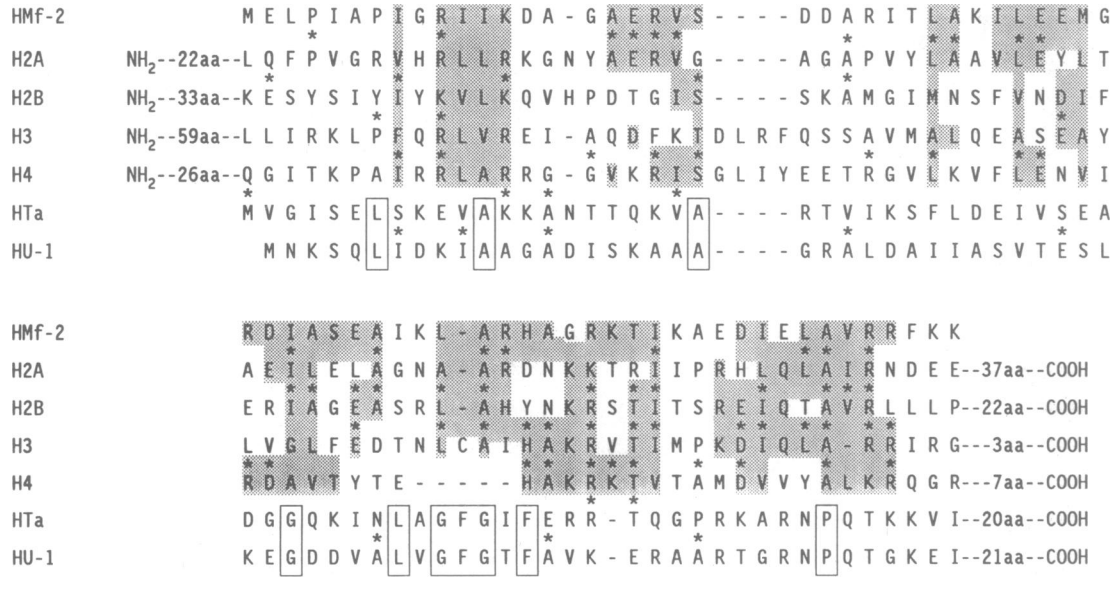

FIG. 4. Alignment of the amino acid sequences of HMf-2; eucaryal histones H2A, H2B, H3, and H4; and the prokaryotic histone-like proteins HTa and HU-1. The consensus sequences derived for the histones (22) are shown. The N-terminal residues that are not specifically identified in the histone H2A, H2B, and H4 sequences and that are not present in HMf-2 can be deleted without loss of viability in yeast (23). Amino acids found both in the HMf-2 sequence and in the other sequences are identified by asterisks. Shading emphasizes regions in HMf-2 and in the four histone sequences where clusters of identical or conserved amino acids occur. Amino acids that are conserved in several prokaryotic histone-like proteins and that are predicted to allow these proteins to form a common secondary structure are boxed in the HTa and HU-1 sequences (21). These amino acids are not conserved in HMf-2. aa, Amino acids.

revealed that the DNA molecules within these structures were constrained into small loops (Fig. 2 *C* and *D*).

Primary Sequence of the HMf-2-Encoding Gene and Flanking Regions. A 1.1-kbp region of the *M. fervidus* genome was sequenced and the HMf-2-encoding gene (designated *hmfB*) was identified. The N-terminal amino acid sequence predicted by the DNA sequence is identical to the sequence determined by microsequencing (Fig. 3). *hmfB* encodes 69 amino acids that give a calculated molecular mass of 7658 daltons for HMf-2. A 74-bp sequence that contains a consensus methanogen promoter BoxA sequence (9) is tandemly repeated directly upstream of *hmfB*. Immediately downstream of *hmfB* are tandemly repeated oligo(dT) sequences that conform exactly to the consensus sequence established for transcription terminators in thermoacidophilic archaea (10) and that were described previously in *M. fervidus* (3, 5). Despite extensive experimentation using Southern hybridization procedures, we have been unable to obtain evidence for the presence of a second gene in the genome of *M. fervidus* closely related to *hmfB*. We therefore must consider it possible that both HMf-1 and HMf-2 are products of *hmfB*.

HMf Sequences Are Extensively Conserved in Eucaryal Histones. Histone-like proteins have been isolated and characterized from several bacteria and archaea (reviewed in ref. 11). Two or more very similar polypeptides are frequently found within the same cells (12–20). Several, but not all, of these polypeptides have conserved sequences and appear to have a common structure and ancestry (11, 20, 21). Extensively studied examples are HU-1 from the bacterium *E. coli* and HTa from the archaeon *Thermoplasma acidophilum*. These are clearly related to each other and a statistically valid relationship to eucaryal histones has been demonstrated (12, 13). An alignment of HMf-2, HTa, and HU-1 with the consensus sequences derived for eucaryal histones H2A, H2B, H3, and H4 (22) is shown in Fig. 4. HMf-2 is much more similar to all the eucaryal histones than to either HTa or HU-1. Of the 69 amino acid residues in HMf-2, 19, 19, 17, and 20 are conserved in the consensus sequences of histones H2A, H2B, H3, and H4, respectively. If conservative amino acid substitutions are accepted as maintaining protein identity, then HMf-2 is 52%, 48%, 51%, and 55% identical to the consensus sequences for histones H2A, H2B, H3, and H4, respectively. In contrast, even when conservative substitutions are allowed, both HTa and HU-1 are <30% identical to HMf-2 or to any of the consensus histone sequences. They are, however, 48% identical to each other and conform to a common secondary structure (21) that is not predicted for HMf-2 (Fig. 4).

DISCUSSION

Differences in the gross anatomy of the nuclear material—namely, the presence of multiple, paired, and distinct chromosomes and a nuclear membrane—were, historically, the primary basis for the separation of eukaryotes (Eucarya) and prokaryotes (Archaea and Bacteria; ref. 1). The subsequent discovery that eucaryal chromatin is consistently organized into well defined nucleosomes appears to extend this division. Although there are many reports of histone-like proteins and EM descriptions of DNA–protein particles in prokaryotes, there is as yet no convincing evidence for a conserved nucleosome-like structure in all prokaryotic species (10, 24). The discovery that HMf is more closely related to eucaryal histones than to any known bacterial or archaeal histone-like protein is provocative. Maintaining both structural and functional integrity of macromolecules in hyperthermophiles must severely limit the number of acceptable changes in their primary sequences. Hyperthermophiles are therefore likely to exhibit very slow rates of evolutionary change. Evidence supporting this prediction has already been obtained from comparisons of the primary structures of methyl coenzyme M reductase in mesophilic, thermophilic, and hyperthermophilic methanogens (3). The similarity of HMf to eucaryal histones may also reflect a very slow rate of evolutionary change predicted by this argument for *M. fervidus*. It is, of course, tempting to speculate that a nucleosome-like organization employing HMf has also been maintained by *M. fervidus*.

This work was supported by contract N00014-86-K-0211 from the Office of Naval Research. The N-terminal sequencing of HMf-1 and HMf-2 was performed at the Harvard University Microchemistry Facility. K.S. was supported by an Ohio State University postdoctoral fellowship.

1. Woese, C. R., Kandler, O. & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4576–4579.
2. Fabry, S. & Hensel, R. (1988) *Gene* **64**, 189–196.
3. Weil, C. F., Cram, D. S., Sherf, B. A. & Reeve, J. N. (1988) *J. Bacteriol.* **170**, 4718–4726.
4. Fabry, S., Lang, J., Nierman, T., Vingron, M. & Hensel, R. (1989) *Eur. J. Biochem.* **179**, 405–413.
5. Haas, E. S., Daniels, C. J. & Reeve, J. N. (1989) *Gene* **77**, 257–263.
6. Hensel, R. & König, H. (1988) *FEMS Microbiol. Lett.* **49**, 75–79.
7. Spiess, E. & Lurz, R. (1988) *Methods Microbiol.* **20**, 293–323.
8. Heilmann, H. & Reeve, J. N. (1982) *Gene* **17**, 91–100.
9. Brown, J. W., Daniels, C. J. & Reeve, J. N. (1989) *CRC Crit. Rev. Microbiol.* **16**, 287–338.
10. Zillig, W., Palm, P., Reiter, W., Gropp, F., Pühler, G. & Klenck, H. (1988) *Eur. J. Biochem.* **173**, 473–482.
11. Drlica, K. & Rouviere-Yaniv, J. (1987) *Microbiol. Rev.* **51**, 301–319.
12. Searcy, D. G. & DeLange, R. J. (1980) *Biochim. Biophys. Acta* **609**, 197–200.
13. DeLange, R. J., Williams, L. C. & Searcy, D. G. (1981) *J. Biol. Chem.* **256**, 905–911.
14. Laine, B., Chartier, F., Imbert, M., Lewis, R. & Sautiere, P. (1986) *Eur. J. Biochem.* **161**, 681–687.
15. Choli, T., Wittmann-Liebold, B. & Reinhardt, R. (1988) *J. Biol. Chem.* **263**, 7087–7093.
16. Choli, T., Henning, P., Wittmann-Liebold, B. & Reinhardt, R. (1988) *Biochim. Biophys. Acta* **950**, 193–203.
17. Reddy, T. R. & Suryanarayana, T. (1989) *J. Biol. Chem.* **264**, 17298–17308.
18. Chartier, F., Crest, G., Laine, B. & Sautiere, P. (1989) *J. Chromatogr.* **466**, 331–337.
19. Chartier, F., Laine, B., Bèlaiche, D. & Sautiere, P. (1989) *J. Biol. Chem.* **264**, 17006–17015.
20. Chartier, F., Laine, B., Bèlaiche, D., Touzel, J.-P. & Sautiere, P. (1989) *Biochim. Biophys. Acta* **1008**, 309–314.
21. Tanaka, I., Appelt, K., Dijk, J., White, S. W. & Wilson, K. S. (1984) *Nature (London)* **310**, 376–381.
22. Wells, D. & McBride, C. (1988) *Nucleic Acids Res.* **17**, r311–r345.
23. Kayne, P. S., Kim, U.-J., Han, M., Mullen, J. R., Yoshizaki, F. & Grunstein, M. (1988) *Cell* **55**, 27–39.
24. Pettijohn, D. E. (1988) *J. Biol. Chem.* **263**, 12793–12796.