CMLS, Cell. Mol. Life Sci. 54 (1998) 1350–1364
1420-682X/98/121350-15 $ 1.50 + 0.20/0
© Birkhäuser Verlag, Basel, 1998

CMLS Cellular and Molecular Life Sciences

Review

Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome

K. Sandman*, S. L. Pereira and J. N. Reeve

Department of Microbiology, Ohio State University, 484 W. 12th Avenue, Columbus (Ohio 43210, USA), Fax +1 614 292 8120, e-mail: sandman.1@osu.edu

Received 5 August 1998; accepted 1 September 1998

Abstract. All cells employ architectural proteins to confine and organize their chromosomes, and to prevent the otherwise thermodynamically favored collapse of concentrated DNA into compact structures. To accomplish this, prokaryotes have evolved a variety of phylogenetically unrelated, small, basic, sequence-independent DNA-binding proteins that include histones in Euryarchaeota, and members of the HU family in many Bacteria. In contrast, virtually all Eukarya employ histones, and recently a metabolism-based hypothesis proposed that the eukaryal nucleus originated from a hydrogen-consuming, histone-containing Archaeon. Histones may have prevailed during the evolution of the Eukarya because of their extended interactions with DNA and, as noted, the histone fold now exists not only in histones but also as a structural motif in eukaryal transcription factors.

Key words. DNA packaging; nucleoid; macromolecular crowding; Archaea; histone; nucleosome; chromatin.

Introduction

The DNA packaging problem is usually described in terms of accommodating a long DNA molecule within a physically small space while maintaining rapid access to the confined genetic information; however, equally important is preventing the spontaneous aggregation of DNA at the high concentrations imposed by the confinement. In all cells, architectural proteins facilitate chromosome organization, compaction, replication and expression, and here we review the organization of prokaryotic chromosomes, and compare prokaryotic and eukaryal chromosomal architectural proteins. Most investigations have focused on *Escherichia coli*, and first we present a detailed description of the *E. coli* nucleoid; however, the data from other prokaryotes, both bacterial and archaeal, suggest that the *E. coli* paradigm may be only one of several solutions that have evolved to solve the packaging problem. It has been argued that the evolution of histones and DNA condensation into nucleosomes, which apparently occurred in the euryarchaeotal lineage, facilitated genome expansion and the development of Eukarya [1], and histones do now form the basis of chromosome organization in virtually all Eukarya.

[Note: Condensation is used in the DNA/chromosome literature to describe both the ordered structure of DNA under concentrated conditions [2], and the compact structure of chromosomes during cell division.

* Corresponding author.

These may not be equivalent, and we use the term only in the first sense.]

The *E. coli* paradigm

Nucleoid structure

Many complementary approaches, from light and electron microscopy (EM), biochemistry and molecular genetics have contributed to our understanding of the *E. coli* nucleoid, the irregularly shaped internal region of the cell that contains the chromosome and its associated proteins. The 4.6-Mb circular chromosome is organized into supercoiled domains, and by determining the number of DNA strand breaks required to relax all supercoiling, Worcel and Burgi [3] calculated that there were 12–80 independent domains per chromosome. This value was later refined to 33 to 53 domains by Sinden and Pettijohn [4], who noted also that the number of domains varied with growth conditions. EM of isolated nucleoids demonstrated the presence of many DNA loops extending from a central point, intertwined in plectonemic supercoils [5], and light microscopy of protease-treated chromosomes released from ethidium bromide-stained cells has revealed the presence of dense nodes from which emanate 20–50 long arms of looped DNA forming rosette structures [6]. The number of nodes per chromosome varied with growth conditions, with the highest numbers present in chromosomes from rapidly growing cells. The intense fluorescence of the nodes indicated densely packed DNA, although the presence or absence of proteins in these nodes has not been determined. Trun and Marko [7] calculated that the *E. coli* chromosome, as a random coil, would form a sphere with a ~10 μm radius, considerably larger than an E. *coli* cell, but with supercoiling into 100 separate domains the size would be reduced to approximately that of the nucleoid.

The size and stability of individual domains in the *Salmonella* nucleoid has been investigated by determining the frequency of cointegrate resolution between copies of the same transposon separated by increasing lengths of chromosomal DNA [8]. While recombination frequency decreased with increasing distance between the transposons, no fixed barriers to cointegrate resolution were detected. The boundaries of individual supercoiled domains were apparently variable and differed in different cells in the population. Selection of conditional mutants with altered numbers of domains identified gyrase and topoisomerase IV as determinants of domain boundaries, consistent with the physical barriers to supercoil diffusion in the chromosome being structures such as knots and tangles that require type II topoisomerase activity for resolution [9].

On a more localized scale, EM of cryo-preserved and freeze-substituted *E. coli*, immunogold-stained with antisera raised against RNA polymerase, topoisomerase I and HU, has identified DNA projections that extend from the nucleoid into the cytoplasm as regions active in transcription [10]. These projections are not visible by light microscopy, but their existence may explain the irregular shape observed for the active nucleoid, and why the inhibition of transcription or translation results in the nucleoid becoming spherical. The *E. coli* nucleoid apparently has a fluid structure, with chromosomal regions moving to the cytoplasm interface for transcription and association with ribosomes, or to the membrane for protein export, and DNA within these projections returns to the central mass of the nucleoid when gene expression is inhibited.

Murphy and Zimmerman [11] recently developed a lysozyme, low salt and exposure to spermidine-based protocol to isolate nucleoids intact from exponentially growing cells. Protein was present at a five fold weight excess relative to DNA. DNase I treatment of these nucleoids released only HU, *f*actor for *i*nversion *s*timulation (FIS), H-NS and the subunits of RNA polymerase, a much smaller complement of proteins than found in conventional preparations of *E. coli* DNA-binding proteins, which usually also contain nucleic acid-binding ribosomal proteins. Macromolecular crowding by polyethylene glycol or dextran was found to stabilize the isolated nucleoids [12]. The presence of these polymers that exclude water preserved the compact structure of the nucleoids in vitro at temperatures and under salt conditons that would otherwise have resulted in denaturation. Macromolecular crowding by the estimated 340 mg of RNA plus protein per milliliter of cytoplasm may stabilize nucleoids in vivo [13] and may help maintain the integrity of nucleoids in extremophiles growing under 'extreme' conditions.

The high concentration of DNA in the nucleoid (50–100 mg/ml) [13] plus the macromolecular crowding effects of the cytoplasm result in chromosomal DNA being condensed in vivo, with a volume decreased relative to that of fully hydrated B-form DNA present in dilute solutions in vitro [2]. Under such condensed conditions, DNA exhibits quite different reactivities from those observed in dilute solutions in vitro. Reaction rates are enhanced as the local concentrations of the reactants are increased [2, 14]. As noted above, a compact spherical nucleoid results when transcription or translation is inhibited, consistent with the nucleoid shape being determined by external forces from the cytoplasm.

Nucleoid proteins

HU, the most abundant *E. coli* nucleoid protein [11], appears to be evenly distributed in the nucleoid based

on fluorescence microscopy of fluorescein-labeled HU introduced into living cells [15]. The surface loops identified as HU binding sites by immuno-EM [10] were not, however, detectable in these experiments. DNA binding by HU is not sequence-specific, and with ~20,000 dimers present per 4.6 Mb genome [16], there is sufficient for one dimer to bind per 230 bp of *E. coli* DNA. In vitro, one HU dimer can bind in an ordered array every 9 bp [17], however if such close binding occurs in vivo, it must be restricted to a few sites.

HU binding introduces negative superhelical tension into a DNA molecule [18], or stated differently, HU can constrain negative superhelicity by binding to the DNA. Approximately 50% of the superhelicity of the *E. coli* chromosome is maintained by protein-DNA interactions, and although a mutant lacking HU has only minor defects in supercoiling, HU is nevertheless believed to be a major contributor to the superhelicity of the *E. coli* chromosome. Recently, the concentration of HU and the activity of topoisomerase I were shown to act in concert to modulate the superhelicity of the *E. coli* chromosome [19].

HU binding bends DNA sharply, mediating covalent circularization of short linear DNA fragments by DNA ligase in vitro [20] and functioning in vivo as an architectural factor within multiprotein complexes that catalyze transposition and DNA inversion events that require DNA looping. HU binds with high affinity to four-way junction DNAs [21], synthetic oligonucleotide structures that simulate in vitro the crossing of DNA strands that must occur during recombination. This ability of HU to recognize crossed DNA molecules may reflect a role in maintaining chromosome architecture by facilitating the movement of adjacent DNA within the nucleoid. Crossed DNA strands must occur frequently in the highly compacted nucleoid [22], and HU may act to prevent tangling.

HU preparations are populations of 20-kDa polypeptide dimers, predominantly heterodimers of α and β monomers [23], encoded by *hupA* and *hupB*; however only the $\alpha\beta$ heterodimer and the $\alpha 2$ homodimer are active in vivo [18]. Based on the crystal structure determined for the closely related homodimer of the *Bacillus stearothermophilus* protein, HUBst [24], *E. coli* HU dimers are lobster-shaped. Two long flexible β-ribbon arms extend from a globular base to contact the DNA, and each carries an array of regularly spaced arginine and lysine residues appropriately positioned to interact with six consecutive phosphate groups on one DNA strand [25]. Computer-docking simulations of the HUBst structure with B-form DNA show a reasonable fit that is further improved by widening the minor groove to introduce a kink into the DNA, consistent with the known ability of HU to bend DNA. The tips of the DNA-binding arms could not be resolved in the HUBst crystals; however nuclear magnetic resonance (NMR) molecular dynamic analyses indicate that the tips of the arms are extremely flexible, and have the potential to move and encircle the DNA [26].

Integration host factor (IHF) is a close relative of HU, and also a heterodimer of subunits encoded by *himA* and *himD*. Although IHF is abundant, present at 30 to 50% the level of HU [27], it has been thought of as a sequence-specific DNA-binding protein, with a degenerate consensus sequence. The crystal structure established for an IHF-DNA complex, however, revealed that IHF binds in a manner virtually identical to that proposed for HU, with two β-ribbon arms inserted into the minor groove introducing a sharp bend of at least 160° [28]. The arms do not contact nucleotide-specific structures, but rather interact with the phosphodiester backbone, and the consensus sequence for IHF binding apparently therefore reflects dinucleotides that readily accommodate narrowing of the minor groove. Still unexplained is the discrepancy between the high IHF abundance and the relatively low number of IHF binding sites [27], despite this including the ~100 repetitive reiterative IHF bacterial interspersed mosaic element (RIB) elements in the *E. coli* genome [29].

The other two proteins identified by Murphy and Zimmerman [11] as *E. coli* nucleoid components, H-NS and FIS, appear to modulate the expression of a wide range of genes. H-NS, a 30-kDa homodimer present at 3000 to 5000 molecules per chromosome, regulates transcription of environmentally responsive genes, sometimes acting as a repressor and in other cases as an activator [30]. Consistent with a role as an architectural chromosomal protein, H-NS has been localized to the nucleoid in vivo by immuno-EM [31], and shown to constrain supercoils [32] and to bind preferentially to bent DNA sequences in vitro [33]. H-NS exhibits two types of DNA binding in vitro dependent on the protein:DNA ratio; however, the preferred binding to curved DNAs in which the precise geometry of the curve is important [34], which occurs at subsaturating H-NS levels in vitro, is likely to be the physiologically significant form in vivo [32]. With an appropriately shaped DNA sequence providing an initiation site, cooperative binding of H-NS dimers or higher oligomers could occur at defined locations along the *E. coli* chromosome. The *E. coli stpA* gene encodes a protein with 58% amino acid sequence identity to H-NS that appears to have overlapping functions with H-NS in gene regulation and unique functions as an RNA chaperone [35].

FIS is a 22-kDa homodimer whose abundance varies with the growth phase, with 25,000 to 50,000 molecules per genome present during exponential growth but fewer than 100 molecules per genome present in station-

ary phase. FIS also functions as a transcription activator or repressor, with one or more FIS dimers binding to specific sites, generating severely bent DNA sequences within nucleoprotein structures. One helix-turn-helix (HTH) DNA binding motif has been identified per FIS monomer in a homodimer [36], and FIS preferred binding sites reflect not sequence-specific binding, but rather the ability of a DNA sequence to accommodate the bending induced by FIS binding [37], as now similarly established for IHF preferred binding sites.

The effects of depriving *E. coli* of these nucleoid proteins, singly and in combinations, has been studied by introducing null mutations into the encoding structural genes. All single and double mutants remain viable, although they exhibit highly pleiotropic phenotypes, consistent with the involvement of the chromosomal proteins in many different reactions, and with partial redundancy in facilitating essential functions. A viable triple mutant, lacking HU, IHF and H-NS, could not be constructed [38].

Growth phase regulation has been found for all *E. coli* nucleoid proteins, consistent with the structure and superhelicity of the nucleoid being very different in growing versus stationary phase cells. Murphy and Zimmerman were, in fact, able to purify nucleoids only from exponentially growing *E. coli* cells [11]; stationary phase cells were refractory to their technique. HU abundance increases on entry into stationary phase, and whereas $\alpha 2$ homodimers predominate in logarithmically growing cells, increased synthesis of β subunits occurs at the onset of stationary phase resulting in monomers reassociating and cells containing predominantly $\alpha\beta$ heterodimers [18]. FIS binds to the *hupA* and *hupB* promoters, stimulating *hupA* transcription and repressing *hupB* transcription [39], but FIS abundance decreases dramatically on entry into stationary phase which results in *hupB* expression. IHF and H-NS also increase 5 to 10-fold in abundance on entry into stationary phase [27, 40]. The cellular factors that translate changes in nutrient availability into global changes in *E. coli* gene expression, namely the stringent response and the *rpoS*-encoded sigma factor, participate in regulating nucleoid protein abundance [39, 41].

Labeling the *E. coli* origin of replication in vivo with green fluorescent protein (GFP) has revealed that nucleoid organization is coordinated with the cell cycle [42]. In nucleoids with a single origin of replication, the origin is located near one cell pole; however, immediately after replication, the two origins separate and maintain positions at opposite poles of the nucleoid until cell division occurs. There must therefore be a structure within the nucleoid that directs the segregation of origins, and there is evidence that *mukB* encodes a component of a partitioning mechanism that segregates daughter chromosomes in *E. coli*. MukB is a myosin-like protein and mutations in *mukB* result in a dispersed nucleoid phenotype [43]. Multiple rounds of genome replication occur concurrently in *E. coli* cells during rapid exponential growth, and any partitioning structure must therefore be incorporated within the nucleoid. As MukB appears to be a major component of this structure, it could legitimately be considered a chromosomal architectural protein.

In summary, the *E. coli* chromosome is a dynamic structure divided into supercoiled domains with boundaries that change with the growth rate and cell metabolism. DNA compaction results from the exogenous pressure of macromolecular crowding from the cytoplasm, and from supercoiling introduced by architectural proteins and topoisomerase activities. Exponentially growing and stationary phase cells have different nucleoid structures, biochemistries and levels of gene expression that are reflected in the presence or absence of specific architectural proteins.

Other bacterial nucleoids

Hinnebusch and Bendich [6] state that *Agrobacterium tumefasciens*, *Streptomyces lividans* and *Pyrococcus* strain ES4 (fig. 1) have rosette-shaped chromosomal structures similar to *E. coli*, but describe the presence of a completely different structure in two *Borrelia* species. Unlike the dense nodes with radiating DNA strands observed in *E. coli*, these spirochetes have a network of DNA strands extending throughout the cell. *Borrelia burgdorferi* has 8–11 copies of a linear chromosome and many plasmids and is therefore polyploid. Although specific linking structures have not been visualized by microscopy, attempts to separate individual replicons from the network have been unsuccessful, suggesting the existence of such replicon-associating structures. A similar description was reported for the nucleoid of *Epulopiscis* [44] and *Chlamydia* [45], also intracellular symbionts.

Most microbial nucleoid proteins appear to have very restricted phylogenetic distributions, H-NS and FIS, for example, are present in only *E. coli* and close relatives, but HU homologs are strikingly widespread. The sequences of HU homologs, from diverse bacteria, mitochondria and chloroplasts can be aligned readily, without gaps, on the basis of conserved hydrophobic residues that generate the global fold and conserved basic residues that interact with DNA phosphates [16, 46]. Most HU homologs are homodimers; however, two similar but not identical polypeptides are present in enteric bacteria [46], and in *E. coli*, all three possible polypeptide dimers can form [18]. The $\alpha\beta$ heterodimer that predominates in stationary phase *E. coli* cells con-

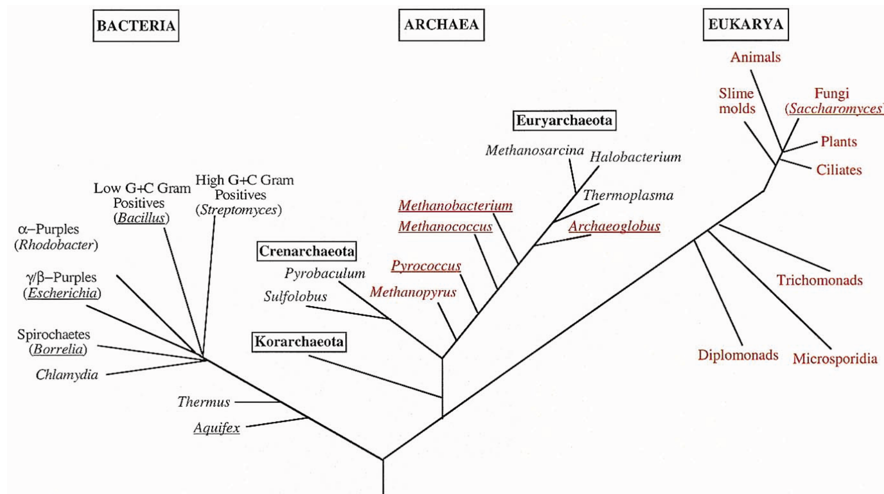

Figure 1. Universal phylogenetic tree in rooted form, based on rRNA sequence comparisons [143] of organisms described herein. Branch lengths do not represent exact evolutionary distances, but the branching order is correctly represented. The names of organisms with histones are displayed in red, and those whose genome has been sequenced are underlined.

fers greater long-term survival than is exhibited by strains lacking either α, β, or both HU polypeptides. A phylogenetic analysis of HU sequences suggests that horizontal gene transfer may initially have generated a strain with two nonidentical copies of HU-encoding genes [46]. The monomer-monomer interface was apparently compatible, and a long-term survival advantage was apparently conferred by the resulting heterodimer that would have exerted the selective pressure needed to maintain both genes.

Bacillus subtilis, often investigated as a representative Gram-positive Bacteria, has a nucleoid visually indistinguishable from that of *E. coli* [44], but other than HBsu, an HU-homolog, *B. subtilis* has nucleoid proteins unrelated to those of *E. coli*. *B. subtilis* vegetative cells synthesize HBsu as an essential nucleoid-associated homodimer [47] that cannot be replaced by the prophage-encoded HU homolog YonN. As *B. subtilis* cells enter stationary phase, HBsu expression levels drop [48], probably as the result of the completion of replication, and sporulation begins. It is not known if HBsu is normally present in the developing spore, but spores lacking HBsu do germinate, although with a slight delay [48]. As sporulation proceeds, a new family of DNA-binding proteins, known generically as the small acid-soluble spore proteins (SASPs), is synthesized in the prespore compartment. SASPs bind to the spore chromosome, and although gene expression continues during spore formation, SASP binding does change the architecture of the nucleoid, and eventually SASP binding saturates the chromosome and silences genes expression [49]. Labeling with GFP has revealed that *B. subtilis* also has a structure within the nucleoid that segregates origins of replication [50–52], with chromosome partitioning directed by SpoOJ and Smc [50, 53]. Smc is a homolog of MukB in *E. coli* [54]. A transition state regulator, AbrB, has also been proposed to be a nucleoid protein on the basis of having an in vivo expression pattern similar to FIS in *E. coli*, and because *abrB* null mutations have pleiotropic effects on the expression of many unrelated genes [55].

Chlamydia, an intracellular parasite, similarly has two cellular forms with different nucleoid structures [56]. The metabolically active reticulate body (RB) has a nucleoid composed of loosely packed fine fibrils that extend throughout the cytosol, as described above for *Borrelia*, whereas the nucleoid of the metabolically inert but infectious extracellular elementary body (EB) appears as a condensed, electron-dense mass [45]. The reorganization of the chlamydial nucleoid that occurs during RB to EB transitions is mediated by Hc1 and Hc2 [57, 58], two related small, basic proteins that share sequence homology with the eukaryotic linker histone H1 [57, 59–61]. Hc1 and Hc2 also compact the *E. coli* chromosome in vivo following their expression as recombinant proteins [62, 63]. Hc1 binds cooperatively to DNA in vitro, generating highly condensed spherical bodies, and forming these structures in vivo probably causes the decline in transcription, translation and replication that coincides with Hc1 synthesis in *E. coli* [62, 64–66]. Although the other protein components of the chlamydial nucleoid have not been identified, chlamydial genomic DNA encodes an IHF homolog [67].

Archaeal chromosomes and architectural proteins

Archaea and Bacteria have similar cellular dimensions and DNA content, and therefore face essentially the same DNA-packaging problems. However most investigations of archaeal chromosome structure have originated as attempts to identify the factors that protect DNA in vivo from the extreme conditions of temperature, pH and osmolarity that are 'normal' growth conditions for archaeal extremophiles. Despite this mindset, as macromolecular crowding from the cytoplasm presumably exerts a condensing force on archaeal nucleoids, and as renaturation rates for complementary DNA strands are dramatically increased in this condensed DNA [68], macromolecular crowding could be sufficient to overcome the denaturing effects of high temperatures on the genomes of hyperthermophiles.

Bacterial chromosomes appear always to be negatively supercoiled in vivo, whereas topology measurements of archaeal plasmids suggest that alternative states of supercoiling may exist in Archaea. Negative supercoiling, comparable to that in Bacteria, was found for plasmids in mesophilic Archaea but plasmids from thermophilic Archaea were relaxed, which was not the case for plasmids from thermophilic Bacteria [69]. Although plasmid topology may not necessarily reflect chromosome supercoiling, these results do suggest that DNA supercoiling patterns may be different in thermophilic Archaea.

A light microscopy study of *Sulfolobus* by Poplawski and Bernander [70] revealed that nucleoid structures are different in exponentially growing versus stationary phase cells. During exponential growth, the nucleoid is a single, organized structure with regions exhibiting different levels of fluorescence following 4,6-diamino-λ-phenylindole (DAPI) staining, whereas stationary phase nucleoids have evenly dispersed staining and occupy a larger volume in the cell. Flow cytometry revealed that stationary phase *Sulfolobus* cells have two copies of the chromosome, which were visibly detected as two separate nucleoids in some cells [71].

Most archaeal chromosomal proteins have been studied only following their extraction and purification, as genetic techniques and conditions for manipulating the growth of these organisms are, in most cases, very limited precluding in vivo experiments. Except for HTa (see below) and the Smc partitioning proteins, chromosomal architectural proteins purified from *Archaea* are not related to the known bacterial proteins, and genes encoding homologs of the *E. coli* proteins described above are not present in sequenced archaeal genomes.

Thermoplasma acidophilum, a thermoacidophile, contains an abundant DNA-binding protein, HTa, that stabilizes DNA against thermal denaturation in vitro [72], and which based on its amino acid sequence [73], is the only archaeal member of the HU family identified to date [46]. An initially proposed relationship between HTa and eukaryal histones [74] has not subsequently been supported by structural studies. The short region of histone H2A that does have sequence similarity to a region of HTa is within the HSH histone fold of H2A [75] (see below), and this has a three-dimensional (3D) structure totally unrelated to that of HU.

MC1 was identified as the most abundant chromosomal protein in *Methanosarcina barkeri* [76], and closely related homologs of this 11-kDa protein have now also been characterized from other Methanosarcinaceae [77] and from halophiles [78]. EM of MC1-DNA complexes has revealed that MC1 binding has no effect on the contour length but does introduce sharp kinks into the DNA molecule [79]. MC1 binds preferentially to a specific sequence [80], but in light of the explanation of similar observations for the *E. coli* architectural proteins IHF and FIS, the preference may reflect the ability of a particular DNA sequence to adapt to MC1-induced bending rather than specific interactions with a base-pair sequence.

Mixtures of small, acid-soluble DNA-binding polypeptides have been purified from *Sulfolobus* species, and a subset of these 7-kDa proteins, named Sac7, Sso7 or Ssh7 depending on the *Sulfolobus* species of origin, have been studied intensively. Binding by these proteins introduces negative superhelicity [81, 82] and promotes reannealing of complementary DNA strands in vitro [83]. HSNP-C′, identical to Sac7d [84], has been localized to the nucleoid in vivo by immuno-EM [85]. Based on NMR studies, Sac7d and Sso7d in solution [86, 87] are rich in β-sheet structure and contain several helices, although the length and packing of the C-terminal helix differs in the two proteins [86]. High-resolution structures of Sac7d and Sso7d complexed with short DNA molecules reveal binding in the minor groove, and inserting amino acid side chains into the DNA results in DNA bending with the protein wrapped around the outer circumference of the bend [88, 89].

Archaeal histones were first identified in the hyperthermophile *Methanothermus fervidus*, and purified on the basis that archaeal histone-DNA complexes migrate faster during agarose gel electrophoresis than the protein-free DNA molecules [90]. HMf (the histone from *M. fervidus*) was localized to the nucleoid by immuno-EM [85], and EM of archaeal histone-DNA complexes assembled in vivo [91] and in vitro [90, 92] has shown the beads-on-a-string configuration that is classically described for nucleosomes in eukaryal chromatin [93] (see below). Deproteinization demonstrated that the DNA was looped around a protein core [90], consistent with the decrease in contour length observed by EM for DNA following archaeal histone binding [94].

Figure 2. Alignment of archaeal histone sequences and eukaryal core histone consensus sequences [144]. The positions and lengths of α helices determined for HMfB [113] are indicated above its sequence. Vertical lines mark the positions of hydrophobic residues that specify monomer-monomer interactions. An asterisk indicates a stop codon in the DNA sequence, and a hyphen indicates a gap introduced to improve the alignment. The numbers in parentheses preceding and following the eukaryal sequences indicate the number of residues outside the histone fold domain that were omitted in this alignment. HMf histones are from *Methanothermus fervidus* [90, 100]; MT from *Methanobacterium thermoautotrophicum* [100, 107]; HFo from *Methanobacterium formicicum* [101]; HTz from *Thermococcus zilligii* [103]; HPy from *Pyrococcus* strain GB3a [105]; AF from *Archaeoglobus fulgidus* [108]; MJ from *Methanococcus jannaschii* [106]; Mka from *Methanopyrus kandleri* [104]; HMv from *Methanococcus voltae* [128].

While changes in nucleoid structure have not been directly investigated in *M. fervidus*, the polypeptide composition of HMf preparations purified from *M. fervidus* does change with growth phase [92], in a pattern similar to that observed for HUα and HUβ in *E. coli* [18]. HMf preparations contain two closely-related 7.5-kDa polypeptides, HMfA and HMfB, that form both homodimers and heterodimers, and HMfA predominates in HMf preparations isolated from exponentially growing *M. fervidus* cells, whereas the HMfB abundance increases with the transition into stationary phase. Binding by HMfA and HMfB changes DNA topology differently in vitro, with HMfB binding introducing greater compaction into the DNA [92].

All histones are polypeptide dimers in solution [92, 95], but tetramers form during DNA wrapping. Micrococcal nuclease (MN) digestion of *M. fervidus* chromatin and of HMf-DNA complexes assembled in vitro, demonstrated the presence of tetramer-containing complexes that protect ~60 bp from MN digestion [91, 96]. There is sufficient HMf in *M. fervidus* cells to wrap the chromosome entirely in structures that contain a histone tetramer and ~60 bp of DNA (S. L. Pereira and J. N. Reeve, unpublished).

Topoisomerase I-mediated topology assays have demonstrated that HMf binding to circular DNAs can introduce either negative or positive superhelicity, depending on the histone to DNA ratio [97]. As the ratio in vivo [98] is approximately that at which the negative to positive supercoiling transition occurs in vitro, both topologies may be introduced in vivo, depending on local variations in HMf concentration and torsional stress differences in different regions of the chromosome.

The amino acid sequences of HMfA and HMfB are related to those of the four eukaryal nucleosome core histones, H2A, H2B, H3 and H4, and this facilitated an alignment of these eukaryal histone sequences [90] (fig. 2), that has now been validated by 3D structural studies [75, 99] (see below). Histones and histone-encoding genes have now also been characterized from several other members of the Euryarchaeota [100–109] (fig. 1), both mesophilic and thermophilic species, indicating that archaeal histones have functions that extend beyond the protection of DNA in hyperthermophiles. Histones are the predominant chromosomal architectural proteins in these Archaea.

Eukaryal chromatin

Although the eukaryal nucleus is larger than most prokaryotic cells, eukaryal genomes are also much larger and DNA confinement and gene expression from packaged DNA is still a challenge. Eukaryal chromosomes, visible during mitosis as discrete structures by light

microscopy, contain several levels of compaction and organization, but at the most basic level, eukaryal DNA is organized by small, basic, abundant proteins, histones, into nucleosomes. Each nucleosome contains 146 bp of DNA wrapped around a histone octamer [75], and adjacent nucleosomes result in the 'beads-on-a-string' structure observed for negatively stained chromosomes by EM [93]. As eukaryal chromosomes are linear, any supercoiling in these molecules must result from protein binding, and is primarily a result of DNA wrapping during nucleosome assembly by histones.

The nucleosome core histones, H2A, H2B, H3 and H4, associate to form H2A-H2B and H3-H4 heterodimers, each of which is present twice per nucleosome. The two H3-H4 heterodimers assemble to form a central tetramer that is flanked by H2A-H2B heterodimers creating a histone octamer in which the histone dimers are arranged in a linear protein supercoil. Histones are extremely conserved presumably because of severe constraints imposed by the essential multiple interactions with DNA, with one another, and with other chromosome-associated proteins. H2A, H2B, H3 and H4 monomers all have a common, globular domain, termed the histone fold, consisting of three amphipathic α helices separated by short loops and β-strand regions [99]. The histone fold exists only in monomers assembled into dimers, as each α-helix has an extensive hydrophobic surface that must be buried by dimerization to form and maintain the native structure of the histone fold [110–112].

Comparison of archaeal and eukaryal histones

Figure 2 shows an alignment of archaeal histone sequences with consensus sequences for the eukaryal nucleosome core histones. Only a short gap is required to align the sequences, and conserved residues are readily apparent. In pairwise comparisons based on this alignment, H2A, H2B, H3 and H4 are more similar to HMfB than to each other, consistent with the archaeal sequence resembling a common ancestor of the four eukaryal core histones. The conclusion that archaeal and eukaryal histones are homologs, based initially on amino acid sequence similarities [90], has been validated by high-resolution 3D structures [113–115]. The arrangement of the six α helices in an $(HMfB)_2$ dimer (fig. 3) and in the histone fold domains of an (H3-H4) heterodimer is virtually identical, with a root-mean-square deviation of 2.1 Å for the α-carbon atoms [115]. Comparing contemporary archaeal and eukaryal histone structures may provide clues to the molecular evolution events that must have occurred during the transition from archaeal to eukaryal chromatin, probably an important step in the evolution of eukaryotes from prokaryotes.

Archaeal histones range in length from 66 to 69 residues and are essentially only histone folds, whereas the eukaryal histones are larger, with N- and C-terminal amino acid sequences that extend beyond the histone fold. These regions are not essential for nucleosome organization, and can be proteolytically removed without loss of structure [116]. They extend beyond the core of the nucleosome and play roles in gene expression and contribute to the higher order structure of eukaryal chromatin [117]. H3 and H4 extensions contact adjacent nucleosomes, facilitating tighter packing within chromatin fibers, and provide sites for posttranslational modifications that are known to be important in gene regulation.

As noted above, HMfA and HMfB form both homodimers and heterodimers, whereas the eukaryal histones form only heterodimers, and always with the same partner. The histone monomer-monomer interface within a dimer is stabilized by extensive hydrophobic interactions, with all six α helices in a dimer contributing to the hydrophobic core. Virtually all of the residues involved in dimer stabilization are identical in HMfA and HMfB, and HMfA and HMfB monomers therefore have the same hydrophobic surface. Using the coordinates from the crystal structures of HMfA and HMfB homodimers, an (HMfA-HMfB) heterodimer can be generated in silico (K. Decanniere and U. Heinemann, unpublished), although heterodimers cannot be biochemically purified in the absence of homodimers (R. A. Grayling and J. N. Reeve, unpublished). Archaea have different numbers of histone-encoding genes, ranging from two in *M. fervidus* and other organisms to five in *Methanococcus jannaschii* of which three are chromosomal and two are plasmid-encoded [106]. The hydrophobic residues predicted to be at the monomer-monomer interface are well conserved in all five *M. jannaschii* polypeptides, and therefore most dimer types presumably can form, although some part-

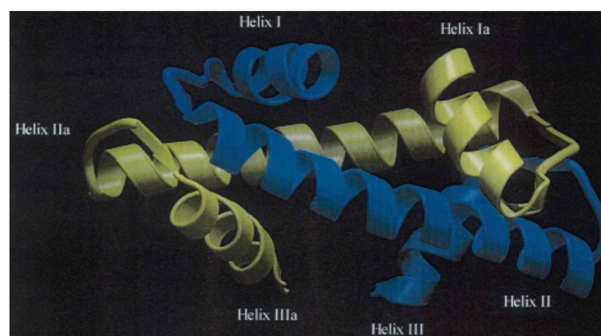

Figure 3. Structure of an $(HMfB)_2$ dimer [113], with one polypeptide chain in blue and the other in yellow. Helices are labeled I, II and III from the N-terminus.

nerships may be fixed. Different dimers and therefore tetramers could contribute significantly to local variations in chromosome architecture and gene expression. Hydrophobic residues are also present at the same locations in the histone folds of the eukaryal histones (fig. 2), but the actual residues present vary, consistent with H2A-H2B and H3-H4 dimers having monomers with complementary but not identical surface shapes, and explaining why only heterodimerization with the appropriate partner is feasible.

Although histones are dimers in solution, both archaeal and eukaryal histones assemble into higher polymers to wrap DNA. The crystal structure of the eukaryal nucleosome [75] reveals that dimer-dimer interactions occur in four-helix bundles formed by the $\alpha 2$ and $\alpha 3$ helices. Such interactions occur at three sites in the nucleosome, between the two H3 molecules in the central $(H3\text{-}H4)_2$ tetramer, and between H4 and H2B at the interface of each H3-H4 dimer with a flanking H2A-H2B dimer. The H3 interactions in the central tetramer form a more stable complex than is formed between the central tetramer and the flanking H2A-H2B dimers. Consistent with this, the $(H3\text{-}H4)_2$ tetramer remains intact under conditions where the H2A-H2B dimers dissociate from the octamer [75], for example, during transcription. It has been proposed that the $(H3\text{-}H4)_2$ tetramer interface could exist in two configurations, which would result in the DNA being constrained in either a positive or a negative supercoil [118]. This would require flexibility and a reorientation of the dimer-dimer interface within the $(H3\text{-}H4)_2$ tetramer [119] that apparently is not feasible at the (H2A-H2B)-(H4-H3) interface. The association of H2A-H2B heterodimers with the $(H3\text{-}H4)_2$ tetramer may lock the tetramer into the negatively supercoiling configuration. A comparison of the H3 residues identified as responsible for tetramer formation in an $(H3\text{-}H4)_2$ tetramer with the HMfB sequence suggests that a virtually identical four-helix bundle could form within an $(HMfB)_4$ tetramer [115].

Cross-linking studies have confirmed that archaeal histones form tetramers, but no higher order structures have been detected within archaeal histone-DNA complexes assembled in vivo or in vitro [91, 120]. As archaeal histones can constrain DNA in either positive or negative supercoils in vitro, depending on the protein:DNA ratio [97], possibly dimer-dimer reorientation does occur within a four-helix-bundle formed by two $\alpha 2$ and two $\alpha 3$ helices in an archaeal histone homotetramer [121].

The data available to date suggest that archaeal nucleosomes resemble the structure formed by the central $(H3\text{-}H4)_2$ tetramer within a eukaryal nucleosome [122]. The histones have similar structures, and both structures contain histone tetramers that protect $\sim 60-70$ bp of DNA from MN digestion. Although nucleosome-DNA contacts are not sequence-specific, some DNA sequences do inherently have shapes that facilitate wrapping and therefore nucleosome assembly, and it is the $(H3\text{-}H4)_2$ tetramer that recognizes and responds to this positioning information [123, 124]. One strong positioning element, six or more tandem repeats of the trinucleotide CTG [125], has also been shown to be recognized by archaeal histones as an archaeal nucleosome positioning sequence [145].

Histone evolution

A helix-strand-helix (HSH) structure occurs twice in every histone monomer [126], and the evolution of the contemporary histone fold probably therefore began with the duplication of a sequence that encoded this motif. The resulting polypeptide would have had to homodimerize to shield the hydrophobic surface. While this HSH duplication is predicted by the 3D structure of contemporary histones, it is no longer evident in their amino acid sequences, which must subsequently have undergone extensive divergence. Duplication of this single ancestral histone gene and divergence of the duplicated copy to encode a modified but still functional histone would have resulted in a situation similar to that now found in *M. fervidus*. Two closely related histone monomers (84% identity) are present that can assemble into three dimers [92], and potentially into six different tetramers. After gene duplication, a competition may occur between gene loss, due to redundancy, and maintenance of both genes due to gained function. The close similarity of the two histone genes retained by *M. fervidus* suggests that very few changes were sufficient to result in gene products with altered functions that provided a selective advantage sufficient for their maintenance.

Mutations that resulted in changes in the hydrophobic residues at the monomer-monomer interface must have led to the transition from two genes encoding monomers that could form both homodimers and heterodimers, to two genes encoding polypeptides that formed only heterodimers. Changing the length or shape of a hydrophobic side chain would have changed the packing in the interior of both homodimers and heterodimers, and changes in a potential partner might then initially have favored, and eventually mandated heterodimer formation. A similar scenario may account for the situation in *Methanopyrus kandleri*. This hyperthermophilic methanogen has only one histone gene, but this encodes two tandemly arranged histone folds [104] which must assemble into a structure equivalent to an $(HMfB)_2$ dimer. The hydrophobic residues in the his-

tone folds differ from one another and from the residues in other archaeal histones, consistent with the hydrophobic surface presented by each histone fold having a unique shape. Although *M. kandleri* has lost the potential to generate three different dimers from two separate histone gene products, a change in the residues at one of the two potential tetramer interfaces, favoring the formation of only one type of tetramer, might have provided a selective advantage. The residues at the positions that direct tetramerization (or dimerization in this case) are different in the two histone folds of the *M. kandleri* protein, consistent with only one orientation for four-helix-bundle formation.

As all Eukarya appear to have H2A, H2B, H3 and H4, the genes that encode these different histones presumably evolved before the divergence of the Eukarya. Two alternative possibilities are the duplication of an already diverged pair of histone genes followed by further sequence divergence, or the sequential duplication and divergence of one histone gene. The absence of the same five residues in $\alpha 1$ of H4 and H2A, which are present in H3 and H2B, argues for the first alternative [122]. However, as a single recombinant archaeal histone synthesized in *E. coli* from one histone gene can form homotetramers that assemble archaeal nucleosomes [92], the suggestion that all four core histones had to have evolved before nucleosome assembly could occur [127] no longer seems valid. Archaeal histone homopolymerization itself poses an interesting problem. There are two $\alpha 2$ and $\alpha 3$ helices in a homodimer, each of which should be equally capable of interacting with the corresponding regions of a second homodimer to form a four-helix bundle. Therefore, with two identical potential interaction sites on each dimer, polymerization should not stop at the tetramer stage. Perhaps, in a manner analogous to the role proposed for the $(H3\text{-}H4)_2$ tetramer reorientation, the orientation of the dimer-dimer interface in an archaeal histone tetramer either permits or prevents additional polymerization.

Divergence of the histone fold

Histone fold-encoding genes have continued to duplicate and diverge within both the Archaea and the Eukarya, although it is unclear whether the histone folds now present in some larger proteins still wrap DNA or function only in directing specific heterodimer formation. In addition to five bona fide histones, the *M. jannaschii* genome also encodes a histone fold-containing polypeptide, MJ1647, with an $\sim$30-residue C-terminal extension (fig. 2). Functions for MJ1647, and for HMvA, a homolog in *Methanococcus voltae* [128], are unknown, but recombinant MJ1647, synthesized in *E. coli*, does form homodimers and compacts DNA (W. Li, K. Sandman and J. N. Reeve, unpublished).

Multiple eukaryal transcription factors, including activators CBF [129], TAFII of transcription factor IID [130, 131], PCAF [132] and SPT3 [131] and the repressor Dr1-DRAP [133], contain histone folds within larger structures. These proteins bind DNA, but in each case through an association with a sequence-specific DNA-binding protein. Possibly the accessory protein provides sequence recognition function, and forms a DNA-protein complex to which the histone fold components bind preferentially [95]. The relative positions of the DNA-binding elements in the histone folds of the eukaryal histones and in the TAFII tetramers are very different, and therefore the same types of DNA-protein interactions may not occur [115].

Additional putative histone folds have been identified [134] using the conserved sequences and arrangement of three α helices of eukaryal histones as a motif identification tool. An important aspect of the histone fold, however, is mandatory dimerization with a second histone fold [110–112], and monomeric histone folds are unknown. A practical test of the existence of this fold might therefore include identifying the dimerization partner. While some of the proteins identified by this motif-searching approach are known to interact with other polypeptides and have DNA-metabolizing activity, for example, son-of-sevenless and Tn3 resolvase [135–138], others exist as monomers, for example, *Sulfolobus* DNA polymerase [139, 140]. These seem unlikely to have a histone fold.

Conclusions

The essential functions of architectural chromosomal proteins

Chromosomal architectural proteins are abundant proteins that maintain the macromolecular configuration of the chromosome. On a local scale, they may distort and dictate the path of short stretches of DNA, and on a larger scale, under the crowded solution conditions in the nucleoid, they must maintain DNA compaction and prevent aggregation. Having chromosomal architectural proteins is apparently essential; however, their structural diversity in prokaryotes suggests that there are multiple solutions to the DNA-packaging problem. Chromosomal architectural proteins have common features, small size and a structure that positions basic residues appropriately to interact with DNA; however, only one structure, the histone fold, dominates in the Eukarya. HU dimers contact six phosphates on each DNA strand over a 9–10 bp region [25], and the archaeal Sac7-type of proteins contact 8 bp [88, 89]; however, archaeal histones assemble structures that constrain 60–70 bp [91, 96]. Possibly it was this ability to

organize a much larger region of DNA into a compact and reversible nucleoprotein complex that led to the domination of histones and a structure that became the $(H3\text{-}H4)_2$ core of the eukaryal nucleosome.

The origin of histones in Eukarya

Phylogenetic trees based on ribosomal RNA (rRNA) sequences group Archaea and Eukarya on one branch (fig. 1), and indicate that these two groups share a common ancestor after the separation of Archaea and Bacteria. Consistent with this, many features of the central dogma of molecular biology, most notably the components of the transcription machinery, are similar in Archaea and Eukarya, and different from their bacterial counterparts. In this regard, histones are present throughout the Eukarya, and are present in Archaea but only in one branch, namely the Euryarchaeota (fig. 1), and a recent proposal for the origin of the eukaryal nucleus would explain this distribution of histones [141]. Martin and Müller, in the 'hydrogen hypothesis,' propose that Eukarya originated from a symbiotic association of an H_2-generating α-proteobacterium and an H_2-consuming methanogen, similar to symbioses observed today in some protists [142]. These two microorganisms became linked in an anaerobic environment on the basis of their shared hydrogen metabolism and gene transfer from the α-proteobacterium to the methanogen then expanded their metabolic interdependence. The association became cemented and irreversible following transfer to an aerobic environment. The methanogen was then absolutely dependent on the metabolism acquired from the bacterium for survival. The eukaryal nucleus in the hypothesis is derived from the methanogen and most cellular metabolism from the bacterium, which explains why archaeal and eukaryal information storage and expression systems have common features, including the use of histones for genome packaging, while most eukaryal metabolism has a bacterial origin. During the transition from prokaryote to eukaryote there was a large increase in DNA content, and Minsky et al. have argued that only a histone-based system of chromosome organization could have accommodated this increase [1]. Histones organize DNA around a protein core, insulating and preventing the collapse of individual DNA strands within highly concentrated DNA into irreversibly condensed forms. Having an expandable system of DNA packaging was obviously essential for the success of eukaryal evolution, and if other symbiotic associations existed between non-histone-containing prokaryotes, their subsequent evolution and expansion may have been limited by an inappropriate chromosomal architecture.

Acknowledgements. The authors thank Arnold Bendich for helpful discussions, Pat Higgins and Purificacion López-García for sharing manuscripts prior to publication, and James R. Brown for assistance in generating figures. Archaeal histone research is supported in the authors' laboratory by NIH grant GM53185.

1 Minsky A., Ghirlando, R. and Reich Z. (1997) Nucleosomes: a solution to a crowded intracellular environment. J. Theor. Biol. **188:** 379–385
2 Bloomfield V. A. (1996) DNA condensation. Curr. Opin. Struct. Biol. **6:** 334–341
3 Worcel A. and Burgi E. (1972) On the structure of the folded chromosome of *Escherichia coli*. J. Mol. Biol. **71:** 127–147
4 Sinden R. R. and Pettijohn D. E. (1981) Chromosomes in living *Escherichia coli* cells are segregated into domains of supercoiling. Proc. Natl. Acad. Sci. USA **78:** 224–228
5 Pettijohn D. E. (1996) The Nucleoid. In: *Escherichia coli* and *Salmonella typhimurium*, pp. 158–166, Neidhardt F., Curtiss R., Ingraham J. L., Lin E. C. C., Low K. B., Magasanik B. et al. (eds.), ASM Press, Washington, DC
6 Hinnebusch B. J. and Bendich A. J. (1997) The Bacterial nucleoid visualized by fluorescence microscopy of cells lysed within agarose: comparisons of *Escherichia coli* and spirochetes of the genus *Borrelia*. J. Bacteriol. **179:** 2228–2237
7 Trun N. J. and Marko J. F. (1998) Architecture of a bacterial chromosome. ASM News **64:** 276–283
8 Higgins N. P., Yang X., Fu Q. and Roth J. R. (1996) Surveying a supercoil domain by using the $\gamma\delta$ resolution system in *Salmonella typhimurium*. J. Bacteriol. **178:** 2825–2835
9 Staczek P. and Higgins N. P. (1998) Gyrase and topo IV determine chromosome domain size in vivo. Mol. Microbiol. **29:** 1435–1448
10 Dürrenberger M., Bjornsti M.-A., Uetz T., Hobot J. A. and Kellenberger E. (1988) Intracellular location of the histone-like protein HU in *Escherichia coli*. J. Bacteriol. **170:** 4757–4768
11 Murphy L. D. and Zimmerman S. B. (1997) Isolation and characterization of spermidine nucleoids from *Escherichia coli*. J. Struct. Biol. **119:** 321–335
12 Murphy L. D. and Zimmerman S. B. (1997) Stabilization of compact spermidine nucleoids from *Escherichia coli* under crowded conditions: implications for in vivo nucleoid structure. J. Struct. Biol. **119:** 336–346
13 Zimmerman S. B. and Murphy L. D. (1996) Macromolecular crowding and the mandatory condensation of DNA in bacteria. FEBS Lett. **390:** 245–248
14 Murphy L. D. and Zimmerman S. B. (1994) Macromolecular crowding effects on the interaction of DNA with *Escherichia coli* DNA-binding proteins: a model for bacterial nucleoid stabilization. Biochim. Biophys. Acta **1219:** 227–284
15 Shellman V. L. and Pettijohn D. E. (1991) Introduction of proteins into living bacterial cells: distribution of labeled HU protein in *Escherichia coli*. J. Bacteriol. **173:** 3047–3059
16 Drlica K. and Rouvière-Yaniv J. (1987) Histonelike proteins of bacteria. Microbiol. Rev. **51:** 301–319
17 Bonnefoy E. and Rouvière-Yaniv J. (1991) HU and IHF, two homologous histone-like proteins of *Escherichia coli*, form different protein-DNA complexes with short DNA fragments. EMBO J. **10:** 687–696
18 Claret L. and Rouvière-Yaniv J. (1997) Variation in HU composition during growth of *Escherichia coli*: the heterodimer is required for long term survival. J. Mol. Biol. **273:** 93–104
19 Bensaid A., Almeida A., Drlica K. and Rouvière-Yaniv J. (1996) Cross-talk between topoisomerase I and HU in *Escherichia coli*. J. Mol. Biol. **256:** 292–300
20 Hodges-Garcia Y., Hagerman P. J. and Pettijohn D. E. (1989) DNA ring closure mediated by protein HU. J. Biol. Chem. **264:** 14621–14623
21 Pontiggia A., Negri A., Beltrame M. and Bianchi M. E. (1993) Protein HU binds specifically to kinked DNA. Mol. Microbiol. **7:** 343–350

22 Zlatanova J. and vanHolde K. (1998) Binding to four-way junction DNA: a common property of architectural proteins? FASEB J. **12:** 421–431
23 Rouvière-Yaniv J. and Kjeldgaard N. O. (1979) Native *Escherichia coli* HU protein is a heterotypic dimer. FEBS Lett. **106:** 297–300
24 Tanaka I., Appelt K., Dijk J., White S. W. and Wilson K. S. (1984) 3-Å resolution structure of a protein with histone-like properties in prokaryotes. Nature **310:** 376–381
25 White S. W., Appelt K., Wilson K. S. and Tanaka I. (1989) A protein structural motif that bends DNA. Proteins: **5:** 281–288
26 Vis H., Mariani M., Vorgias C. E., Wilson K. S., Kaptein R. and Boelens R. (1995) Solution structure of the HU protein from *Bacillus stearothermophilus*. J. Mol. Biol. **254:** 692–700
27 Ditto M. D., Roberts D. and Weisberg R. A. (1994) Growth phase variation of integration host factor level in *Escherichia coli*. J. Bacteriol. **176:** 3738–3748
28 Rice P. A., Yang S., Mizuuchi K. and Nash H. A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell **87:** 1295–1306
29 Broccard F. and Prentki P. (1993) Specific interaction of IHF with RIBs, a class of bacterial repetitive DNA elements located at the 3′ end of transcription units. EMBO J. **12:** 5019–5027
30 Atlung T. and Ingmer H. (1997) H-NS: a modulator of environmentally regulated gene expression. Mol. Microbiol. **24:** 7–17
31 Spurio R., Dürrenburger M., Falconi M., LaTeana A., Pon C. L. and Gualerzi C. O. (1992) Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. Mol. Gen. Genet. **231:** 201–211
32 Tupper A. E., Owen-Hughes T. A., Ussery D. W., Santos D. S., Ferguson D. J. P., Sidebotham J. M. et al. (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. EMBO J. **13:** 258–268
33 Owen-Hughes T. A., Pavitt G. D., Santos D. S., Sidebotham J. M., Hulton C. S. J., Hinton J. C. D. et al. (1992) The Chromatin-associate protein H-NS interacts with curved DNA to influence DNA topology and gene expression. Cell **71:** 255–265
34 Jordi B. J. A. M., Fielder A. E., Burns C. M., Hinton J. C. D., Dover N., Ussery D. W. et al. (1997) DNA Binding is not sufficient for H-NS-mediated repression of *proU* expression. J. Biol. Chem. **272:** 12083–12090
35 Sondén B. and Uhlin B. E. (1996) Coordinated and differential expression of histone-like proteins in *Escherichia coli*: regulation and function of the H-NS analog *StpA*. EMBO J. **15:** 4970–4980
36 Kostrewa D., Granzin J., Stock D., Choe H.-W., Labahn J. and Saenger W. (1992) Crystal structure of the factor for inversion stimulation FIS at 2.0 Å resolution. J. Mol. Biol. **226:** 209–226
37 Sandmann C., Cordes F. and Saenger W. (1996) Structure model of a complex between the factor for inversion stimulation (FIS) and DNA: modeling protein-DNA complexes with dyad symmetry and known protein structures. Proteins **25:** 486–500
38 Yasuzawa K., Hayashi N., Goshima M., Kohno K., Imamoto F. and Kano Y. (1992) Histone-like proteins are required for cell growth and constraint of supercoils in DNA. Gene **122:** 9–15
39 Claret L. and Rouvière-Yaniv J. (1996) Regulation of HUα and HUβ by CRP and FIS in *Escherichia coli*. J. Mol. Biol. **263:** 126–139
40 Dersch P., Schmidt K. and Bremer E. (1993) Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. Mol. Microbiol. **8:** 875–889
41 Aviv M., Giladi H., Schreiber G., Oppenheim A. B. and Glaser G. (1994) Expression of the genes coding for the *Escherichia coli* integration host factor are controlled by growth phase, *rpoS*, ppGpp and by autoregulation. Mol. Microbiol. **14:** 1021–1031
42 Gordon G. S., Sitnikov D., Webb C. D., Teleman A., Straight A., Losick R. et al. (1997) Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. Cell **90:** 1113–1121
43 Niki H., Jaffe A., Imamura T., Ogura T. and Hiraga S. (1991) The new gene *mukB* codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of *E. coli*. EMBO J. **10:** 183–193
44 Robinow C. and Kellenberger E. (1994) The bacterial nucleoid revisited. Microbiol. Rev. **58:** 211–232
45 Costerton J. W., Poffenroth L., Wilt J. C. and Kordova N. (1976) Ultrastructural studies of the nucleoids of the pleomorphic forms of *Chlamydia psittaci* 6BC: a comparison with bacteria. Can. J. Microbiol. **22:** 16–28
46 Oberto J. and Rouvière-Yaniv J. (1996) *Serratia marcescens* contains a heterodimeric HU protein like *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. **178:** 293–297
47 Köhler P. and Marahiel M. A. (1997) Association of the histone-like protein HBsu with the nucleoid of *Bacillus subtilis*. J. Bacteriol. **179:** 2060–2064
48 Micka B. and Marahiel M. A. (1992) The DNA-binding protein HBsu is essential for normal growth and development in *Bacillus subtilis*. Biochimie **74:** 641–650
49 Setlow P. (1995) Mechanisms for the prevention of damage to DNA in spores of *Bacillus species*. Ann. Rev. Microbiol. **49:** 29–54
50 Lin D. C. and Grossman A. D. (1998) Identification and characterization of a bacterial chromosome partitioning site. Cell **92:** 675–685
51 Lewis P. J. and Errington J. (1997) Direct evidence for active segregation of *oriC* regions of the *Bacillus subtilis* chromosome and co-localization with the SpoOJ partitioning protein. Mol. Microbiol. **25:** 945–954
52 Webb C. D., Graumann P. L., Kahana J. A., Telemen A. A., Silver P. A. and Losick R. (1998) Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. Mol. Microbiol. **28:** 883–892
53 Britton R. A., Lin D. C. and Grossman A. D. (1998) Characterization of a prokaryotic SMC protein involved in chromosome partitioning. Genes Dev. **12:** 1254–1259
54 Erickson H. P. (1997) FtsZ, a tubulin homologue in prokaryote cell division. Trends Cell Biol. **7:** 362–367
55 O'Reilly M. and Devine K. M. (1997) Expression of AbrB, a transition state regulator from *Bacillus subtilis*, is growth phase dependent in a manner resembling that of FIS, the nucleoid binding protein from *Escherichia coli*. J. Bacteriol. **179:** 522–529
56 Moulder J. W. (1985) Comparative biology of intracellular parasitism. Microbiol. Rev. **49:** 298–337
57 Hackstadt T., Baehr W. and Ying Y. (1991) *Chlamydia trachomatis* developmentally regulated protein is homologous to eukaryotic histone H1. Proc. Natl. Acad. Sci. USA **88:** 3937–3941
58 Hackstadt T., Brickman T. J., Barry C. E. III and Sager J. (1993) Diversity in the *Chlamydia trachomatis* histone Hc2. Gene **132:** 137–141
59 Widom J. and Klug A. (1985) Structure of the 300 Å chromatin filament: X-ray diffraction from oriented samples. Cell **43:** 207–213
60 Tao S., Kaul R. and Wenman W. M. (1991) Identification and nucleotide sequence of a developmentally regulated gene encoding a eukaryotic histone H1-like protein from *Chlamydia trachomatis*. J. Bacteriol. **173:** 2818–2822
61 Perara E., Ganem D. and Engel J. N. (1992) A developmentally regulated chlamydial gene with apparent homology to eukaryotic histone H1. Proc. Natl. Acad. Sci. USA **89:** 2125–2129

62 Barry C. E. III, Hayes S. F., and Hackstadt T. (1992) Nucleoid condensation in *Escherichia coli* that express a chlamydial histone homolog. Science **256:** 377–379

63 Brickmann T. J., Barry III C. E. and Hackstadt T. (1993) Molecular cloning and expression of the *hctB* encoding a strain-variant chlamydial histone-like protein with DNA-binding activity. J. Bacteriol. **175:** 4274–4281

64 Christiansen G., Pedersen L. B., Koehler J. E., Lundemose A. G. and Birkelund S. (1993) Interaction between the *Chlamydia trachomatis* histone H1-like protein (Hc1) and DNA. J. Bacteriol. **175:** 1785–1795

65 Barry C. E. III, Brickman T. J. and Hackstadt T. (1993) Hc1-mediated effects on DNA structure: a potential regulator of chlamydial development. Mol. Microbiol. **9:** 273–283

66 Pedersen L. B., Birkelund S. and Christiansen G. (1994) Interaction of the *Chlamydia trachomatis* histone H1-like protein (Hc1) with DNA and RNA causes repression of transcription and translation in vitro. Mol. Microbiol. **11:** 1085–1098

67 Stephens R. S., Kalman S., Lammel C., Fan J., Marathe R., Aravind L. et al. (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science **282:** 754–759

68 Sikorav J.-L. and Church G. M. (1991) Complementary recognition in condensed DNA: accelerated DNA renaturation. J. Mol. Biol. **222:** 1085–1108

69 Charbonnier F. and Forterre P. (1994) Comparison of plasmid DNA topology among mesophilic and thermophilic Eubacteria and Archaebacteria. J. Bacteriol. **176:** 1251–1259

70 Poplawski A. and Bernander R. (1997) Nucleoid structure and distribution in thermophilic Archaea. J. Bacteriol. **179:** 7625–7630

71 Bernander R. and Poplawski A. (1997) Cell cycle characteristics of thermophlic Archaea. J. Bacteriol. **179:** 4963–4969

72 Stein D. B. and Searcy D. G. (1978) Physiologically important stabilization of DNA by a prokaryotic histone-like protein. Science **202:** 219–221

73 DeLange R. J., Williams L. C. and Searcy D. G. (1981) A histone-like protein (HTa) from *Thermoplasma acidophilum*. J. Biol. Chem. **256:** 905–911

74 Searcy D. G. and DeLange R. J. (1980) *Thermoplasma acidophilum* histone-like protein. Biochim. Biophys. Acta **609:** 197–200

75 Luger K., Mäder A. W., Richmond R. K., Sargent D. F. and Richmond T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature **389:** 251–260

76 Chartier F., Laine B., Sautière P., Touzel J.-P. and Albagnac G. (1985) Characterization of the chromosomal protein HMb isolated from *Methanosarcina barkeri*. FEBS Lett. **183:** 119–123

77 Chartier F., Laine B., Belaiche D. and Sautière P. (1989) Primary structure of the chromosomal proteins MC1a, MC1b and MC1c from the archaebacterium *Methanothrix soehngenii*. J. Biol. Chem. **264:** 17006–17015

78 Derkacheva N. I. and Kagramanova V. K. (1994) Archaebacterial chromosomes. Mol. Biol. **28:** 178–184

79 Toulmé F., LeCam E., Teyssier C., Delain E., Sautière P., Maurizot J.-C. and Culard F. (1995) Conformational changes of DNA minicircles upon the binding of the archaebacterial histone-like protein MC1. J. Biol. Chem. **270:** 6286–6291

80 Teyssier C., Laine B., Gervais A., Maurizot J.-C. and Culard F. (1994) Archaebacterial histone-like protein MC1 can exhibit a sequence-specific binding to DNA. Biochem. J. **303:** 567–573

81 López-García P., Knapp S., Ladenstein R. and Forterre P. (1998) In vitro DNA binding of the archaeal proteon Sso7d induced negative supercoiling at temperatures typical for thermophilic growth. Nucleic. Acids Res. **26:** 2322–2328

82 Mai V. Q., Chen X., Hong R. and Huang L. (1998) Small abundant DNA binding proteins from the thermoacidophilic archaeon *Sulfolobus shibatae* constrain negative DNA supercoils. J. Bacteriol. **180:** 2560–2563

83 Guagliardi A., Napoli A., Rossi M. and Ciaramella M. (1997) Annealing of complementary DNA strands above the melting point of the duplex promoted by an archaeal protein. J. Mol. Biol. **267:** 841–848

84 Celestina F. and Suryanarayana T. (1995) Nucleic acid binding properties of a helix stabilising nucleoid protein from the thermoacidophilic archaeon *Sulfolobus acidocaldarius* that condenses DNA into compact structures. Indian J. Biochem. Biophys. **32:** 447–455

85 Bohrmann B., Kellenberger E., Arnold-Schulz-Gahmen B., Sreenivas K., Suryanarayana T., Stroup D. et al. (1994) Localization of histone-like proteins in thermophilic *Archaea* by immunogold electron microscopy. J. Struct. Biol. **112:** 70–78

86 Edmondson S. P., Qiu L. and Shriver J. W. (1995) Solution Structure of the DNA-Binding Protein Sac7d from the hyperthermophile *Sulfolobus acidocaldarius*. Biochem. **34:** 13289–13304

87 Baumann H., Knapp S., Lundbäck T., Ladenstein R. and Härd T. (1994) Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus solfataricus*. Struct. Biol. **1:** 808–819

88 Robinson H., Gao Y.-G., McCrary B. S., Edmondson S. P., Shriver J. W. and Wang A. H.-J. (1998) The hyperthermophile chromosomal protein Sac7d sharply kinks DNA. Nature **392:** 202–205

89 Agback P., Baumann H., Knapp S., Ladenstein R. and Härd T. (1998) Architecture of nonspecific protein-DNA interactions in the Sso7d-DNA complex. Nature Struct. Biol. **5:** 579–584

90 Sandman K., Krzycki J. A., Dobrinski B., Lurz R. and Reeve J. N. (1990) HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus fervidus*, is most closely related to histones. Proc. Natl. Acad. Sci. USA **87:** 5788–5791

91 Pereira S. L., Grayling R. A., Lurz R. and Reeve J. N. (1997) Archaeal nucleosomes. Proc. Natl. Acad. Sci. USA **94:** 12633–12637

92 Sandman K., Grayling R. A., Dobrinski B., Lurz R. and Reeve J. N. (1994) Growth-phase-dependent synthesis of histones in the archaeon *Methanothermus fervidus*. Proc. Natl. Acad. Sci. USA **91:** 12624–12628

93 Olins A. L. and Olins D. E. (1974) Spheroid chromatin units (ν-bodies). Science **183:** 330–332

94 Howard M. T., Sandman K., Reeve J. N. and Griffith J. D. (1992) HMf, a histone-related protein from the hyperthermophilic archaeon *Methanothermus fervidus*, binds preferentially to DNA containing phased tracts of adenines. J. Bacteriol. **174:** 7864–7867

95 Sandman K., Zhu W., Summers M. F. and Reeve J. N. (1998) Evolution of the histone fold. In: Thermophiles: the key to molecular evolution and the origin of life? pp. 243–253, Weigel J. and Adams M. W. W. (eds), Taylor and Francis, London

96 Grayling R. A., Bailey K. A. and Reeve J. N. (1997) DNA binding and nuclease protection by the HMf histones from the hyperthermophilic archaeon *Methanothermus fervidus*. Extremophiles **1:** 79–88

97 Musgrave D. R., Sandman K. M. and Reeve J. N. (1991) DNA binding by the archaeal histone HMf results in positive supercoiling. Proc. Natl. Acad. Sci. USA **88:** 10397–10401

98 Stroup D. and Reeve J. N. (1992) Histone HMf from the hyperthermophilic archaeon *Methanothermus fervidus* binds to DNA in vitro using physiological conditions. FEMS Microbiol. Lett. **91:** 271–276

99 Arents G., Burlingame R. W., Wang B.-C., Love W. E. and Moudrianakis E. N. (1991) The nucleosomal core histone octamer at 3.1Å resolution: a tripartite protein assembly and a left-handed superhelix. Proc. Natl. Acad. Sci. USA **88:** 10148–10152

100 Tabassum R., Sandman K. M. and Reeve J. N. (1992) HMt, a histone-related protein from *Methanobacterium thermoautotrophicum ΔH*. J. Bacteriol. **174:** 7890–7895

101 Darcy T. J., Sandman K. and Reeve J. N. (1995) *Methanobacterium formicicum*, a mesophilic methanogen, contains three HFo histones. J. Bacteriol. **177:** 858–860
102 Ronimus R. S. and Musgrave D. R. (1995) A comparison of the DNA binding properties of histone-like proteins derived from representatives of the two kingdoms of the Archaea. FEMS Microbiol. Lett. **134:** 79–84
103 Ronimus R. S. and Musgrave D. R. (1996) Purification and characterization of a histone-like protein from the Archaeal isolate AN1, a member of the *Thermococcales*. Mol. Microbiol. **20:** 77–86
104 Slesarev A. I., Belova G. I., Kozyavkin S. A. and Lake J. A. (1998) Evidence for an early prokaryotic origin of histones H2A and H4 prior to the emergence of eukaryotes. Nucleic Acids Res. **26:** 427–430
105 Sandman K., Perler F. B. and Reeve J. N. (1994) Histone-encoding genes from *Pyrococcus*: evidence for members of the HMf family of archaeal histones in a non-methanogenic archaeon. Gene **150:** 207–208
106 Bult C. J., White O., Olsen G. J., Zhou L., Fleischmann R. D., Sutton G. G. et al. (1996) Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii*. Science **273:** 1058–1073
107 Smith D. R., Doucette-Stamm L. A., Deloughery C., Lee H., Dubois J., Aldredge T. et al. (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: functional analysis and comparative genomics. J. Bacteriol. **179:** 7135–7155
108 Klenk H.-P., Clayton R. A., Tomb J.-F., White O., Nelson K. E., Ketchum K. A. et al. (1998) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature **390:** 364–370
109 Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y., Yamamoto S. et al. (1998) Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Res. **5:** 55–76
110 Karantza V., Baxevanis A. D., Freire E. and Moudrianakis E. N. (1995) Thermodynamic studies of the core histones: ionic strength and pH dependence on H2A-H2B dimer stability. Biochemistry **34:** 5988–5996
111 Karantza V., Freire E. and Moudrianakis E. N. (1996) Thermodynamic studies of the core histones: pH and ionic strength effects on the stability of the (H3-H4)/$(H3\text{-}H4)_2$ system. Biochemistry **35:** 2037–2046
112 Li W., Grayling R. A., Sandman K., Edmondson S., Shriver J. W. and Reeve J. N. (1998) Thermodynamic stability of archaeal histones. Biochemistry **37:** 10563–10572
113 Starich M. R., Sandman, K., Reeve J. N. and Summers M. F. (1996) NMR structure of HMfB from the hyperthermophile, *Methanothermus fervidus*, confirms that this archaeal protein is a histone. J. Mol. Biol. **255:** 187–203
114 Zhu W., Sandman K., Lee G. E., Reeve J. N. and Summers M. F. (1998) NMR structure and comparison of the archaeal histone HFoB from the mesophile *Methanobacterium formicicum* with HMfB from the hyperthermophile *Methanothermus fervidus*. Biochemistry **37:** 10573–10580
115 Luger K. and Richmond T. J. (1998) DNA binding within the nucleosome core. Curr. Opin. Struct. Biol. **8:** 33–40
116 Ausio J., Dong F. and vanHolde K. E. (1989) Use of selectively trypsinized nucleosome core particles to analyze the role of the histone tails in the stabilization of the nucleosome. J. Mol. Biol. **206:** 451–463
117 Luger K. and Richmond T. J. (1998) The histone tails of the nucleosome. Curr. Opin. Genet. Dev. **8:** 140–146
118 Hamiche A., Carot V., Alilat M., DeLucia F., O'Donohue M.-F., Révet B. et al. (1996) Interaction of the histone $(H3\text{-}H4)_2$ tetramer of the nucleosome with positively supercoiled DNA minicircles: potential flipping of the protein from a left- to a right-handed superhelical form. Proc. Natl. Acad. Sci. USA **93:** 7588–7593
119 Hamiche A. and Richard-Foy H. (1998) The switch in the helical handedness of the histone $(H3\text{-}H4)_2$ tetramer within a nucleoprotein particle requires a reorientation of the H3-H3 interface. J. Biol. Chem. **273:** 9261–9269
120 Grayling R. A., Sandman K. and Reeve J. N. (1996) DNA stability and DNA binding proteins. Adv. Protein Chem. **48:** 437–467
121 Reeve J. N., Sandman K. and Daniels C. J. (1997) Archaeal histones, nucleosomes, and transcription initiation. Cell **89:** 999–1002
122 Pereira S. L. and Reeve J. N. (1998) Histones and nucleosomes in *Archaea* and *Eukarya*: a comparative analysis. Extremophiles **2:** 141–148
123 Dong F. and van Holde K. E. (1991) Nucleosome positioning is determined by the $(H3\text{-}H4)_2$ tetramer. Proc. Natl. Acad. Sci. USA **88:** 10596–10600
124 Spangenberg C., Eisfeld K., Stünkel W., Luger K., Flaus A., Richmond T. J. et al. (1998) The mouse mammary tumor virus promoter positioned on a tetramer of histones H3 and H4 binds nuclear factor 1 and OTF1. J. Mol. Biol. **278:** 725–739
125 Godde J. S. and Wolffe A. P. (1996) Nucleosome assembly on CTG triplet repeats. J. Biol. Chem. **271:** 15222–15229
126 Arents G. and Moudrianakis E. N. (1995) The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. Proc. Natl. Acad. Sci. USA **92:** 11170–11174
127 vanHolde K. E. (1989) Chromatin, Springer, New York
128 Agha-Amiri K. and Klein A. (1993) Nucleotide sequence of a gene encoding a histone-like protein in the archaeon *Methanococcus voltae*. Nucleic Acids Res. **21:** 1491
129 Ouzounis C. A. and Kyrpides N. C. (1996) Parallel origins of the nucleosome core and eukaryotic transcription from Archaea. J. Mol. Evol. **42:** 234–239
130 Kokubo T., Gong D.-W., Wootton J. C., Horikoshi M., Roeder R. G. and Nakatani Y. (1994) Molecular cloning of *Drosophila* TFIID subunits. Nature **367:** 484–487
131 Birck C., Poch O., Romier C., Ruff M., Mengus G., Lavigne A.-C. et al. (1998) Human $TAF_{II}28$ and $TAF_{II}18$ interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. Cell **94:** 239–249
132 Ogryzko V. V., Kotani T., Zhang X., Schlitz R. L., Howard T., Yang X.-J. et al. (1998) Histone-like TAFs within the PCAF histone acetylase complex. Cell **94:** 35–44
133 Mermelstein F., Yeung K., Cao J., Inostroza J. A., Erdjument-Bromage H., Eagelson K. et al. (1996) Requirement of a corepressor for Dr1-mediated repression of transcription. Genes Dev. **10:** 1033–1048
134 Baxevanis A. D., Arents G., Moudrianakis E. N. and Landsman D. (1995) A variety of DNA-binding and multimeric proteins contain the histone fold motif. Nucleic Acids Res. **14:** 2685–2691
135 Bonfini L., Karlovich C. A., DasGupta C. and Banerjee U. (1992) The *Son of sevenless* gene product: a putative activator of Ras. Science **255:** 603–606
136 Karlovich C. A., Bonfini L., McCollam L., Rogge R. D., Daga A., Czech M. P. et al. (1995) In vivo functional analysis of the Ras exchange factor *Son of sevenless*. Science **268:** 576–579
137 Mabilat C., Lourencao-Vital J., Goussard S. and Courvalin P. (1992) A new example of physical linkage between Tn1 and Tn21: the antibiotic multiple resistance region of plasmid pCFFO4 encoding extended-spectrum β-lactamase TEM-3. Mol. Gen. Genet. **235:** 113–121
138 Yang W. and Steitz T. A. (1995) Crystal structure of the site-specific recombinase $\gamma\delta$ resolvase complexed with a 34 bp cleavage site. Cell **82:** 193–208

139 Pisani F. M., De Martino C. and Rossi M. (1992) A DNA polymerase from the archaeon *Sulfolobus solfataricus* shown sequence similarity to family B DNA polymerase. Nucleic Acids Res. **20:** 2711–2716
140 Pisani F. M. and Rossi M. (1994) Evidence that an archaeal α-like DNA polymerase has a modular organization of its associated catalytic activities. J. Biol. Chem. **269:** 7887–7892
141 Sandman K. and Reeve J. N. (1998) Origin of the eukaryotic nucleus. Science **280:** 501–503
142 Martin W. and Müller M. (1998) The hydrogen hypothesis of the first eukaryote. Nature **392:** 37–41
143 Maidak B. L., Olsen G. J., Larsen N., Overbeek T., McCaughey M. J. and Woese C. R. (1997) The RDP (Ribosomal Database Project). Nucleic Acids Res. **25:** 109–111
144 Wells D. and Brown D. (1991) Histone and histone gene compilation and alignment update. Nucleic Acids Res. **19:** 2173
145 Sandman K. and Reeve J. N. (1998) Archaeal nucleosome positioning by CTG-repeats. J. Bacteriol., in press