Extremophiles (2000) 4:43–51

ORIGINAL PAPER

Wen-Tyng Li · Kathleen Sandman · Suzette L. Pereira
John N. Reeve

MJ1647, an open reading frame in the genome of the hyperthermophile *Methanococcus jannaschii*, encodes a very thermostable archaeal histone with a C-terminal extension

Received: September 2, 1999 / Accepted: October 18, 1999

Abstract All archaeal histones studied to date have similar lengths, 66 to 69 amino acid residues that form three α-helices separated by two β-strand loop regions which together constitute a histone fold. In contrast, the eukaryal nucleosome core histones are larger, 102 to 135 residues in length, with N-terminal and C-terminal extensions flanking the histone fold that participate in gene regulation and higher-order chromatin assembly. In the *Methanococcus jannaschii* genome, MJ1647 was annotated as an open reading frame predicted to encode an archaeal histone with an approximately 27-amino-acid C-terminal extension, and we here document the DNA binding and assembly properties and thermodynamic stability parameters of the recombinant product of MJ1647 synthesized in *Escherichia coli* with (rMJ1647) and without (rMJ1647Δ) the C-terminal extension. The presence of the C-terminal extension did not prevent homodimer formation or inhibit DNA binding, but the complexes formed by rMJ1647, presumably archaeal nucleosomes containing a $(rMJ1647)_4$ tetramer, were apparently less stable than those formed by $(rMJ1647\Delta)_4$. The presence of the C-terminal extension increased the thermostability of rMJ1647 when compared with rMJ1647Δ in 0.2 M KCl at pH 4 but not in the absence of KCl at pH 1. Based on thermal unfolding transitions, rMJ1647 and rHAfB generated by expression of AF0337 cloned from the genome of the related hyperthermophile *Archaeoglobus fulgidus* in *E. coli* were found to have higher thermodynamic stabilities than all previously studied archaeal histones.

Key words Histone fold · Recombinant protein · Thermostability · DNA packaging · *Archaeoglobus fulgidus* · Circular dichroism

Communicated by K. Horikoshi

W.-T. Li · K. Sandman · S.L. Pereira · J.N. Reeve (✉)
Department of Microbiology, The Ohio State University, Columbus, OH 43210, USA
Tel. +1-614-292-2301; Fax +1-614-292-8120
e-mail: reeve.2@osu.edu

Introduction

The eukaryal nucleosome core histones H2A, H2B, H3, and H4 have only limited primary sequence homologies (Wells and McBride 1989; Baxevanis and Landsman 1996), but they all have a structurally similar globular domain known as the histone fold formed by three α-helices (α1, α2, and α3) separated by two short loops (L1 and L2) (Arents and Moudrianakis 1995; Luger et al. 1997). The histone fold is stabilized by dimer formation and subunits in several eukaryal general and regulatory transcription factors are paired through histone fold domain interactions (Hoffmann et al. 1996; Nakatani et al. 1996; Burley et al. 1997; Ogryzko et al. 1998; Birck et al. 1998; Liberati et al. 1998; Michel et al. 1998; Moqtaderi et al. 1998). The eukaryal histones H2A, H2B, H3, and H4 have N- and C-terminal sequences that extend beyond the histone fold that provide targets for posttranslational regulatory modifications and participate in the higher-order assembly of nucleosomes into chromatin (Elgin 1995; Wan et al. 1995; Luger et al. 1997) but all the archaeal histones studied to date lack such N- and C-terminal extensions (Fig. 1) and are therefore essentially only histone folds. The archaeal histones must still dimerize to maintain structure; however, they are able to form both homodimers and heterodimers whereas the eukaryal histones form only (H2A + H2B) and (H3 + H4) heterodimers (Sandman et al. 1998). Agha-Amiri and Klein (1993) reported the sequence of a gene cloned from *Methanococcus voltae* that was predicted to encode an archaeal histone (designated HMvA) with an approximately 30-amino-acid C-terminal extension, and MJ1647 in the *Methanococcus jannaschii* genome is an open reading frame that encodes an amino acid sequence 55% identical to that of HMvA and is similarly annotated as encoding an archaeal histone with a C-terminal extension (Bult et al.

A.

```
                         ←L1→                                ←L2→            rMJ1647Δ
             [...α1....]      [............α2.............]    [...α3....]  R70✖           [....α4...]
                   10        20        30        40        50        60       ▼          80        90
rMJ1647   MLPKATVKRIMKQHTDFNISAEAVDELCNMLEEIIKITTEVAEQNARKEGRKTIKARDIKQCDDERLKRKIMELSERTDKMPILIKEMLNVITSEL
          *:**:*******  *:: :* * *  *  :* *:*  ** :** ** *: *********  ** **** **:::****:*: :*: *:* ** ***
HMvA      MIPKGTVKRIMKDNTEMYVSTESVVALVDILQEMIVTTTKIAEENAAKDKRKTIKARDIEECDAERLKEKILQVSERTEKVNMLANEILHVIASELERY
           :* : : **:*      :* :    *  :* *:         *   *    ****** **        *
HMfB     MELPIAPIGRIIKDAGAERVSDDARITLAKILEEMGRDIASEAIKLARHAGRKTIKAEDIELAVRRFKK
HMfA    MG------------N------------A---V-----EE-----V---K---------------RKMFK
HFoB     -------------N-----------EA---A---K-ET--T--V---K------V--S-V----K-L
HPyA1   MG------VD-L-RK-------EE-AKI--EY---YAIEVSKK-VEF--------V-----K--IKS
HAfB     ----L--VE-LLRK---S---E--KVE---AI--YAMQ-GKK-AE--K------V-VD--K--L-EL
              #  ##       #   #   ## ##  ##   #  #   #       #    #  #
```

B.

chicken histone H2A	NH_2---22aa--*histone fold*---37aa--COOH
chicken histone H2B	NH_2---33aa--*histone fold*---22aa--COOH
chicken histone H3	NH_2---63aa--*histone fold*----2aa--COOH
chicken histone H4	NH_2---26aa--*histone fold*----7aa--COOH
general transcription factor dTAF42	NH_2---17aa--*histone fold*--192aa--COOH
general transcription factor dTAF62	NH_2----3aa--*histone fold*--519aa--COOH
global negative regulator NCB1	NH_2---56aa--*histone fold*---22aa--COOH
global negative regulator NCB2	NH_2----7aa--*histone fold*---68aa--COOH
yeast CCAAT-box regulator HAP3	NH_2---39aa--*histone fold*---34aa--COOH
yeast CCAAT-box regulator HAP5	NH_2--157aa--*histone fold*---15aa--COOH

Fig. 1A,B. Alignment of archaeal histone amino acid sequences and conservation of the histone fold. **A** The amino acid sequences of rMJ1647 and HMvA are aligned with conserved (*) and similar (:) residues identified between the sequences. Residues conserved (*) or similar (:) in rMJ1647, HMvA, and in HMfB are identified above the HMfB sequence. Positions at which the archaeal histone sequences shown below the HMfB sequence are identical to the HMfB sequence are indicated by hyphens. The regions that form α1, α2, α3 of the histone fold are indicated; the location of the nonsense mutation (×) that terminates the MJ1647 open reading frame creating MJ1647Δ and the region predicted by NNPREDICT (Kneller 1991) to form an α-helix (α4) within the C-terminal extensions are indicated above the sequences. The locations of hydrophobic residues that contribute to the histone core and help maintain the histone fold are indicated (#) below the sequences. MJ1647 is from *Methanococcus jannischii* (Bult et al. 1996), HMvA1 is from *Methanococcus voltae* (Agha-Amiri and Klein 1993), HMfB and HMfA are from *Methanothermus fervidus* (Sandman et al. 1994a), HFoB is from *Methanobacterium formicicum* (Darcy et al. 1995), HPyA1 is from *Pyrococcus* strain GB-3a (Sandman et al. 1994b), and HAfB is from *Archaeoglobus fulgidus* (Klenk et al. 1998). **B** The lengths of the N- and C-terminal extensions that flank the histone fold in the eukaryal nucleosome core histones (Wells and McBride 1989; Baxevanis and Landsman 1996) and in representative pairs of nonhistone transcription factors that assemble into complexes through histone fold interactions (Kokubo et al. 1994; McNabb et al. 1995; Hoffmann et al. 1996; Gadbois et al. 1997)

1996). The *M. jannaschii* genome also appears to encode five archaeal histones that lack N- and C-terminal extensions, raising the question of whether the MJ1647 gene product behaves as a histone. To address this issue, and to determine properties conferred by the C-terminal extension, recombinant (r) MJ1647 gene product has been synthesized in *Escherichia coli* with and without the C-terminal extension (rMJ1647Δ), and both proteins have been shown to dimerize and form complexes with DNA in vitro consistent with their assembly into archaeal nucleosomes (Pereira et al. 1997; Sandman et al. 1998; Bailey et al. 1999). In addition, as it is known that archaeal histones with similar primary, secondary, and tertiary structures can have inherently very different thermodynamic stabilities (Li et al. 1998), we have determined the thermostabilities of rMJ1647 and rMJ1647Δ, and of rHAfB encoded by AF0337 in the genome of the related archaeal hyperthermophile *Archaeoglobus fulgidus* (Klenk et al. 1998). All three proteins have native structures that are more resistant to heat-induced unfolding than the structures of all previously studied archaeal histones (Li et al. 1998), and rMJ1647Δ unfolded at lower temperatures than rMJ1647, indicating that the C-terminal extension contributes to the overall stability of the rMJ1647 fold.

Materials and methods

Subcloning, expression of MJ1647 in *E. coli*, and purification of rMJ1647

E. coli clone AMJEZ84 that contained the MJ1647 region of the *M. jannaschii* genome cloned in pUC18 was purchased from the American Type Culture Collection (Rockville, MD, USA). *Eco*RI digestion of the plasmid DNA purified from this clone generated a 1714-bp fragment of *M. jannaschii* genomic DNA from which a 322-bp region that contained MJ1647 was amplified by polymerase chain reaction (PCR) technology with flanking *Bam*HI and *Nsi*I sites added by using oligonucleotide primers with the sequences 5′-CGGGATCCGGAGGTATATATTT-ATGTTACC-3′ and 5′-CCAATGCATCTATAATTCA-GATGTTATTAC-3′ (Ransom Hill Bioscience, Ramona, CA, USA). The amplified DNA and the expression vector pRAT4 (Peränen et al. 1996) were digested with *Bam*HI and *Nsi*I, mixed, ligated (Sambrook et al. 1989), and the ligation products used to transform XL-1 Blue supercompetent *E. coli* cells (Stratagene, La Jolla, CA, USA). Plasmid DNA (designated pMJ1647) was isolated

from an ampicillin-resistant transformant, sequenced to confirm the construction, and transformed into *E. coli* B834 (DE3) (Novagen, Madison, WI, USA) to obtain rMJ1647 synthesis. Cultures were grown aerobically at 37°C to OD_{600} ~0.4 in Luria-Bertani medium (Sambrook et al. 1989) that contained 100 μg ampicillin/ml, isopropyl β-D-thiogalactopyranoside (IPTG) was then added (final concentration of 400 μM), and incubation continued for 3 h. rMJ1647 was purified using the procedures used to purify rHMfB (Sandman et al. 1995; Li et al. 1998), except the buffer used was 50 mM 2-(*N*-morpholino)-ethanesulfonic acid (MES, pH 6). The composition and concentration of all archaeal histone preparations were determined, after acid hydrolysis, by amino acid analysis (Li et al. 1998).

Site-specific mutagenesis of MJ1647; synthesis and purification of rMJ1647Δ

The arginine codon (AGA) at position 69 of MJ1647 was changed to a termination codon (TGA) in an aliquot of pMJ1647 by using a QuikChange site-directed mutagenesis kit (Strategene, La Jolla, CA, USA) with the primers 5′-ATGATGAAAGATTAAAGTGAAAGATTATGGAAC-3′ and 5′- GTTCCATAATCTTTCACTTTAATCTTTCATCATC-3′ (Ransom Hill Bioscience). The mutation was confirmed in pMJ1647Δ by DNA sequencing, and rMJ1647Δ was purified from *E. coli* B834 (DE3) containing pMJ1647Δ following IPTG induction and quantitated as described for rMJ1647.

Cloning and expression of *hafB* (AF0337) in *E. coli* and purification of rHAfB

The *hafB* gene (AF0337) was PCR-amplified from *A. fulgidus* genomic DNA by using primers with the sequences 5′-GAAGGATCCAAAGCTCCCTCAGAGCAAGCTTG and 5′-GCGAATTCCCAATACGATATTTTTGC (Ransom Hill Biosciences) that hybridized to sequences flanking *hafB* (Klenk et al. 1998) and added *Bam*HI and *Eco*RI sites, respectively, to the amplified product. The PCR-amplified DNA was digested with *Bam*HI and *Eco*RI, ligated with *Bam*HI- plus *Eco*RI-digested pRAT4 (Peränen et al. 1996), and the ligation products used to transform XL-1 Blue super-competent *E. coli* cells (Stratagene). The sequence of plasmid DNA isolated from an ampicillin-resistant transformant was determined to confirm the construction and this plasmid was transformed into *E. coli* BL21(DE3) (Novagen) for rHAf*B* synthesis. Expression of *hafB* was induced by addition of IPTG to exponentially growing cultures of the *E. coli* BL21(DE3) transformant, and rHAfB was purified by the procedures used to purify rHMfB (Sandman et al. 1995; Li et al. 1998).

Agarose gel electrophoretic mobility shift assay (EMSA)

Archaeal histone–DNA complexes formed with DNA molecules longer than 1.2 kbp migrate faster through 1% (w/v) agarose gels during electrophoresis at 1 V/cm than do the DNA molecules alone (Sandman et al. 1990). This property was used with linear pBR322 DNA to assay and compare the DNA-binding and compacting activities of rMJ1647, rMJ1647Δ, and rHAfB. Control experiments employed rHMfB as previously described (Sandman et al. 1995).

Formaldehyde protein–protein cross-linking

Aliquots (10 μl) of MJ1647 and MJ1647Δ dissolved in 25 mM glycine buffer (~1 mg/ml; pH 1) were incubated at 95°C for 10 min and then allowed to cool slowly to room temperature. Cross-linking buffer (20 μl of 150 mM KCl; 25 mM potassium phosphate) was added, the solution dialyzed for 1 h against cross-linking buffer before formaldehyde was added (final concentration of 400 mM), and incubation continued for 1 h at room temperature. The cross-linking reaction was terminated by addition of 3 μl 0.5 M $(NH_4)HCO_3$ and incubation for 15 min at room temperature. Following dialysis for 1 h against 100 mM NaCl, 50 mM MES (pH 6), the proteins present were separated by tricine sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Grayling et al. 1995) and detected by immunoblotting using polyclonal anti-HMf IgG as the primary antibody, antirabbit IgG peroxidase conjugate as secondary antibody (Sigma, St. Louis, MO, USA), and complex visualization by chemiluminescence (LumiGLO kit; Kirkgarrd and Perry Labs., Gaithersburg, MD, USA).

Circular dichroism spectropolarimetry and determination of thermal unfolding transitions

Circular dichroism (CD) spectra of recombinant archaeal histone solutions ranging in concentration from 0.3 to 6 μM of histone monomer were obtained at 25°C using an AVIV 62A-DS spectropolarimeter (AVIV, Lakewood, NJ, USA) with a 1-mm-pathlength cylindrical quartz cell and averaging times of 2 to 5 s. Temperature-induced changes in the CD measurement at 222 nm (θ_{222}) were determined at 1°C intervals from 0° to 99°C using a 10-mm-pathlength quartz cell with an averaging time of 5 s and the temperature maintained within ±0.2°C. Baseline measurements were determined using deionized water and subtracted from the experimental data. A two-state model in which a histone dimer unfolds directly into two random coil monomers with negligibly populated intermediate states was shown previously to fit the thermal unfolding data obtained for $(rHMfB)_2$, $(rHMfA)_2$, $(rHFoB)_2$, and $(rHPyA1)_2$ dimers (Li et al. 1998). Based on the excellent fit of the predictions of this model to the experimental data obtained here, this thermal unfolding pathway appears also to be valid for $(rMJ1647)_2$, $(rMJ1647\Delta)_2$, and $(rHAfB)_2$. The comparative data presented for $(rHMfA)_2$, $(rHMfB)_2$, $(rHFoB)_2$, and $(rHPyA1)_2$ were obtained using the protein preparations described previously (Li et al. 1998).

Results

rMJ1647 and rMJ1647Δ dimer formation, DNA binding, and complex stability

MJ1647 encodes an amino acid sequence that is only ~35% identical to that of HMfB, the most studied archaeal histone (Sandman et al. 1998), but does contain residues P4, R10, S21, A47, R52, K53, T54, and D59 (HMfB residue numbering system) that are conserved in all archaeal histones (see Fig. 1). Based on the crystal structure of the eukaryal nucleosome (Luger et al. 1997), R10, S21, K53, and T54 are predicted to participate directly in DNA binding, and R52 and D59 form a buried intramolecular bond essential for histone fold stabilization. From the alignment of residues and the structures established for $(rHMfB)_2$ and $(rHFoB)_2$ (Starich et al. 1996; Zhu et al. 1998), the sequences predicted to form α1, α2, α3, L1, and L2 that constitute the histone folds of rMJ1647 and rMJ1647Δ, HMvA, and HAfB are identified in Fig. 1. The sites located primarily along the buried faces of the three α-helices at which hydrophobic residues interact to maintain the histone dimer configuration are indicated, and the conservation of hydrophobic residues at these sites in rMJ1647, HMvA, and HAfB is consistent with these polypeptides also forming dimers. The C-terminal extensions of rMJ1647 and HMvA are predicted (Kneller 1991) to contain a fourth α-helix (designated α4 in Fig. 1) that would not constitute part of the histone fold. A longer α-helix (designated αC), with marginal sequence similarly to α4, is present in the C-terminal extension of eukaryal H2B histones.

Incubating preparations of rMJ1647 and rMJ1647Δ with formaldehyde under conditions that covalently cross-linked the histone monomers within $(rHMfB)_2$ dimers (Grayling et al. 1995) generated molecules with electrophoretic mobilities consistent with cross-linked dimers (Fig. 2). A small fraction of the molecules in rMJ1647Δ preparations migrated as dimers without exposure to formaldehyde, indicating the presence of dimers so intrinsically stable that they resisted denaturation by 1% SDS and demonstrating clearly that the C-terminal extension of MJ1647 was not essential for dimer formation. The C-terminal extension was also not essential for DNA binding and compaction by rMJ1647, but rather appeared to reduce the stability of the complexes formed with linear pBR322 DNA when compared with the complexes formed by rMJ1647Δ and rHMfB (Fig. 3). Consistent with archaeal nucleosome formation, the complexes formed by rMJ1647 and rMJ1647Δ binding to linear pBR322 DNA migrated faster during electrophoresis than pBR322 molecules alone, but whereas the rMJ1647Δ- and control rHMfB-containing complexes migrated as a single, discrete band, the rMJ1647-containing complexes migrated to form a diffuse, broad smear consistent with instability and substantial disassembly during electrophoresis.

Thermal unfolding transitions and thermodynamic stability curves

Archaeal histones with similar primary sequences and secondary and tertiary structures have different inherent stabilities and, as the thermal unfolding of archaeal histones is fully reversible, they offer a tractable experimental system to determine and compare quantitative structure–stability relationships at the individual residue level (Li et al. 1998; Zhu et al. 1998). As previously documented, in 200 mM KCl at pH 4, rHFoB from the mesophile *Methanobacterium formicicum* exhibits both low and high temperature-induced unfolding, whereas rHMfA, rHMfB, and rHPyA1 from the hyperthermophiles *M. fervidus* and *Pyrococcus* strain GB-3a, respectively, exhibit only high temperature unfolding (Li et al. 1998). Full unfolding transitions could also be observed for MJ1647Δ and rHAfB under these conditions, but not for MJ1647, which remained ~60% folded at 99°C, the highest temperature at which CD measurements could be made (Fig. 4A). The presence of the C-terminal extension therefore increased the stability of rMJ1647 relative to rMJ1647Δ, resulting in the retention of

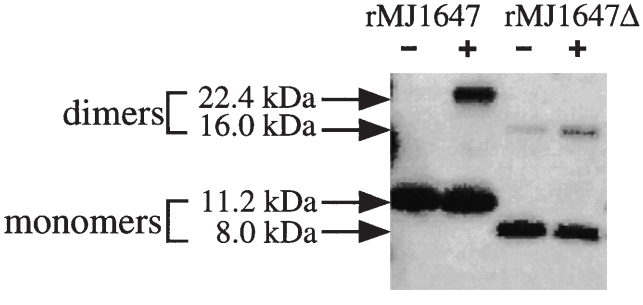

Fig. 2. SDS-PAGE separation of aliquots of rMJ1647 and rMJ1647Δ with (+) and without (−) exposure to formaldehyde cross-linking. The proteins present were detected by immunoblotting using polyclonal anti-HMf antibodies and the histone monomers and dimers were identified by their electrophoretic mobilities relative to size standards. rMJ1647 and rMJ1647Δ monomers have calculated masses of 11.2 and 8 kDa, respectively

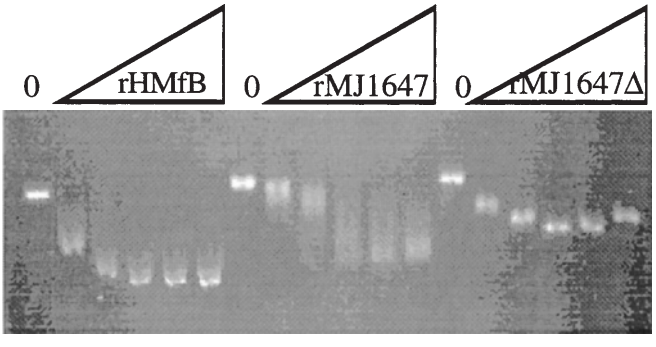

Fig. 3. Electrophoretic mobility shift assays of the complexes formed by rHMfB, rMJ1647, and rMJ1647Δ with linearized pBR322. Complexes formed at rHMfB to DNA mass ratios of 0.2, 0.4, 0.6, 0.8, and 1.2, rMJ1647 to DNA at mass ratios of 0.4, 0.8, 1.2, 1.6, and 2.4, and rMJ1647Δ to DNA mass ratios of 0.2, 0.4, 0.6, 1.0, and 1.4 were subjected to electrophoresis through a 0.8% agarose gel at 1 V/cm and detected by ethidium bromide staining. Control tracks that contained protein-free, linear pBR322 DNA molecules are indicated (0)

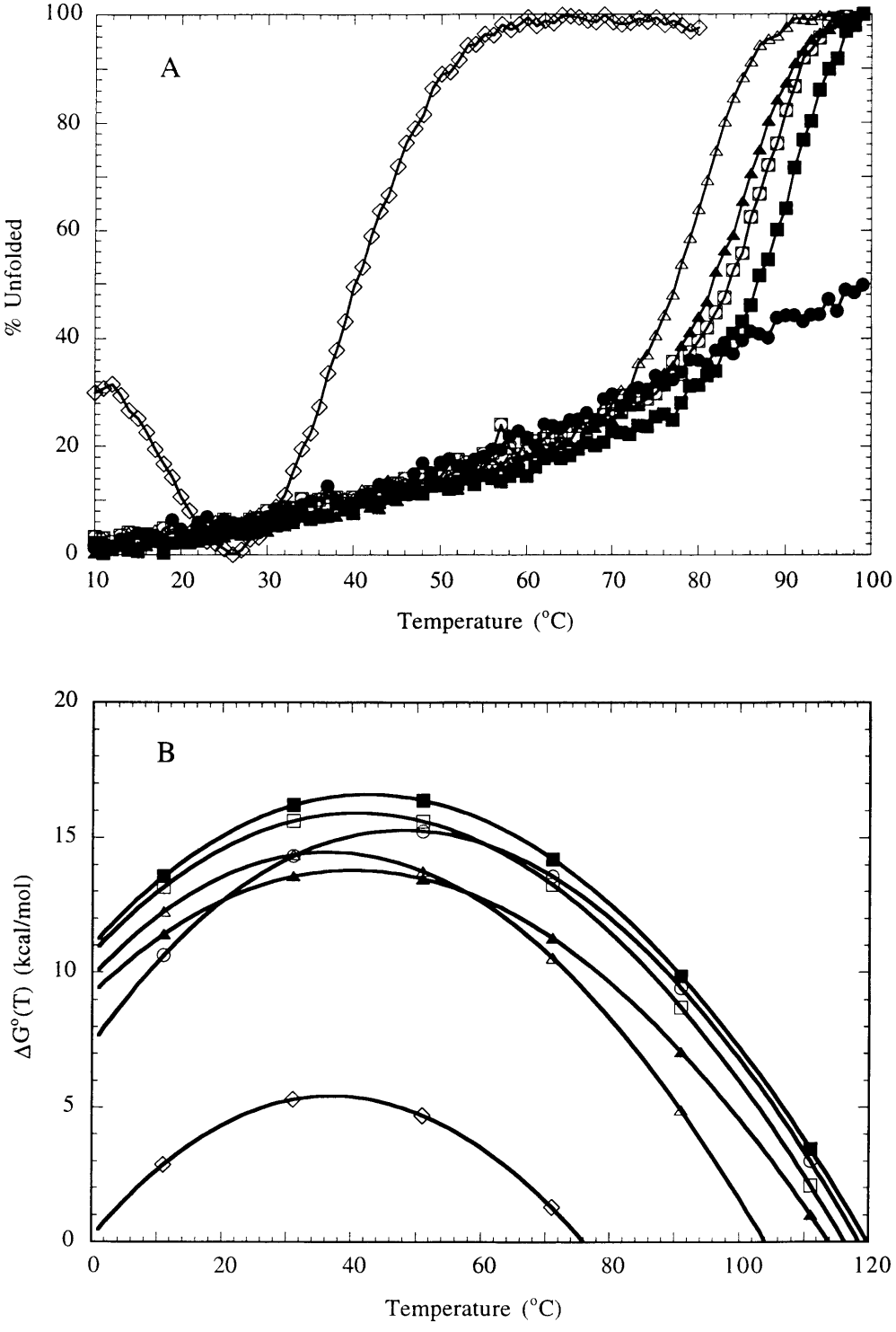

Fig. 4A,B. Thermal unfolding and stability curves of archaeal histones. **A** Unfolding of 0.45 μM rMJ1647 (*solid circles*), 0.38 μM rMJ1647Δ (*open circles*), 0.43 μM rHAfB (*solid squares*), 1.52 μM rHPyA1 (*open squares*), 5.22 μM rHMfB (*solid triangles*), 5.22 μM rHMfA (*open triangles*) and 4.50 μM rHFoB (*open diamonds*) in 0.2 M KCl at pH 4 in 25 mM glycine buffer determined by θ_{222} measurements. The θ_{222} value for 100% unfolded rMJ1647 was determined by measurements above 90°C at pH 1 (see Fig. 5B). **B** Stability curves calculated for rMJ1647Δ (*open circles*), rHAfB (*solid squares*), rHPyA1 (*open squares*), rHMfB (*solid triangles*), rHMfA (*open triangles*), and rHFoB (*open diamonds*) at pH 4 in 0.2 M KCl based on the CD thermal unfolding data

secondary structure above 100°C. The stability curves (Becktel and Schellman 1987) calculated for rHFoB, rHMfA, rHMfB, rHAfB, and rMJ1647Δ dimers, based on the thermal transition data shown in Fig. 4A, are provided in Fig. 4B with the thermodynamic parameters for the unfolding of these dimers listed in Table 1. All the recombinant archaeal histones that originate from hyperthermophiles had higher stabilities at all temperatures than rHFoB, although they have maximum stabilities at essentially mesophilic temperatures ranging from 36° to 48°C. Each stability curve has a slightly different shape, consistent with different molecular interactions contributing to overall fold stability (Fig. 4B). rMJ1647Δ had a marginally lower maximum stability than rHPyA1 although the unfolding transitions of these two proteins appeared almost identical (Fig. 4A), and rHAfB had a higher stability at all temperatures than all the previously studied archaeal histones (Li et al. 1998).

Table 1. Thermodynamic parameters of archaeal histone unfolding[a,b]

Histone	$T°$ (°C)	ΔH_{vh} (kcal/mol)	ΔC_p (cal/deg/mol)	T_{max} (°C)	$\Delta G°_{max}$ (kcal/mol)
rHFoB	75.6 ± 0.0	91.1 ± 0.3	2292 ± 0	38	5.0
rHMfA	104.0 ± 0.5	156.0 ± 2.3	2090 ± 58	36	14.5
rHMfB	113.8 ± 0.5	139.9 ± 1.5	1711 ± 36	40	13.8
rHPyA1	116.3 ± 0.7	158.4 ± 2.3	1886 ± 52	41	15.9
rHAfB	119.8 ± 0.7	162.4 ± 2.3	1882 ± 52	42	16.6
rMJ1647Δ	118.5 ± 0.7	164.0 ± 2.9	2106 ± 68	48	15.3
rMJ1647[c]	nd[c]	nd	nd	nd	nd

[a] Parameters calculated from the thermal unfolding of the archaeal histones in 0.2 M KCl, 25 mM glycine (pH 4) shown in Fig. 4A. $T°$ is the temperature at which the free energy of unfolding of a standard state (1 M) solution is zero; ΔH_{vh} is the van't Hoff enthalpy of dimer unfolding; ΔC_p is the change in heat capacity for dimer unfolding; T_{max} is the temperature at which the dimer has maximum stability; $\Delta G°max$ is the free energy of unfolding of a standard state solution at T_{max}

[b] The errors are standard deviations from the curvature of the χ^2 surface of each fit (Bevington 1969), and represent the quality of the fit

[c] Unfolding was incomplete at 99°C, the upper limit for CD measurements, and therefore these parameters could not be determined (nd) for MJ1647 under these solution conditions

CD spectra and rMJ1647 thermal unfolding at pH 1

To observe complete thermal unfolding transitions for rMJ1647, it was necessary to reduce the pH and salt concentration. Based on θ_{222} ellipticity measurements, both rMJ1647 and rMJ1647Δ retained their predominantly α-helical secondary structure at pH 1 in 25 mM glycine buffer at 25°C, and a difference in θ_{222} was observed between rMJ1647 and rMJ1647Δ consistent with the presence of an additional α-helix, α4, in the C-terminal extension of rMJ1647 (Fig. 1). In contrast, rHMfB was almost completely unfolded as previously reported (Li et al. 1998) under these conditions (Fig. 5A). The heat-induced unfolding transitions followed for rMJ1647 and rMJ1647Δ at pH 1 in the absence of salt were very similar (Fig. 5B). Under these conditions, the C-terminal extension did not substantially increase the unfolding transition temperature of rMJ1647 when compared with rMJ1647Δ, as was the case for unfolding at pH 4 in 0.2 M KCl (Fig. 4A). The thermodynamic parameters calculated from the complete unfoldings of rMJ1647 and rMJ1647Δ observed at low pH and low salt concentrations are listed in Table 2. The $T°$ values calculated for $(rMJ1647)_2$ in 0.1 M KCl, the temperature at which the free energy of unfolding of a standard state (1 M) protein solution would be zero, are above 120°C and are the highest such values so far determined for an archaeal histone dimer.

Table 2. Thermodynamic parameters of rMJ1647 and rMJ1647Δ unfolding[a]

Protein	KCl[b] [M]	pH	$T°$ (°C)	ΔH_{vh} (kcal/mol)	ΔC_p (cal/deg/mol)
rMJ1647	–	1	104.0 ± 3.8	105.3 ± 7.6	1331 ± 178
		3	118.7 ± 1.4	121.4 ± 2.7	1374 ± 60
	0.1	1	122.1 ± 3.2	105.6 ± 5.1	1181 ± 133
		2	121.0 ± 3.2	108.6 ± 5.1	1219 ± 133
rMJ1647Δ	–	1	104.5 ± 0.3	106.0 ± 0.7	1316 ± 17
	0.2	4	118.5 ± 0.7	164.0 ± 2.9	2106 ± 68

[a] $T°$ is the temperature at which the free energy of unfolding of a standard state (1 M) solution is zero; ΔH_{vh} is the van't Hoff enthalpy of dimer unfolding; ΔC_p is the change in heat capacity for dimer unfolding. The errors are standard deviations from the curvature of the χ^2 surface of each fit (Bevington 1969) and represent the quality of the fit

[b] Salt concentration in 25 mM glycine buffer

Discussion

The histone fold is only stable when complexed with a second histone fold, and histone fold dimerization involves predominantly hydrophobic interactions between residues positioned along the buried faces of the α-helices (Starich et al. 1996; Luger et al. 1997; Zhu et al. 1998). The original histone fold presumably formed a homodimer, and this is still the case for the archaeal histones although they also form heterodimers (Sandman et al. 1994a) and, in vitro, archaeal histone monomers from different species form heterodimers (unpublished results). In contrast, the eukaryal histones and histone domain-containing transcription factors do not exhibit such promiscuity and assemble with only one partner, usually a nonhomologous polypeptide, to form a heterodimeric complex. The flexibility in dimer formation exhibited by archaeal histones may generally be facilitated by their lack of N- and C-terminal extensions, but rMJ1647 with a C-terminal extension still formed homodimers (Fig. 2), and also heterodimers in vitro not only with rMJ1647Δ but also with rHMfB monomers (results not shown).

The presence of the C-terminal extension of MJ1647 apparently reduced the stability of the complexes formed with pBR322 DNA in vitro (Fig. 3). Archaeal nucleosomes contain a histone tetramer (Pereira et al. 1997; Bailey et al. 1999) in which, based on the eukaryal precedent (Luger et al. 1997), two histone dimers interact through a four-helix bundle formed by two α2s and two α3s. The C-terminal extension of rMJ1647 extends from α3, and in archaeal

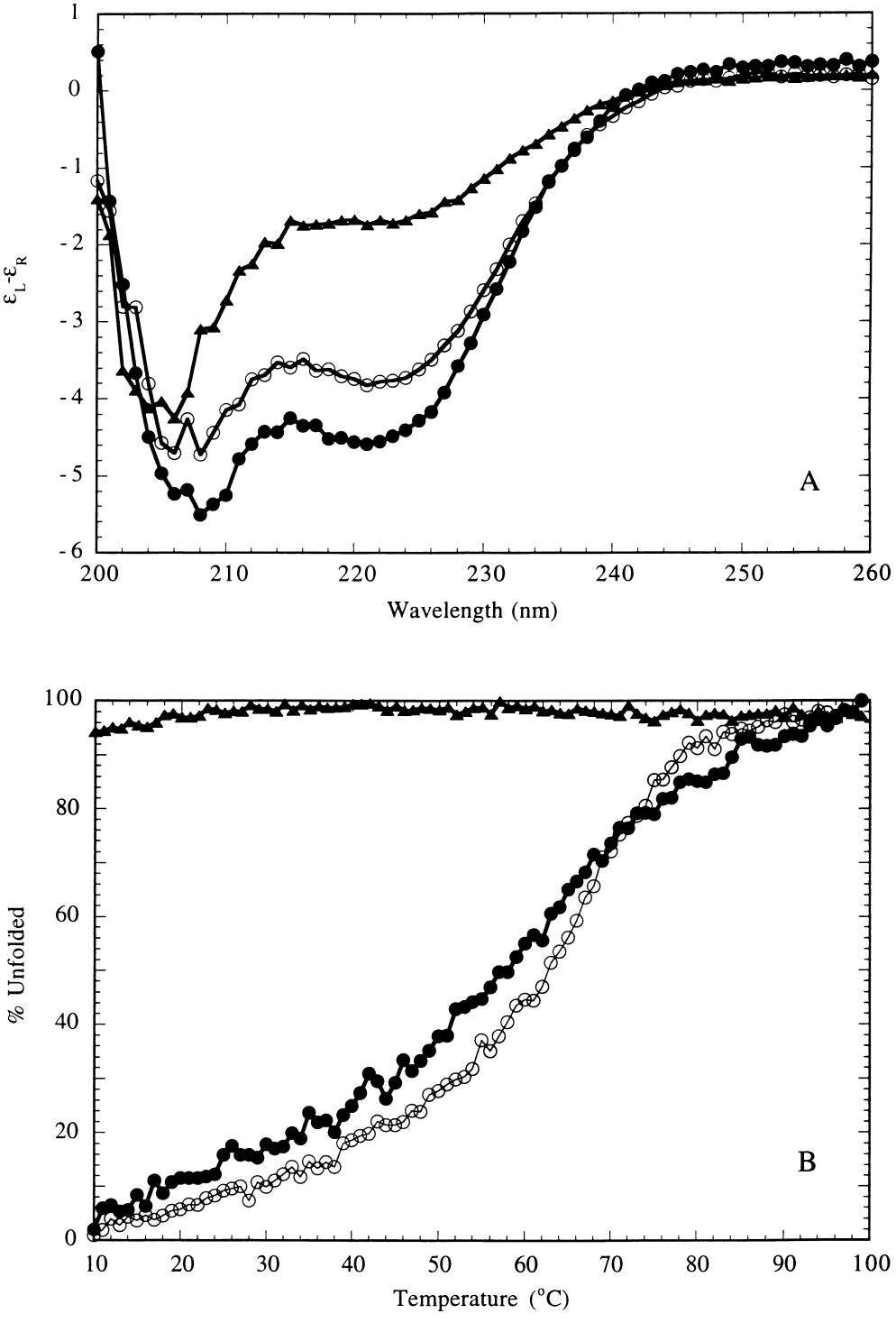

Fig. 5A,B. CD spectra and thermal unfolding transitions at pH 1. **A** CD spectra determined at 25°C for 1.7 μM rMJ1647 (*solid circles*), 2.4 μM rMJ1647Δ (*open circles*), and 2.6 μM rHMfB (*solid triangles*) solutions in 25 mM glycine at pH 1. Circular dichroism intensities (liters/cm × moles of residues) are given as the difference in extinction coefficient for left (ε_L) and right (ε_R) circularly polarized light. **B** Thermal unfolding of rMJ1647 (*solid circles*), rMJ1647Δ (*open circles*), and rHMfB (*solid triangles*) in 25 mM glycine at pH 1 determined by θ_{222} measurements. The θ_{222} value of 100% folded rHMfB was determined in 0.2 M KCl at pH 4

nucleosomes containing a rMJ1647 tetramer, four copies of this extension must occupy space not normally so occupied in archaeal nucleosomes formed by tetramers of the histone-fold-only archaeal histones. Accommodating this additional mass, especially the two C-terminal extensions at the center of the four-helix bundle, may well reduce the overall stability of the archaeal nucleosome core and result in complexes that are unstable during extended electrophoresis (Fig. 3). As the *M. jannaschii* genome also encodes five histone-fold-only archaeal histones, MJ1647 may preferentially form heterodimers with one or more of these smaller histones in vivo, which would reduce the demand for space within the archaeal nucleosome core and result in more stable structures.

Hydrophobic interactions between residues along the antiparallel aligned α2s predominate in dimer formation and maintenance (Luger et al. 1997), and differences in these residues should be reflected in differences in overall

fold stability. At the center of this interaction are residues 31, 35, and 36, and whereas large hydrophobic residues generally occupy these positions in the archaeal histones from hyperthermophiles, rHFoB from the mesophile has A31, K35, and G36 (Fig. 1). rMJ1647 and rHAfB have now also been shown to have very stable structures (Tables 1,2), and consistent with this they have only hydrophobic residues – M31, I35, I36, and A31, Y35, A36, respectively – at these positions. Comparisons of thermal transitions of rMJ1647 and rMJ1647Δ revealed that the C-terminal extension increased the overall stability of rMJ1647 at pH 4 in the presence of 0.2 M KCl (Fig. 4A), but rMJ1647 was so stable under these conditions that complete thermal transitions could not be observed and the quantitative contribution of the C-terminal extension to overall fold stability could not therefore be calculated. Complete unfolding transitions were observed for both rMJ1647 and rMJ1647Δ at pH 1 (Fig. 5B) but, under these partially destabilizing conditions, the presence of the C-terminal extension did not increase fold stability. At lease 37 of the 97 residues in rMJ1647 are potentially charged (D, E, K, and R), consistent with structure stabilization by ion pair formation and a surface-located salt bridge network (Zhu et al. 1998). Ten of these charged residues occur in the C-terminal extension with 7 occurring within first 12 residues of the extension preceding the region predicted to form α4 (Fig. 1). It seems likely therefore that in the presence of salt and at a higher pH this highly charged region also helps to stabilize the overall fold of rMJ1647 by forming additional ionic interactions.

In conclusion, we have confirmed that rMJ1647, and by inference HMvA, forms homodimers in solution and binds to DNA in vitro forming compact structures consistent with assembly into archaeal nucleosomes. Having C-terminal extensions, these molecules appear to represent an intermediate step in the evolution of the larger N- and C-terminal extension containing eukaryal histones and histone-fold-containing transcription factors, and it is now important to determine if the C-terminal extensions of MJ1647 and HMvA provide similar targets for analogous or homologous regulatory posttranslational modifications. As expected for proteins from hyperthermophiles, rMJ1647 and rHAfB have very stable structures, and rMJ1647Δ is less stable than rMJ1647 under physiologically relevant solution and temperature conditions (Martin et al. 1999), indicating that the C-terminal extension contributes to overall fold stability.

Acknowledgments This work was supported by grants from the National Institutes for Health (GM53185) and from the U.S. Department of Energy (DE-FG02-87ER13731). We are very grateful to J.W. Shriver for his continued help and advice, and for providing the data analysis software, and we thank M. Nguyen for technical assistance.

References

Agha-Amiri K, Klein A (1993) Nucleotide sequence of a gene encoding a histone-like protein in the archaeon *Methanococcus voltae.* Nucleic Acids Res 21:149

Arents G, Moudrianakis EN (1995) The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. Proc Natl Acad Sci USA 92:11170–11174

Bailey, KA, Chow CS, Reeve JN (1999) Histone stoichiometry and DNA circularization in archaeal nucleosomes. Nucleic Acids Res 27:532–536

Baxevanis AD, Landsman D (1996) Histone sequence database: a compilation of highly-conserved nucleoprotein sequences. Nucl Acids Res 24:245–247

Becktel WJ, Schellman JA (1987) Protein stability curves. Biopolymers 26:1862–1877

Bevington PR (1969) Data reduction and error analysis for the physical sciences. McGraw-Hill, New York

Birck C, Poch O, Romier C, Ruff M, Mengus G, Lavigne A-C, Davidson I, Moras D (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

Bult CJ, White O, Olsen G, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb J-F, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JT, Nguyen D, Utterback TR, Kelley JM, Peterson JD, Sadow PW, Hanna MC, Cotton MD, Roberts KM, Hurst MA, Kaine BP, Borodovsky M, Klenk H-P, Fraser CM, Smith HO, Woese CR, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii.* Science 273:1058–1073

Burley SK, Xie X, Clark KL, Shu F (1997) Histone-like transcription factors in eukaryotes. Curr Opin Struct Biol 7:94–102

Darcy TJ, Sandman K, Reeve JN (1995) *Methanobacterium formicicum,* a mesophilic methanogen, contains three HFo histones. J Bacteriol 177:858–860

Elgin SR (1995) Chromatin structure and gene expression. IRL Press, Oxford

Gadbois EL, Chao DM, Reese JC, Green MR, Young RA (1997) Functional antagonism between RNA polymerase II holoenzyme and global negative regulator NC2 in vivo. Proc Natl Acad Sci USA 94:3145–3150

Grayling RA, Becktel WJ, Reeve JN (1995) Structure and stability of histone HMf from the hyperthermophilic archaeon *Methanothermus fervidus*. Biochemistry 34: 8441–8448

Hoffmann A, Chiang C-M, Oelgeschläger T, Xie X, Burley S, Nakatani Y, Roeder RG (1996) A histone octamer-like structure within TFIID. Nature (Lond) 380:356–359

Kneller D (1991) NNPREDICT, protein secondary structure prediction. http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html

Klenk H-P, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, Richardson DL, Kerlavage AR, Graham DE, Krypides NC, Fleischmann RD, Quackenbush J, Lee NH, Sutton GG, Gill S, Kirkness EF, Dougherty BA, McKenney K, Adams MD, Loftus B, Peterson S, Reich CI, McNeil LK, Badger JH, Glodek A, Zhou L, Overbeek R, Gocayne JD, Weidman JF, McDonald L, Utterback T, Cotton MD, Spriggs T, Artiach P, Kaine BP, Sykes SM, Sadow PW, D'Andrea KP, Bowman C, Fujii C, Garland SA, Mason TM, Olsen GJ, Fraser CM, Smith HO, Woese CR, Venter JC (1998) The complete genome sequence of the hyper-thermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature (Lond) 390:364–370

Kokubo T, Gong D-W, Wootton JC, Horikoshi M, Roeder RG, Nakatani Y (1994) Molecular cloning of *Drosophila* TFIID subunits. Nature (Lond) 367:484–487

Li W-T, Grayling RA, Sandman K, Edmonson S, Shriver JW, Reeve JN (1998) Thermodynamic stability of archaeal histones. Biochemistry 37:10563–10572

Liberati C, di Silvio A, Ottolenghi S, Mantovani R (1998) NF-Y binding to twin CCAAT boxes: role of Q-rich domains and histone fold helices. J Mol Biol 285:1441–1455

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8Å resolution. Nature (Lond) 389:251–260

McNabb DS, Xing Y, Guarente L (1995) Cloning of yeast HAP5: a novel subunit of a heterotrimeric complex required for CCAAT binding. Genes Dev 9:47–58

Martin DM, Ciulla RA, Roberts MA (1999) Osmoadaptation in *Archaea.* Appl Environ Microbiol 65:1815–1825

Michel B, Komarnitsky P, Buratowski S (1998) Histone-like TAFs are essential for transcription in vivo. Mol Cell 2:663–673

Moqtaderi Z, Keaveney M, Struhl K (1998) The histone H3-like TAF is broadly required for transcription in yeast. Mol Cell 2:675–682

Nakatani Y, Bagby S, Ikura M (1996) The histone folds in transcription factor TFIID. J Biol Chem 271:6575–6578

Ogryzko VV, Kotani T, Zhang X, Schiltz RL, Howard T, Yang X-J, Howard BH, Qin J, Nakatani Y (1998) Histone-like TAFs within the PCAF histone acetylase complex. Cell 94:35–44

Peränen J, Rikkonen M, Hyvonen M, Kaariainen L (1996) T7 vector with a modified T7*lac* promoter for expression of proteins in *E. coli.* Anal Biochem 236:371–373

Pereira SL, Grayling RA, Lurz R, Reeve JN (1997) Archaeal nucleosomes. Proc Natl Acad Sci USA 94:12633–12637

Sambrook J, Maniatis T, Fritsch EF (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Plainview, NY

Sandman K, Krzycki JA, Dobrinski B, Lurz R, Reeve JN (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

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

Sandman K, Perler FB, Reeve JN (1994b) Histone-encoding genes from *Pyrococcus*: evidence for members of the HMf family of archaeal histones in a non-methanogenic archaeon. Gene 150:207–208

Sandman K, Grayling RA, Reeve JN (1995) Improved N-terminal processing of recombinant proteins synthesized in *Escherichia coli.* Bio/Technology 13:504–506

Sandman K, Pereira SL, Reeve JN (1998) Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome. Cell Mol Life Sci 54:1350–1364

Starich MR, Sandman K, Reeve JN, Summers MF (1996) NMR structure of HMfB from the hyperthermophile, *Methanothermus fervidus*, confirms that this archaeal protein is a histone. J Mol Biol 255:187–203

Wan JS, Mann RK, Grunstein M (1995) Yeast histone H3 and H4 N termini function through different *GAL1* regulatory elements to repress and activate transcription. Proc Natl Acad Sci USA 92:5664–5668

Wells D, McBride C (1989) A comprehensive compilation and alignment of histones and histone genes. Nucleic Acids Res 17:r311–r346

Zhu W, Sandman K, Lee GE, Reeve JN, Summers MF (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