The EMBO Journal Vol. 21 No. 17 pp. 4654–4662, 2002

Structure of Alba: an archaeal chromatin protein modulated by acetylation

B.N.Wardleworth, R.J.M.Russell, S.D.Bell[1], G.L.Taylor[2] and M.F.White[2]

Centre for Biomolecular Science, St Andrews University, North Haugh, St Andrews, Fife KY16 9ST and [1]MRC Cancer Cell Unit, Hutchison MRC Centre, Hills Road, Cambridge CB2 2XZ, UK

[2]Corresponding authors
e-mail: mfw2@st-andrews.ac.uk or glt2@st-andrews.ac.uk

B.N.Wardleworth and R.J.M.Russell contributed equally to this work

Eukaryotic DNA is packaged into nucleosomes that regulate the accessibility of the genome to replication, transcription and repair factors. Chromatin accessibility is controlled by histone modifications including acetylation and methylation. Archaea possess eukaryotic-like machineries for DNA replication, transcription and information processing. The conserved archaeal DNA binding protein Alba (formerly Sso10b) interacts with the silencing protein Sir2, which regulates Alba's DNA binding affinity by deacetylation of a lysine residue. We present the crystal structure of Alba from *Sulfolobus solfataricus* at 2.6 Å resolution (PDB code 1h0x). The fold is reminiscent of the N-terminal DNA binding domain of DNase I and the C-terminal domain of initiation factor IF3. The Alba dimer has two extended β-hairpins flanking a central body containing the acetylated lysine, Lys16, suggesting three main points of contact with the DNA. Fluorescence, calorimetry and electrophoresis data suggest a final binding stoichiometry of ~5 bp DNA per Alba dimer. We present a model for the Alba–DNA interaction consistent with the available structural, biophysical and electron microscopy data.
Keywords: acetylation/Alba/archaea/chromatin/crystal structure

Introduction

Architectural DNA binding proteins are essential in all cellular organisms, and function both to compact and to regulate the availability of the genetic material. In eukaryotes this role is principally fulfilled by the nucleosome, whilst eubacteria possess double-stranded DNA (dsDNA) binding proteins such as HU (reviewed in Sandman *et al*., 1998). The third domain of life, the archaea, display striking similarities to the eukarya in their information processing pathways (for example DNA replication and transcription) (Bell and Jackson, 1998; Cann and Ishino, 1999). Archaea do not possess recognizable homologues of the eubacterial HU protein, although a subset have histone proteins that form tetrameric nucleosomes wrapping ~70 bp dsDNA (Sandman *et al*., 2001). However, histones are not ubiquitous in archaea and may not be sufficiently abundant to fulfil an equivalent role to that of eukaryotic histones in DNA packaging (reviewed in Sandman and Reeve, 2000). In addition to DNA packaging, eukaryotic nucleosomes play an important role in regulation of the accessibility of chromatin to transcription and repair. Control is asserted by reversible covalent modification of the N-terminal histone tails, which are lysine-rich and accessible in the nucleosome structure (Luger and Richmond, 1998); indeed histone modification is regarded as a fundamental feature of eukaryotic gene expression (reviewed in Struhl, 1999). Histone acetylation by a variety of histone acetyltransferases (HATs) serves to increase the accessibility of the chromatin structure for cellular processes (reviewed in Wu and Grunstein, 2000). On the other hand, hypoacetylation of histone tails leads to a tighter chromatin structure, and transcriptional repression or silencing (Cheung *et al*., 2000; Wu and Grunstein, 2000).

The dsDNA binding protein Alba, formerly named Sso10b, is conserved in most sequenced archaeal genomes, including all the thermophiles and hyperthermophiles whose genomes have been completed, and is distributed both in the euryarchaeota (which encode histone proteins) and the crenarchaeota (which do not) (Forterre *et al*., 1999). Several species, including *Sulfolobus solfataricus*, *Archaeoglobus fulgidus*, *Aeropyrum pernix* and *Methanopyrus kandleri*, have two copies of the Alba gene; the significance of this is unknown at present. A selection of sequences is shown aligned in Figure 1. Homologues of Alba are also found in eukaryotes, including higher plants, trypanosomes and possibly vertebrates (Bell *et al*., 2002).

Alba is a dimeric protein with 10 kDa subunits (Xue *et al*., 2000). The protein is abundant (4–5% of total soluble protein), binds dsDNA tightly but without apparent sequence specificity (Xue *et al*., 2000; Bell *et al*., 2002) and appears from chromatin immunoprecipitation (ChIP) experiments to be distributed uniformly and ubiquitously on the chromosome (S.D.Bell, data not shown)—properties that suggest a role in the architecture of archaeal chromatin. Electron microscopic studies of Alba from *Sulfolobus acidocaldarius* suggest that it binds DNA duplexes without significant compaction, affording protection against degradation by the nuclease DNase I (Lurz *et al*., 1986). A sub-fraction of the protein from *Sulfolobus* cell extracts is found in a stable complex with the silencing protein Sir2 (Bell *et al*., 2002). Sir2 has histone deacetylase (HDAC) activity and is conserved across the three domains of life (Frye, 2000; Imai *et al*., 2000; Landry *et al*., 2000). Analysis by MALDI mass spectrometry of the bulk population of Alba purified from late-log phase *Sulfolobus* cells has demonstrated that most of the protein is present in an acetylated form. The acetylation site has been mapped to Lys16, and acetylation of this

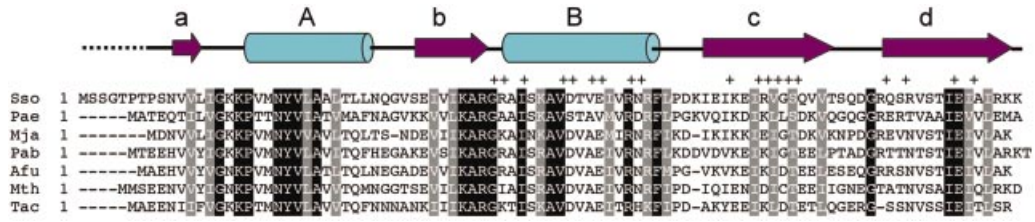


Fig. 1. Sequence alignment of archaeal Alba homologues. More divergent homologues of Alba exist in trypanosomes, higher plants and vertebrates (Bell *et al*., 2002). Sso (*S.solfataricus* P74761), Pae (*Pyrobaculum aerophilum* Q8ZVL3), Mja (*Methanococcus jannaschii* Q57665), Pab (*Pyrococcus abyssi* Q9VIN3), Afu (*Archaeoglobus fulgidus* O28323), Mth (*Methanothermobacter thermautotrophicus* O27527), Tac (*Thermoplasma acidophilum* Q9HJQ5) are aligned. The secondary structure of *Sulfolobus* Alba derived from the crystal structure is shown above the alignment. Residues involved in the main dimer interface are marked with a '+'.

Table I. Crystallographic data and refinement statistics

Data collection	Native hexagonal	Peak hexagonal	Inflection hexagonal	Remote hexagonal	Tetragonal
Wavelength (Å)	1.542	0.979	0.980	0.939	
Resolution (Å)	30–2.6	30–2.7	30–2.7	30–2.7	30.0–2.8
Top shell	(2.69–2.6)	(2.8–2.7)	(2.8–2.7)	(2.8–2.7)	(2.9–2.8)
R_{merge} (%)[a]	4.6 (32.2)	5.3 (31.6)	4.8 (33.4)	6.1 (52.9)	6.8 (44.7)
$\langle I\rangle/\langle\sigma(I)\rangle$	6.9 (2.7)	14.5 (3.0)	14.7 (2.9)	13.1 (1.9)	15.6 (2.4)
Unique reflections	10 996	10 043	10 005	10 042	4115
Completeness (%)	98.3 (94.7)	100 (100)	100 (100)	100 (99.8)	100 (99.3)
Redundancy	7.6	12.5	12.5	12.5	23.3

Refinement	Hexagonal	Tetragonal
Reflections	9675 (work set) 1112 (test set)	3697 (work) 417 (test)
Resolution range (Å)	30.0–2.6	30.0–2.8
Protein atoms	1380	690
Solvent atoms	185	11
R-factor (%)[b] work	23.5	25.1
R-factor (%)[b] free	28.5	29.5
Average *B* value (Å^2)	50 (monomer A) 63 (monomer B)	60
R.m.s.d. bonds (Å)	0.007	0.009
R.m.s.d. angles (°)	1.3	1.3

[a]$R_{merge} = \Sigma|I(k) - \langle I\rangle|/\Sigma I(k)$, where $I(k)$ is the value of the kth measurement of the intensity of a reflection, $\langle I\rangle$ is the mean value of the intensity of that reflection and the summation is over all measurements. Figures in parentheses refer to the highest resolution shells.
[b]*R*-factor = $\Sigma_{hkl}|F_{obs}(hkl) - F_{calc}(hkl)|/\Sigma_{hkl} F_{obs}(hkl)$.

residue strongly reduces the affinity of the protein for DNA (Bell *et al*., 2002). Incubation of acetylated Alba with Sir2 and the NAD^+ co-factor results in deacetylation of the protein and transcriptional repression in an *in vitro* assay, suggesting that mechanisms for the control of gene expression through covalent modification of chromatin components is much more ancient than previously supposed.

We now report the crystal structure of Alba at 2.6 Å resolution, and suggest a model for DNA binding consistent with the available structural and biophysical evidence.

Results

Structure of Alba

To investigate the structure of Alba, we expressed it in *Escherichia coli*, purified the recombinant protein and crystallized it (Wardleworth *et al*., 2001). We now report the structure of a hexagonal crystal form of Alba, solved by MAD phasing, and refined to a resolution of 2.6 Å, containing a dimer in the asymmetric unit (Table I) [Protein Data Bank (PDB) code 1h0x]. We have also determined the structure of a tetragonal crystal form to 2.8 Å, containing a monomer in the asymmetric unit (PDB code 1h0y). The crystal structure of the Alba monomer reveals a mixed α/β structure comprising four β-strands (a to d) and two α-helices (A and B) (Figures 1 and 2). In both crystal forms no electron density is observed for the first eight amino acids, and the only major difference between these two forms is a slight variation in the positioning of the long β-hairpin arm formed by strands c and d. As shown in Figure 2A, this hairpin loop is ordered, with lower than average *B*-factors, in one monomer (monomer A) of the hexagonal crystal form and is involved in extensive contacts with symmetry-related monomers. The hairpin loop of monomer B is disordered and makes no interactions with other monomers in the crystal. The *B*-factor profile of the Alba monomer of the tetragonal crystal form is very similar to that of monomer A in the hexagonal form, with the hairpin loop also involved in extensive crystal contacts. The topology of Alba has been observed previously in the C-terminal

domain of the *E.coli* translation initiation factor IF3 (Biou *et al*., 1995), but with a long β-hairpin arm extending from the body of the protein. Interestingly, Alba also has a high level of structural homology to the N-terminal domain of the DNase I protein (Suck *et al*., 1988), with a root mean square (r.m.s.) fit of 2.1 Å between the C_α atoms of Alba and residues 1–86 of DNase I (Figure 3).

Alba dimer associations

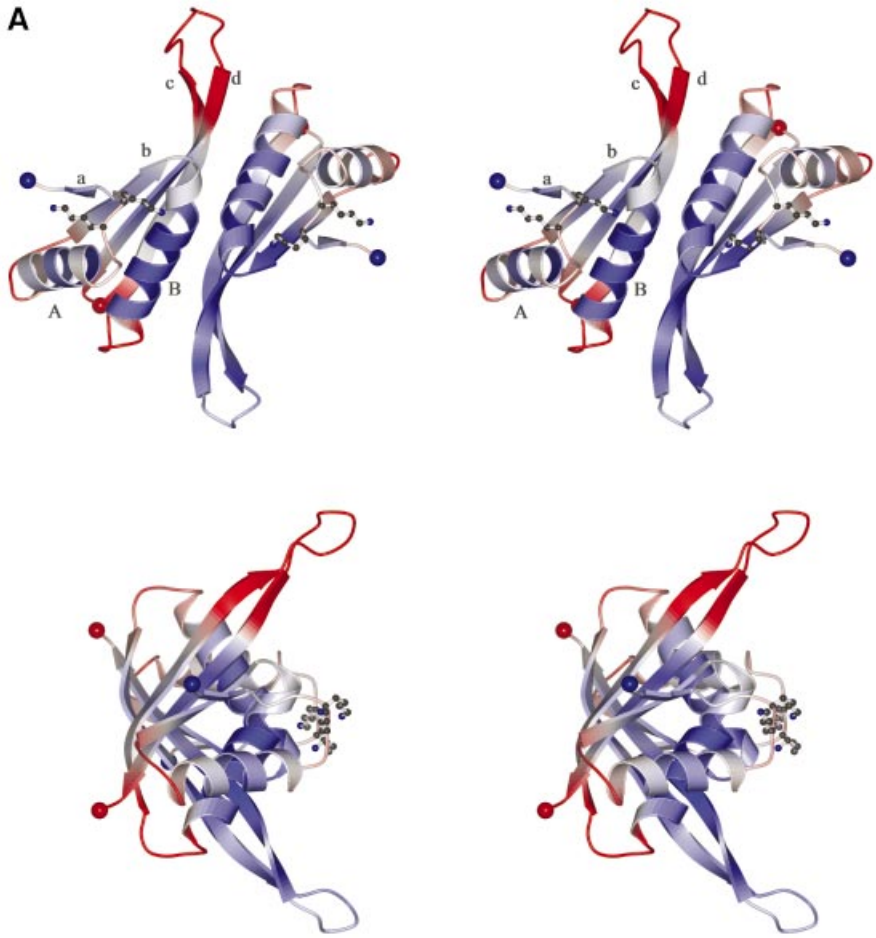
Alba is a dimer in solution and both crystal forms reveal a dimer with a buried surface area of 690 $Å^2$, the interface formed from helix B and strands c and d of both monomers (residues involved in this interface are highlighted by a '+' in Figure 1). The dimer resembles a body with outstretched arms (the β-hairpin loops), and calculation of the electrostatic potential reveals a positively charged surface encompassed by the arms (Figure 2). In both crystal forms there is a second association of monomers with a buried surface area of 460 $Å^2$, the interface formed from helix A and the C-terminal end of helix B and the equivalent regions of a symmetry-related subunit. The potential significance of this second interface is unclear at present, as it does not correspond to the dimer–dimer contacts expected in the model proposed for DNA binding by Alba.

Role of conserved residues

Alba is a highly conserved protein across the crenarchaea and euryarchaea, suggesting a fundamental, conserved function. Many non-polar residues involved in the hydrophobic core of Alba, or the dimer interface, are absolutely conserved (Figure 1). The largest patch of conserved residues on the protein surface is centred on the loop containing Lys16 and Lys17. Tyr22, anchored by flanking residues Asn21 and Val23, plays a particular role in stabilizing this loop, as do Gly15 and Pro18 ('A' in Figure 2B). Lys16 undergoes deacetylation by Sir2 that affects the DNA-binding affinity of Alba, and both Lys16 and Lys17 are known to be important for DNA binding (Bell *et al*., 2002). Conserved Arg42 lies to the sides of the belly of the dimer ('B' in Figure 2B). In the structure presented here, there are several salt-bridges involving Glu36-Lys68, Glu54-Arg57, Asp63-Lys97, Glu66-Arg95 and an ion-pair network involving Lys40-Glu91-Arg71-Glu69. All of these interactions are on the 'back' of the Alba dimer, and may serve to brace and stablilize the structure, as an increase of ion pairs, and in particular networks of ion pairs is a common feature of proteins from hyperthermostable organisms (Karshikoff and Ladenstein, 2001). Only some of these ion-pairs are conserved across the archaea, however, and only Glu91 is totally conserved.

DNA binding by Alba

The interaction of Alba with plasmid DNA was investigated using isothermal titration calorimetry (ITC). Recombinant Alba was titrated into a solution containing supercoiled phiX174 plasmid DNA and the DNA–protein interaction was observed as an exothermic event (Figure 4A and B). The interaction was not easily



interpreted using standard binding models. This was probably due to the cooperative nature of the interaction, as observed previously by gel electrophoretic retardation analysis (Bell *et al*., 2002), with interactions between adjacent Alba dimers contributing to the overall heat effect. The binding stoichiometry could be ascertained clearly, with saturation at ~5–10 bp DNA per dimer of Alba observed at both 25 and 50°C.

A second method used to quantify DNA binding by Alba in solution was the DAPI dye displacement assay first described by Zaitsev and Kowalczykowski (1998). In this experiment, the DNA ligand is first saturated with the minor groove binding fluorescent reporter molecule DAPI. Subsequent addition of a DNA binding protein results in displacement of the DAPI and consequent fluorescence quenching. Recombinant Alba was titrated into a solution containing plasmid phiX174 at 20°C, and DAPI displacement was monitored. The greatest release of DAPI was observed at titration stoichiometries between 30 and 15 bp per Alba dimer, and completion of displacement occurred between 10 and 5 bp per dimer (Figure 4C). The displacement of DAPI by Alba suggests the protein forms interactions in the minor groove of the DNA, or displaces the dye due to distortion of the minor groove.

We also investigated Alba binding plasmid phiX174 DNA by agarose gel electrophoresis. Increasing concentrations of recombinant or native (acetylated) Alba were incubated with supercoiled or relaxed plasmid in binding buffer, followed by separation of bound and free species by electrophoresis through an agarose gel, and visualization by ethidium bromide staining. Increasing concentrations of Alba resulted in a decrease in the mobility of the

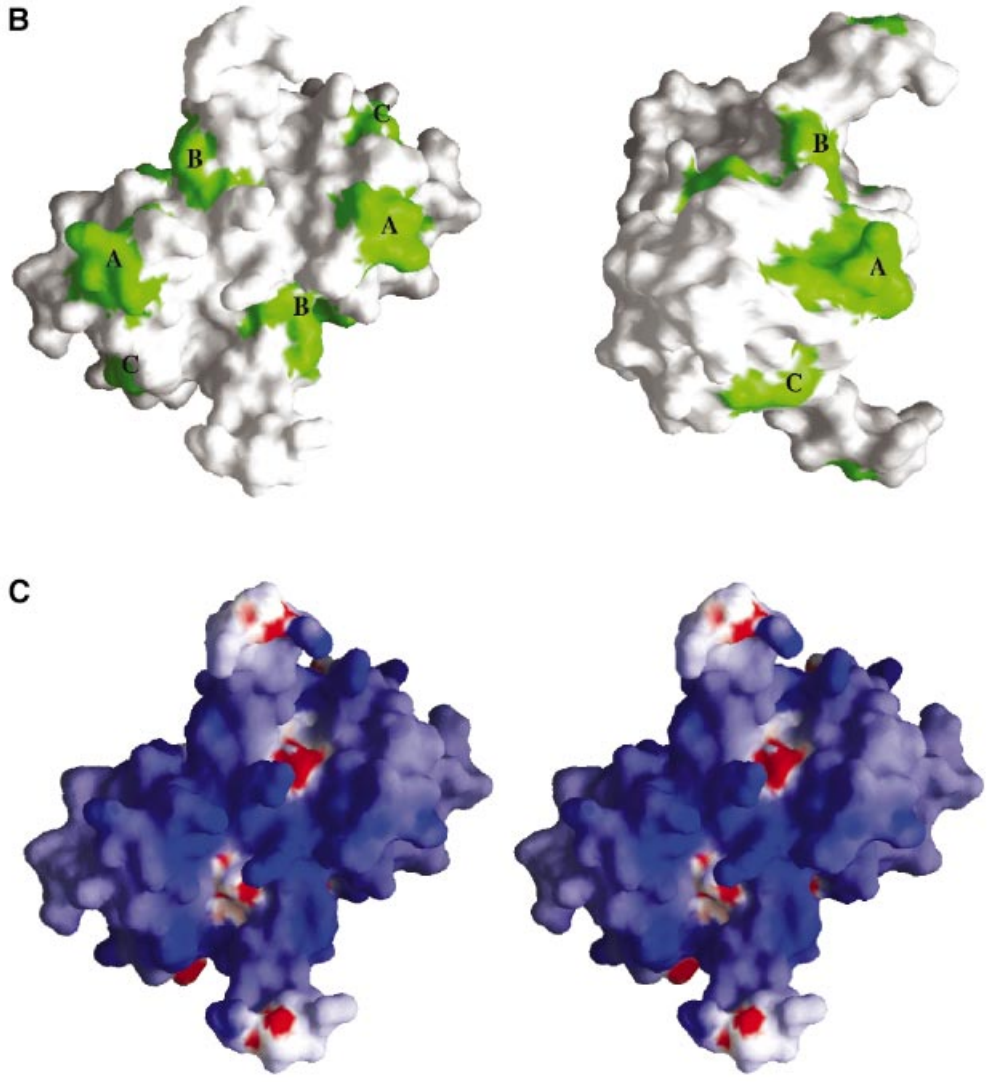


Fig. 2. The Alba dimer. (**A**) Stereo views of the dimer coloured by *B*-factor (dark blue to deep red representing a span of *B*-factor from 30 to 100 Å^2). The lower view is related to the upper view by a 90° rotation around a vertical axis. The strands and helices of one monomer are labelled as in Figure 1. The N- and C-termini are highlighted by blue and red spheres, respectively. Lysines 16 and 17 are also shown. (**B**) Orthogonal views of the dimer showing the location of exposed residues conserved across the Archaea: A (Gly15, Lys17, Pro18, Asn21, Tyr22), B (Lys40, Arg42, Glu91) and C (Phe60). (**C**) Stereo view of the Alba dimer coloured by electrostatic potential. The groove formed between the two loops containing Lys16 and Lys17 is apparent. Figures 2, 3 and 5 were drawn with BOBSCRIPT (Esnouf, 1997) and GL_RENDER (L.Esser and J.Deisenhofer, unpublished).

plasmid in the gel, with the first retardation apparent at a stoichiometry of ~15–20 bp DNA per Alba dimer (Figure 4D). The most significant reduction in DNA mobility was observed at stoichiometries between 5 and 10 bp DNA per dimer. No significant difference in binding stoichiometry was observed between the acetylated and unacetylated forms of the protein, despite the ~30-fold difference in DNA binding affinity of the two species (Bell *et al*., 2002), and there was no observable difference in binding the supercoiled and relaxed forms of the plasmid (Figure 4D; data not shown).

Changes in DNA mobility during agarose gel electrophoresis can result from changes in both the net charge of the migrating complex and the compaction of the DNA (for example, supercoiled DNA migrates considerably more quickly than relaxed plasmid DNA due to its increased compaction). Notably, we did not see any evidence for an increase in DNA mobility on addition of Alba, as has been observed previously for archaeal histone proteins (Sandman *et al*., 1990). In these cases, the increase in mobility has been ascribed to a reduction in the contour length of DNA due to DNA wrapping by the binding protein. Conversely, the data for the Alba–DNA complex support the previous observations by EM that there is no reduction in DNA contour length in Alba complexes (Lurz *et al*., 1986), and thus little or no wrapping of DNA by the protein.

Together these experiments suggest significant changes in the Alba–DNA complex as binding stoichiometries reduce from 20 down to 5 bp DNA per dimer, with saturation occurring between 10 and 5 bp per dimer, independent of the acetylation state of the protein and the presence or absence of negative supercoils. This is in agreement with previous data showing that Alba binds 60 bp duplexes with an initial stoichiometry of ~12 bp per dimer, followed by formation of a more densely packed nucleoprotein complex (Xue *et al*., 2000).

Discussion

The three distinct techniques we used to monitor DNA binding by Alba each suggest a final stoichiometry of 5–10 bp dsDNA per dimer of Alba. The EM and gel electrophoresis data suggest strongly that Alba does not compact DNA significantly. This role may be undertaken by the histone proteins in euryarchaea, and by as yet unknown mechanisms in those archaea (including *Sulfolobus*) that lack histones. This may include other known dsDNA binding proteins such as Sso7d (Agback *et al*., 1998), allowing the creation of higher-order protein–DNA complexes.

Where and how does Alba bind DNA? The co-crystal structure of DNase I with an 8 bp DNA nicked duplex defines the DNase I DNA binding site (Figure 3). In DNase I, a protruding β-hairpin sits in the minor groove of the DNA, and our DAPI displacement data suggest that Alba also binds in the minor groove. The much longer β-hairpin of Alba, however, turns in a different direction, suggesting that although it might sit in the minor groove the DNA would have to be very differently orientated compared with that seen in the DNase I complex. Extended antiparallel β-hairpin arms are a feature of several DNA binding proteins: both the bacterial HU and integration host factor (IHF) proteins have two flexible β-hairpin arms that interact with the minor groove of the DNA, widening it locally and introducing a sharp bend in the bound DNA (White *et al*., 1989; Rice, 1997; reviewed in Jones *et al*., 1999).

In the Alba dimer, the two β-hairpins are separated by ~40 Å, about the length of a single complete turn of B-form duplex DNA. This would allow each hairpin to interact with an equivalent part of the DNA duplex, presumably in the minor groove. The central 'belly' of the Alba dimer would thus be in the correct position to interact with the major groove centred between the two minor grooves. Thus Alba could make three points of contact to a B-DNA helix: a central interaction with the major groove flanked by two interactions with the minor groove (Figure 5b). The central body of the Alba dimer has two loops stabilized by highly conserved residues including Pro18 and Tyr22. These loops are separated by a distance of 22 Å, wide enough to accommodate a DNA duplex (Figures 2A and C, and 5A). At the top of the loops lie residues Lys16 and Lys17, which are known from site-directed mutagenesis studies to be important for DNA binding and are the site for reversible acetylation of the protein (Bell *et al*., 2002). Conserved residue Arg42 would also be in position to interact with the duplex, and may therefore be involved in DNA binding. We have used standard B-form DNA in our model in Figure 5, as we know nothing about the degree or type of DNA distortion that the protein exerts on binding. The model has only minor steric clashes between the DNA duplex and the β-hairpins, and these may be accommodated by changes in the flexible protein arms as well as by local distortion of the DNA.

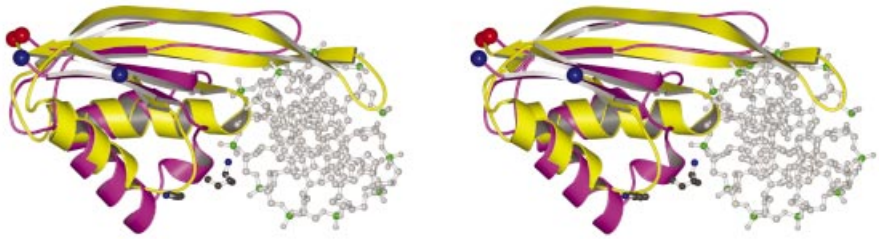


Fig. 3. Comparison of DNase I with Alba. The N-terminal domain, residues 1–86, of DNase I is coloured magenta, in complex with a nicked DNA octamer (PDB code 2DNJ), and showing the β-hairpin that interacts with the DNA minor groove. An Alba monomer is superimposed in yellow, revealing the more extensive β-hairpin of Alba, and suggesting that the orientation of the DNA will be different in the Alba–DNA complex. The phosphorus atoms of the DNA are coloured green, and the side-chains of lysines 16 and 17 of Alba are shown.

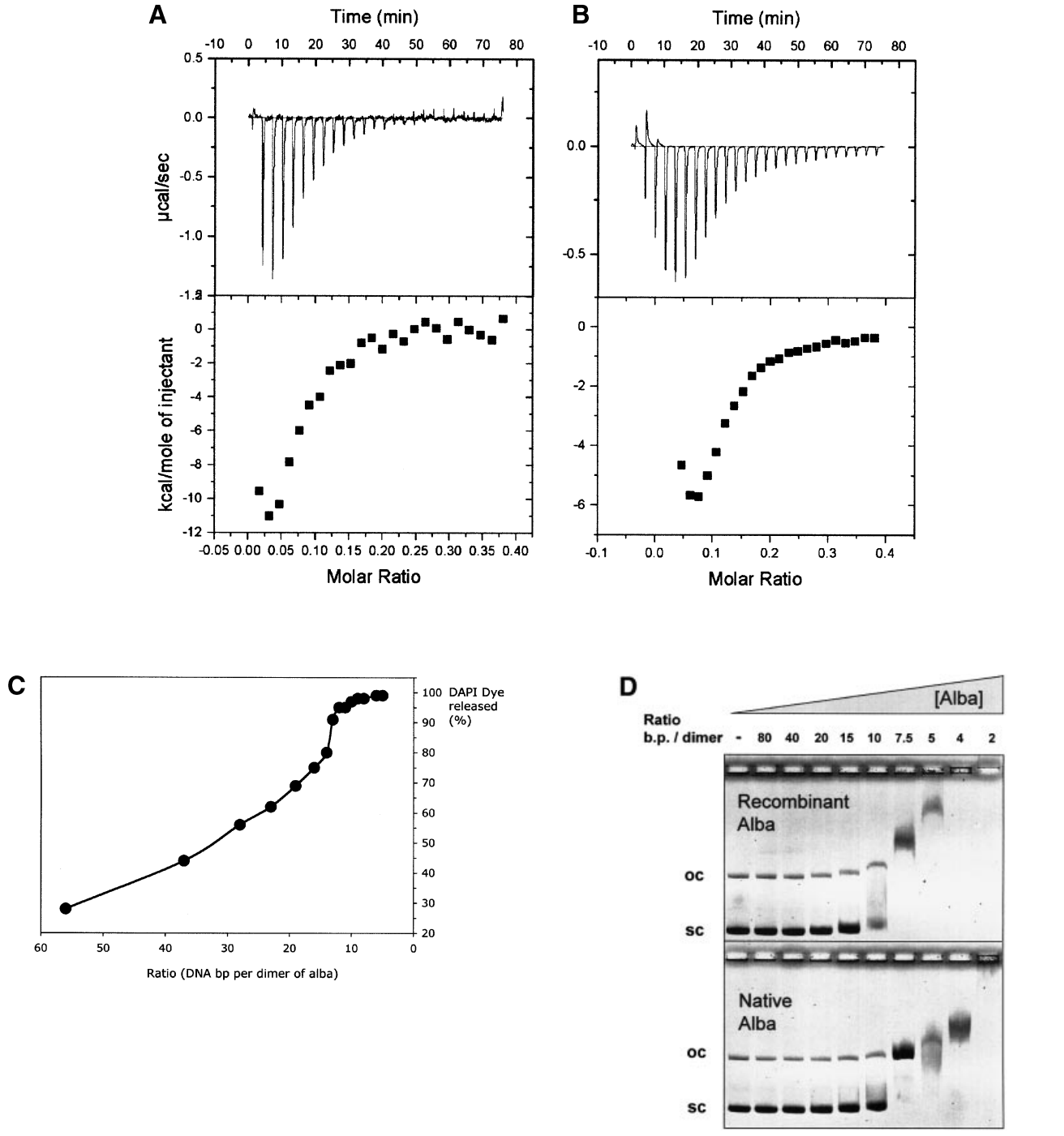
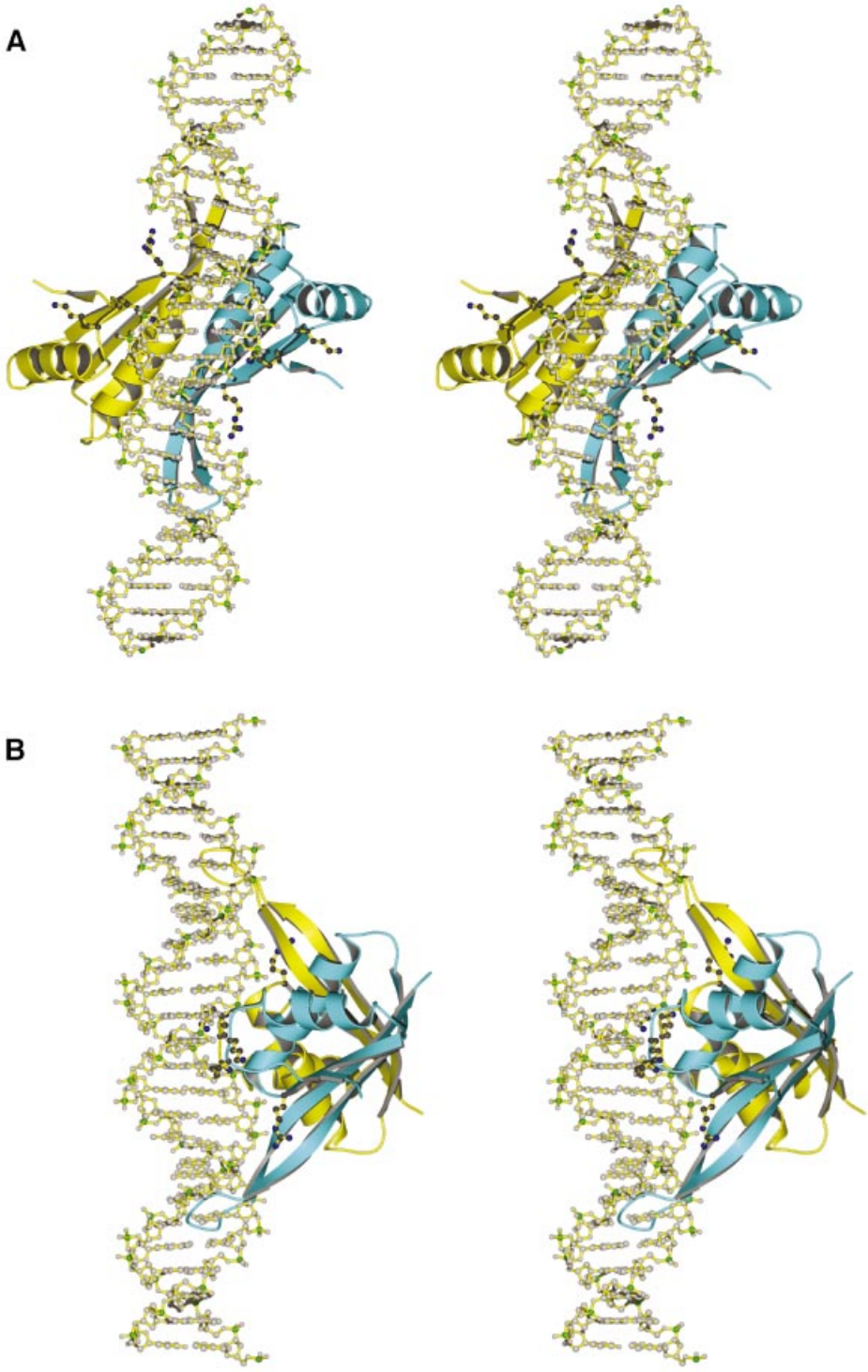


Fig. 4. Interaction of Alba with DNA. (**A** and **B**) ITC. Titration of recombinant Alba into supercoiled pUC18 plasmid DNA at 50°C (A) and 25°C (B). Upper panels: raw data for sequential 10 µl injections of 300 µM recombinant Alba into a solution containing 0.05 µM pUC18 plasmid. Lower panel: integrated heat data. Exothermic binding was observed at both temperatures, with a complex binding curve that could not be fitted using standard models. The *x*-axis denotes the stoichiometry of Alba dimers per base pair of duplex DNA. Saturation of the DNA binding sites was evident at between 10 and 5 bp dsDNA per dimer of Alba. (**C**) DAPI dye displacement assay. Supercoiled phiX174 plasmid (1 nM), saturated with the fluorescent dye DAPI in binding buffer was incubated with increasing concentrations of recombinant Alba protein, and the decrease in fluorescence intensity due to displacement of the DAPI was monitored as described in Materials and methods. Completion of the dye displacement was evident at a ratio of between 10 and 5 bp dsDNA per dimer of Alba. (**D**) Agarose gel electrophoresis. Two hundred and fifty nanograms of supercoiled phiX174 dsDNA (sc) were incubated with increasing concentrations of native (acetylated) or recombinant (unacetylated) Alba prior to electrophoresis in a 0.7% agarose/1× TBE gel. The molar ratio of (base pairs dsDNA to Alba dimers) is indicated. Decreases in the DNA mobility are evident at 20 bp per dimer for both forms of the protein, with the most significant shift occurring between 10 and 7.5 bp per dimer. There is little obvious difference between the stoichiometry of DNA binding of the native and recombinant protein, despite the 30-fold lower binding affinity of the former. The small amount of nicked, open circle (oc) DNA present in the sample is retarded similarly to the supercoiled form (sc).

The model predicts that one dimer of Alba spans around 15 bp of DNA duplex. However, our biophysical measurements suggest that DNA binding is only saturated at much higher protein–DNA stoichiometries. Additional Alba dimers could bind in a similar manner along the duplex, and from simple modelling these would need to be rotated by ~120° around the helix axis to avoid steric clashes. Model building allows the construction of a central DNA helix surrounded by tightly packed dimers of Alba, with a stoichiometry of ~5 bp per dimer, in good agreement with our experimental observations (Figure 5C). This sheathing of the DNA would explain the resistance to DNase I digestion observed previously (Lurz *et al*., 1986).

The model supports the previous mutagenesis data that implicate the acetylated Lys16 residue with a direct role in DNA binding. Thus, we can postulate a mechanism for the control of DNA binding affinity by acetylation of Lys16 due to electrostatic and steric changes in the DNA binding site. This contrasts with the situation in eukaryotic nucleosomes, where lysine modification occurs on N-terminal



extensions and is not thought to influence DNA binding directly. One can speculate that an ancestral form of protein acetylation may have arisen as a means to directly influence the DNA binding affinity of target proteins such as Alba, and that this was subsequently modified over the course of evolution, becoming the sophisticated signalling mechanism we see in present day eukaryotes.

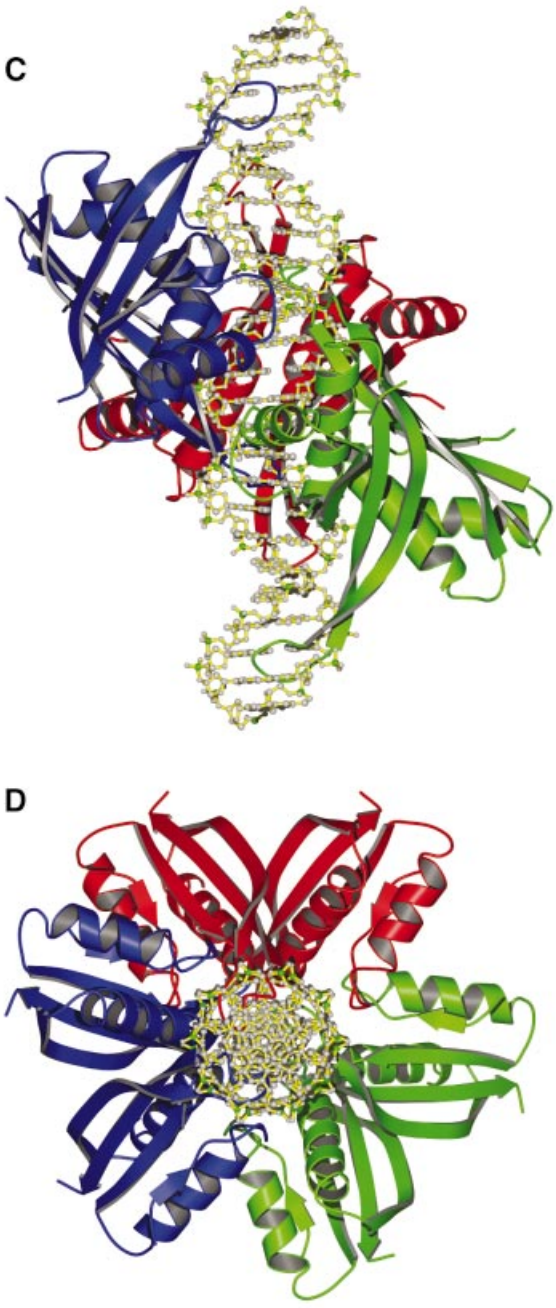


Fig. 5. Models of Alba binding to DNA. (**A** and **B**) Orthogonal stereo views of how one Alba dimer might bind to DNA, with the β-hairpins interacting with the minor grooves and the critical Lys16 sitting either side of the duplex. Residues Lys16, Lys17 and Arg42 are highlighted. (**C** and **D**) Orthogonal views of a model for the binding of multiple Alba dimers along the DNA duplex. Consecutive dimers are rotated by 120° with respect to one another, allowing close packing of Alba molecules along the DNA duplex with minimal steric clash.

It is generally accepted that prokaryotes adopt a 'non-restrictive transcriptional ground state', where promotors are accessible by RNA polymerase unless specifically repressed (Struhl, 1999). In contrast, eukaryotes adopt a restrictive ground state maintained by chromatin and regulated by chromatin modification. The recent identification of covalent modification of the conserved archaeal protein Alba by acetylation, and its interaction with the conserved histone deacetylase Sir2 (Bell *et al*., 2002), challenge these preconceptions and suggest that, at least in archaea, chromatin modification may play an active role in transcriptional silencing. Whilst the role of the Alba protein in chromatin structure and the true significance of its modification by acetylation remain to be determined, the solution of the Alba protein structure adds another piece to the puzzle.

Materials and methods

Crystallography

Recombinant Alba was expressed in *E.coli* and purified as described previously (Wardleworth *et al*., 2001). Crystals of Alba belong to space group $P6_122/P6_522$ with cell dimensions of $a = b = 84.4$ Å, $c = 162.2$ Å as previously reported (Wardleworth *et al*., 2001), and native data were collected to 2.6 Å. For structure solution, the single methionine at position 20 was substituted with seleno-methionine (Se-Met) by expressing the protein in minimal medium under conditions of metabolic inhibition (Van Duyne *et al*., 1993). Crystals of the recombinant Se-Met Alba grew under the same conditions as native protein and were isomorphous with the native. Data were collected at three wavelengths at 100 K on beamline ID14-4 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France). All data were processed using DENZO and SCALEPACK (Otwinowski and Minor, 1997). Heavy atom phasing was carried out using the program SOLVE (Terwilliger and Berendzen, 1999), which identified the location of two selenium atoms, implying a single dimer of Alba in the asymmetric unit with a 70% solvent content. Cross difference Fouriers calculated in SOLVE gave peak heights of 33 sigma for the Se sites, but the overall mean figure of merit to 2.7 Å was only 0.25, albeit with a *Z*-score of 7.5. Density modification with RESOLVE (Terwilliger, 2000) resulted in an overall mean FOM of 0.53 and produced a readily interpretable map at 2.7 Å in space group $P6_522$.

The majority of Alba could be built into the initial map using program O (Jones *et al*., 1991). The initial model was refined against the native data set to 2.6 Å, with NCS-restraints applied between the two monomers, using CNS (Brünger *et al*., 1998), and gave a starting *R*-factor of 36.3%. Cycles of manual rebuilding and refinement produced a final model with NCS restraints being dropped in later rounds of refinement, which lead to a further drop in the free *R*-factor. The final model is missing the first eight N-terminal residues of each monomer, indicating significant flexibility in this region. In addition, residues 78–84 are poorly ordered in one monomer. This region is involved in crystal contacts in the other monomer, presumably reducing the flexibility and thus ordering this region. The final model shows good stereochemistry as defined in PROCHECK (Laskowski *et al*., 1993), with no residues in disallowed regions. Protein folds sharing a similar topology were identified using the DALI programme (Holm and Sander, 1997).

A second crystal form was obtained during attempts to crystallize an Alba–DNA complex. Alba (500 μM) was mixed in a 1:1 ratio with a 20mer oligonucleotide (also 500 μM). Crystals were obtained in 0.3 M NaCl, 0.1 M ammonium sulphate, 0.01 M $MgCl_2$, 0.05 M MES buffer pH 5.6, and 16% PEG8000. The crystals belong to space group $I4_122$ with cell dimensions $a = b = 84.68$ Å, $c = 87.14$ Å. This crystal form was readily solved using molecular replacement, with an Alba monomer in the asymmetric unit, but there was no evidence of bound DNA. Data collection and refinement details are given in Table I. Atomic coordinates and structure factors have been deposited for the hexagonal (1h0x, r1h0xsf) and tetragonal (1h0y, r1h0ysf) crystal forms, respectively.

ITC

ITC experiments were carried out using a VP-ITC device (MicroCal, Northampton, MA). All solutions were degassed. Alba and DNA samples were dialysed extensively against 20 mM Tris–HCl buffer pH 7.5, 10 mM

$MgCl_2$. The binding experiments were performed at 25 or 50°C. A 370 µl syringe with stirring at 400 r.p.m. was used to titrate Alba into a cell containing ~1.4 ml DNA solutions. Each titration consisted of a preliminary 1 µl injection followed by up to 30 subsequent 10 µl injections. Calorimetric data were analysed using MicroCal ORIGIN software.

DAPI fluorimetric dye displacement assay

DAPI (4′,6′-diamidino-2-phenyl-indole) (Molecular Probes) increases in fluorescence upon binding the minor groove of dsDNA. The standard reaction buffer consisted of 20 mM Tris pH 7.5, 10 mM $MgCl_2$ and 400 nM DAPI (400 µl total volume), 6 µM nucleotides and a variable amount of Alba protein. Fluorescence measurements were performed on a Perkin-Elmer LS50B Luminescence Spectrometer. The excitation/emission wavelengths and the slit widths were 345/467 nm and 5/10 nm, respectively. Reaction components were added as follows. After equilibration of the standard buffer at 20°C any background signal due to buffer components was set to zero. The phiX174 dsDNA was added and the resulting fluorescence increased several fold due to the formation of the DAPI–dsDNA complex. Alba was then added and the decrease in fluorescence of the DAPI–dsDNA complex monitored after a 10 min incubation period. Relative fluorescence values were corrected for dilution effects and subtraction of fluorescence due to buffer and DNA components. The percentage of DAPI dye released at each protein concentration was calculated with respect to the maximal dye release observed at saturating protein concentrations.

DNA binding monitored by agarose gel electrophoresis

Agarose gel electrophoresis was used to investigate the stoichiometry of Alba binding to a plasmid DNA molecule. A range of concentrations of purified Alba protein was incubated with 250 ng of supercoiled phiX174 dsDNA in binding buffer (20 mM MES pH 6.5, 100 mM potassium glutamate, 1 mM $MgCl_2$ and 0.1 mg/ml bovine serum albumin), in 10 µl total volume. After 15 min at 20°C, one-sixth volume loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 35% Ficoll type 400) was added, and samples were electrophoresed (in 0.7% agarose, 1× TBE) at 10 V/cm for 90 min. After electrophoresis, gels were stained in ethidium bromide and visualized under UV light.

Acknowledgements

We thank staff at the ESRF, Grenoble, for assistance with data collection, and the EU TMR/LSF grant for allowing access to ESRF. Thanks to Margaret Taylor for technical assistance with the initial crystallization. M.F.W. is a Royal Society University Research Fellow. This work was funded by the BBSRC and the University of St Andrews.

References

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

Bell,S.D. and Jackson,S.P. (1998) Transcription in Archaea. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 41–51.

Bell,S.D., Botting,C.H., Wardleworth,B.N., Jackson,S.P. and White,M.F. (2002) The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science*, **296**, 148–151.

Biou,V., Shu,F. and Ramakrishnan,V. (1995) X-ray crystallography shows that translational initiation factor IF3 consists of two compact α/β domains linked by an α-helix. *EMBO J.*, **14**, 4056–4064.

Brünger,A.T. *et al*. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.

Cann,I.K. and Ishino,Y. (1999) Archaeal DNA replication: identifying the pieces to solve a puzzle. *Genetics*, **152**, 1249–1267.

Cheung,P., Allis,C.D. and Sassone-Corsi,P. (2000) Signaling to chromatin through histone modifications. *Cell*, **103**, 263–271.

Esnouf,R.M. (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.*, **15**, 132–134, 112–133.

Forterre,P., Confalonieri,F. and Knapp,S. (1999) Identification of the gene encoding archeal-specific DNA-binding proteins of the Sac10b family. *Mol. Microbiol.*, **32**, 669–670.

Frye,R.A. (2000) Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochem. Biophys. Res. Commun.*, **273**, 793–798.

Holm,L. and Sander,C. (1997) Dali/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res.*, **25**, 231–234.

Imai,S., Armstrong,C.M., Kaeberlein,M. and Guarente,L. (2000) Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature*, **403**, 795–800.

Jones,S., van Heyningen,P., Berman,H.M. and Thornton,J.M. (1999) Protein–DNA interactions: A structural analysis. *J. Mol. Biol.*, **287**, 877–896.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Karshikoff,A. and Ladenstein,R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a 'traffic rule' for hot roads. *Trends Biochem. Sci.*, **26**, 550–556.

Landry,J., Sutton,A., Tafrov,S.T., Heller,R.C., Stebbins,J., Pillus,L. and Sternglanz,R. (2000) The silencing protein SIR2 and its homologs are NAD-dependent protein deacetylases. *Proc. Natl Acad. Sci. USA*, **97**, 5807–5811.

Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.

Luger,K. and Richmond,T.J. (1998) The histone tails of the nucleosome. *Curr. Opin. Genet. Dev.*, **8**, 140–146.

Lurz,R., Grote,M., Dijk,J., Reinhardt,R. and Dobrinski,B. (1986) Electron microscopic study of DNA complexes with proteins from the Archaebacterium *Sulfolobus acidocaldarius*. *EMBO J.*, **5**, 3715–3721.

Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.

Rice,P.A. (1997) Making DNA do a U-turn: IHF and related proteins. *Curr. Opin. Struct. Biol.*, **7**, 86–93.

Sandman,K. and Reeve,J.N. (2000) Structure and functional relationships of archaeal and eukaryal histones and nucleosomes. *Arch. Microbiol.*, **173**, 165–169.

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.

Sandman,K., Pereira,S.L. and Reeve,J.N. (1998) Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome. *Cell. Mol. Life Sci.*, **54**, 1350–1364.

Sandman,K., Soares,D. and Reeve,J.N. (2001) Molecular components of the archaeal nucleosome. *Biochimie*, **83**, 277–281.

Struhl,K. (1999) Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell*, **98**, 1–4.

Suck,D., Lahm,A. and Oefner,C. (1988) Structure refined to 2 Å of a nicked DNA octanucleotide complex with DNase I. *Nature*, **332**, 464–468.

Terwilliger,T.C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D*, **56**, 965–972.

Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D*, **55**, 849–861.

Van Duyne,G.D., Standaert,R.F., Karplus,P.A., Schreiber,S.L. and Clardy,J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.*, **229**, 105–124.

Wardleworth,B.N., Russell,R., White,M.F. and Taylor,G.L. (2001) Preliminary crystallographic studies of the double-stranded DNA binding protein Sso10b from *Sulfolobus solfataricus*. *Acta Crystallogr. D*, **57**, 1893–1894.

White,S.W., Appelt,K., Wilson,K.S. and Tanaka,I. (1989) A protein structural motif that bends DNA. *Proteins*, **5**, 281–288.

Wu,J. and Grunstein,M. (2000) 25 years after the nucleosome model: chromatin modifications. *Trends Biochem. Sci.*, **25**, 619–623.

Xue,H., Guo,R., Wen,Y., Liu,D. and Huang,L. (2000) An abundant DNA binding protein from the hyperthermophilic archaeon *Sulfolobus shibatae* affects DNA supercoiling in a temperature-dependent fashion. *J. Bacteriol.*, **182**, 3929–3933.

Zaitsev,E.N. and Kowalczykowski,S.C. (1998) Binding of double-stranded DNA by *Escherichia coli* RecA protein monitored by a fluorescent dye displacement assay. *Nucleic Acids Res.*, **26**, 650–654.

Received May 13, 2002; revised July 5, 2002;
accepted July 16, 2002