Characterization of Sac10a, a Hyperthermophile DNA-Binding Protein from *Sulfolobus acidocaldarius*†

Stephen P. Edmondson,*,‡ Mebrahtu A. Kahsai,‡ Ramesh Gupta,§ and John W. Shriver*,‡

Laboratory for Structural Biology, Departments of Chemistry and Biological Sciences, Graduate Program in Biotechnology and Bioengineering, University of Alabama in Huntsville, Huntsville, Alabama 35899, and Department of Biochemistry and Molecular Biology, School of Medicine, Southern Illinois University, Carbondale, Illinois 62901

Received April 23, 2004; Revised Manuscript Received August 12, 2004

ABSTRACT: Sac10a is a member of a group of basic DNA-binding proteins thought to be important in chromatin structure and regulation in the archaeon *Sulfolobus.* We describe here the isolation, gene identification, and biophysical characterization of native Sac10a. The protein exists as a 23.8 kDa homodimer at pH 7 and unfolds with a $T°$ of 122 °C. Dissociation of the dimer into folded globular subunits is promoted by decreased pH and salt concentration. Thermal unfolding of the monomeric subunits occurred with two transitions, indicating two independent domains. The dimer demonstrated a high affinity for duplex poly(dAdT) with a K_D of 5×10^{-10} M and a site size of 17 bp (in 0.15 M KCl, pH 7), with only weak binding ($K_D > 5 \times 10^{-6}$ M) to poly(dA)-poly(dT), poly(dGdC), poly(dG)-poly(dC), and *Escherichia coli* DNA under similar conditions. Binding to poly(dAdT) resulted in distortions in the DNA duplex that were consistent with overwinding as indicated by inversion of the CD spectrum of the DNA. The monomeric subunits are predicted to adopt a winged helix DNA-binding motif which dimerizes through formation of a two-stranded coiled coil involving an extended C-terminal helix with more than four heptad repeats (about 45 Å in length). This is the first example of the conserved archaeal transcription regulator domain COG3432 to be characterized. Sequences for homologous proteins containing both COG3432 and predicted coiled coil domains occur in the genomes of both crenarchaeota (*Sulfolobus*, *Pyrobaculum*, *Aeropyrum*) and euryarchaeota (*Methanosarcina*, *Methanococcus*, *Archaeoglobus*, *Thermoplasma*), with multiple genes in some species. Sac10a shows no sequence similarity to the other *Sulfolobus* chromatin proteins Sac7d, Sac8, Sso10b2, and Alba.

A number of small DNA-binding proteins that are thought to be important in DNA packaging and gene expression have been isolated from *Sulfolobus* (*1*). These are of interest not only due to their ability to function at high temperature but also because aspects of the regulation of archaeal chromatin are similar to those observed in eukaryotes (*2*). The best characterized are Sac7d and Sso7d from *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, respectively (*3*−*7*). These are highly basic, monomeric proteins that bind noncooperatively and without sequence specificity to double stranded DNA (*8*, *9*). They are commonly believed to be chromatin proteins, although their physiological functions have not been defined. Both have proven to be ideal for studies of protein stability as well as the energetics of DNA binding and bending (*10*−*12*).

Significant attention has been focused more recently on the basic 10 kDa proteins of *Sulfolobus* initially characterized by Dijk and Reinhardt (*13*). Sso10b, also called Alba, has a number of properties that indicate that it is probably important in regulating chromatin structure (*2*). For example, the homologous protein Ssh10b has been shown to induce negative supercoils in double stranded DNA (*14*), and its affinity for DNA is regulated by lysine acetylation (*15*). Alba shows a high affinity for an intrinsic Sir2, a deacetylase that is believed to be important in regulating histones in eukaryotes (*15*). Deacetylation of Alba by Sir2 leads to an increase in DNA affinity and repression of transcription in vitro. Crystal structures of Alba and a homologous protein (Sso10b2) have been reported along with proposed structures of the DNA complex (*16*, *17*).

An additional 10 kDa protein known as Sac10a was isolated from *Sulfolobus acidocaldarius* by Reinhardt and co-workers which appeared to be different from Sac10b (Alba) (*13*). Electron micrographs of DNA complexes indicated that DNA binding by Sac10a led to specific structures that differed significantly from that of Sac10b. Sac10a formed clumps at low protein/DNA ratios, but evenly covered double stranded DNA at higher concentrations (*13*, *18*). Most interestingly, Sac10a led to "supertwisting" of closed circular DNA (*18*). These observations, as well as interest in the energetics of thermophile protein folding and DNA binding, have prompted us to initiate a more thorough characterization of the properties of Sac10a and the homologous protein Sso10a from *Sulfolobus solfataricus*. An initial characterization of crystals of Sso10a has been presented (*19*).

† This work was supported by Grant GM49686 from the National Institutes of Health to J.W.S. and S.P.E. and Grant GM55945 to R.G.

* Authors to whom correspondence should be addressed. J.W.S.: tel, 256-824-2477; fax, 256-824-6349; e-mail: shriverj@uah.edu, edmonds@uah.edu.

‡ University of Alabama in Huntsville.

§ Southern Illinois University.

10.1021/bi0491752 CCC: $27.50
Published on Web 09/24/2004

Consistent with the renaming of the Sac7d family of proteins as Sul7d (*2*),[1] we will refer to Sac10a and its homologues as Sul10a. Individual members of the family will continue to be referred to using the current names, e.g. Sac10a and Sso10a.

MATERIALS AND METHODS

Materials. Synthetic polynucleotides were obtained from Amersham Pharmacia (Piscataway, NJ) and had mean lengths of 500 to 1500 nucleotides. Genomic *Escherichia coli* DNA (Sigma Chemical Co., St. Louis, MO) was sheared to moderate lengths by being passed several times through a 26-gauge needle at 4 °C. All DNA samples were dialyzed against the appropriate buffer before use. Concentrations were determined from UV absorbance using extinction coefficients reported previously (*9*). ANS (1-anilino-8-naphthalene sulfonate) was obtained from Fluka (Buchs, Switzerland).

Purification and Characterization. S. acidocaldarius was obtained from W. Zillig and was originally referred to as *S. solfataricus* P1, but was later shown to be *S. acidocaldarius* (*8*, *20*). The strain maintained in this laboratory is referred to as *S. acidocaldarius* RGJM (*8*). Cells were grown as previously described (*8*) to a density sufficient to give an A_{600} of 1.0, cells were harvested by centrifugation, and Sac10a protein was purified by methods similar to those used for purification of native Sac7d. Following dialysis of the soluble fraction of a cell lysate against 0.2 M H_2SO_4, the acid soluble proteins were fractionated by cation exchange chromatography using HP-SP Sepharose (Pharmacia) with a linear 0 to 0.3 M NaCl gradient in 0.01 M KH_2PO_4 buffer (pH 7.0) at room temperature. Sac10a eluted as a single peak at about 0.2 M NaCl. The protein was further purified by reverse phase HPLC using a 3 mL Resource 15RPC (Pharmacia) column with 0.065% TFA and a linear (0−0.05%) gradient of acetonitrile. No change in the CD spectrum of the preparation was noted upon exposure to the TFA and acetonitrile concentrations used here. Sac10a eluted at 0.02% acetonitrile and was stored at 4 °C in the elution solution until needed. Aliquots were dialyzed against the appropriate buffer before use.

The purity and initial estimate of molecular weight of native Sac10a were obtained using SDS/tricine PAGE (14% acrylamide). An accurate molecular weight of the protein was determined by MALDI mass spectrometry (Protein Sciences Facility, University of Illinois). The molecular weight under native conditions in solution was determined by gel filtration using a Superdex 75 HR 10/30 (Pharmacia) column in 0.01 M KH_2PO_4 buffer, 0.15 M NaCl, pH 7.0. The column was calibrated with molecular weight standards (Pharmacia) consisting of aprotinin, cytochrome *C*, carbonic anhydrase, and bovine serum albumin. Trypsin fragments of Sac10a were prepared and sequenced by the Protein Sequencing and Peptide Mapping Laboratory at the University of British Columbia.

[1] Abbreviations: ANS, 1-anilino-8-naphthalene sulfonate; DSC, differential scanning calorimetry; Sul7d, family of 7 kDa basic, DNA-binding proteins from *Sulfolobus* including Sac7d and Sso7d from *S. acidocaldarius* and *S. solfataricus*, respectively; Sul10a, family of 10 kDa basic, DNA-binding proteins from *Sulfolobus* including Sac10a and Sso10a from *S. acidocaldarius* and *S. solfataricus*, respectively; TFA, trifluoroacetic acid.

Sequence Analysis and Structure Prediction. The Sac10a sequence was analyzed using the ProtParam tool (http://us.expasy.org/tools) to obtain basic protein properties including the protein extinction coefficient. Sequence alignment was performed with ClustalX (*21*). Secondary structure prediction was performed with SSPro (*22*), the tendency for coiled coil was predicted using COILS (http://www.ch.embnet.org/software) (*23*), and leucine zipper prediction was performed with 2ZIP (http://2zip.molgen.mpg.de/index.html) (*24*). Similar sequence searches were performed with the NCBI Web site using tblastn (http://www.ncbi.nlm.nih.gov:80/BLAST). Structure prediction was performed with 3D-pssm fold recognition server (http://www.sbg.bio.ic.ac.uk/~3dpssm), the protein−protein BLAST tool (http://www.ncbi.nlm.nih.gov:80/BLAST), and the conserved domain database tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml).

Analytical Ultracentrifugation. Sedimentation equilibrium was performed on a Beckman X-LA analytical ultracentrifuge using an An50Ti rotor and six channel centerpiece with speeds of 20K, 28K, and 35K rpm at 20 °C. Protein concentrations were 0.02, 0.04, and 0.07 mg/mL. Equilibrium concentration gradients were determined from the absorbance at 230 nm. The data were fit globally using WinNonLin (*25*) (with $\bar{\nu} = 0.73$ cm^3/g) both to a monomer model and to a monomer ↔ dimer association model in which the monomer molecular weight was fixed at 11860. The pH and salt dependence of the monomer ↔ dimer equilibrium constant was investigated at a single protein concentration (0.05 mg/mL) at 35K rpm and 20 °C.

Sedimentation velocity experiments were performed at 50000 rpm using two channel centerpieces with an An50Ti rotor. Protein concentrations ranged from 0.04 to 0.8 mg/mL. The homogeneity of the samples was checked with the van Holde−Weischet method (*26*) using the Ultrascan program. All programs used for analysis of centrifugation data (including Ultrascan, Svdberg, and Sednterp) were obtained from the Reversible Associations in Structural and Molecular Biology Web site (http://www.bbri.org/RASMB/rasmb.html). Sedimentation and diffusion coefficients were determined using the program Svdberg (*27*) and extrapolated to zero protein concentration. Sedimentation parameters $s_{20,w}$, $D_{20,w}$, f/f_o, and model dimensions were calculated using Sednterp.

Circular Dichroism. CD spectra and thermal melts were measured using an AVIV 62DS spectropolarimeter as previously described (*8*). The fraction of secondary structure was calculated using the CDPro software package (http://lamar.colostate.edu/~sreeram/CDPro/) (*28*), and the results from three fitting procedures (cdsstr, contin, selcon) were averaged. A reference set of 42 proteins over the wavelength range of 240 to 185 nm was used for the data obtained in 0.15 M NaCl at pH 7.0. A reference set of 48 proteins with a wavelength range of 240 to 190 nm was used for CD spectra measured at different temperatures (5 to 90 °C) in 1 mM glycine at pH 2.5. The reference sets included contributions from unfolded proteins. No distinction was made between the regular and distorted conformations of helix and sheet.

Differential Scanning Calorimetry. DSC was performed on VP-DSC (MicroCal, Northampton, MA) and Nano II microcalorimeters (Calorimetry Sciences, Provo, UT) at a

scan rate of 1 °C/min, as previously described (*29*). Baseline scans of buffer vs buffer were subtracted from the sample data and converted to molar heat capacity per protein monomer. The heat capacity data were fit to a model consisting of a dimer unfolding directly to two random coils using the procedure IgorDenat (written in IgorPro (Wavemetrics, Inc.) and available at http://daffy.uah.edu/thermo) according to the following equations (*30*):

$$C_p(T) = C_{\text{excess}} + C_{\text{baselines}} \quad (1)$$

$$C_{\text{excess}} = \beta[\Delta H^\circ - \Delta C_p(T^\circ - T)]\frac{\partial \alpha}{\partial T} \quad (2)$$

$$C_{\text{baselines}} = (1 - \alpha)[a + b(T^\circ - T)] + \alpha[c + d(T^\circ - T)] \quad (3)$$

where parameters *a*, *b*, *c*, and *d* are the intercepts and slopes of the pre- and post-transition baselines. ΔH° is the standard state enthalpy of dimer unfolding at T°, the reference temperature at which $\Delta G_{\text{un}} = 0$ for a 1.0 M protein solution, and β is the calorimetric to van't Hoff enthalpy ratio ($\Delta H_{\text{cal}}/\Delta H_{\text{vh}}$). The change in heat capacity for dimer unfolding, ΔC_p, was fixed to the value 3200 cal/(mol·K) estimated from the expected change in accessible surface area calculated from the molecular weight as described by Myers et al. (*31*). The progress of the unfolding reaction, α, as a function of temperature is given by

$$\alpha(T) = \frac{[\text{M}]}{2[\text{D}] + [\text{M}]} = \frac{\sqrt{K_{\text{un}}[\text{D}]}}{[\text{P}_t]} \quad (4)$$

where [M], [D], and [P_t] are the concentrations of monomer, dimer, and total protein and K_{un} is the equilibrium constant for unfolding, $K_{\text{un}} = \exp(-\Delta G^\circ/RT)$, where

$$\Delta G^\circ = \Delta H^\circ\left(\frac{T^\circ - T}{T}\right) - \Delta C_p(T^\circ - T) + T\Delta C_p \ln\left(\frac{T^\circ}{T}\right) \quad (5)$$

Fluorescence Spectroscopy. ANS fluorescence spectra were measured as previously described (*32*) using protein and ANS concentrations of 1.0 and 250 μM, respectively.

DNA Binding. Reverse titrations of Sac10a protein with DNA were measured at 25 °C on an SLM 8000C spectroflourimeter with excitation at 295 nm and emission at 355 nm, as previously described for Sac7d (*9*). Intensities were corrected for dilution and the inner filter effect using the extinction coefficients of DNA and Sac10a as described elsewhere (*33*). Data were normalized to $Q = (F - F_o)/F_o$ where F was the observed fluorescence and F_o the initial fluorescence. Titrations were performed at several different starting protein concentrations (1 to 6 μM), and each data set was fit to the McGhee−von Hippel model (*34*) using nonlinear regression. Forward titrations were performed by measuring CD spectra from 320 to 200 nm in a 1 cm path length cell with sequential addition of Sac10a to a solution of DNA at an initial concentration of about 60 μM nucleotides.

RESULTS

Protein Purification. Native Sac10a was extracted from *S. acidocaldarius* cells and purified using essentially the procedure developed for Sac7d (*8*), i.e. an initial fractionation with acid precipitation of other soluble proteins in a cell lysate, followed by cation exchange chromatography. Sac10a was identified as described by Grote et al. (*1*) as a 10 kDa protein (apparent MW defined by SDS PAGE) which eluted prior to Sac7d from a cation exchange column using a linear salt gradient. Sac10a prepared as described in Materials and Methods did not contain detectable amounts of contaminating proteins as judged by SDS gel electrophoresis and reverse phase HPLC. The protein was readily isolated in exponentially growing cells with yields comparable to Sac7d (approximately 10 mg of protein from 16 L of *S. acidocaldarius* grown to a cell density of 1 OD at 600 nm). The yield of protein was substantially reduced at high cell densities as growth reached steady-state phase. For reasons that are not understood, native Sso10a could not be obtained from *S. solfataricus* cells using similar procedures. The molecular weight of native Sac10a was estimated by SDS gel electrophoresis to be 10500. MALDI mass spectrometry indicated a mass of 11858 Da.

Protein Sequence. Amino terminal sequence information was not provided for native Sac10a in the initial characterization of the protein due to a blocked amino terminus (*1*, *13*), and this was confirmed here. Four trypsin fragments of native Sac10a were prepared and sequenced (Figure 1), and three were used to search the *S. acidocaldarius* genomic database for the Sac10a gene (Roger Garrett, personal communication). One open reading frame coded for a protein in exact agreement with the three sequences (total of 31 residues). Perfect agreement with the fourth test sequence (11 residues) was taken as confirmation that this ORF (gid: 16208 saci_707905−708213) coincided with the Sac10a gene and provided the Sac10a sequence (Figure 1). Another ORF (gid:17293 saci_1559631−1559362) showed less sequence similarity with the four fragments with 36% identity and 24% conservative substitutions compared to Sac10a. As shown below, this most likely codes for a homologue of Sac10a which is referred to here as Sac10a2. The Sac10a gene codes for a basic, 102 residue protein with two methionines at the N-terminus. Removal of one methionine gives a calculated molecular weight of 11898, or 40 mass units greater than that determined for the native protein. With two methionines removed the predicted molecular weight is 11767, indicating that both methionines are removed in the native protein and additional posttranslational modifications exist (consistent with the N-terminal blockage noted above). Sac10a contains 16 lysines, 4 arginines, 7 aspartates, and 8 glutamates with a net charge at pH 7 of +5 (plus the charge on the histidine) and a calculated p*I* of 9.21. The protein contains single tryptophan, cysteine, and histidine residues. The calculated absorbance (*35*) for a 1 mg/mL solution in 6 M GdnHCl was 1.412 at 278 nm, and this value was used for determining protein concentrations.

Secondary Structure. Protein secondary structure predicted from the sequence indicates that Sac10a is predominantly α-helical (Figure 1), with only short β-sheet segments located at the N-terminus and near the middle of the sequence prior to an extended (33 residue) C-terminal α-helix. The C-terminal helix is predicted to form a two-stranded coiled coil, with a probability greater than 0.15 predicted by COILS (*23*) from D69 to E101 using a window size of 28 residues. This comprises more than 30% of the sequence, spanning more

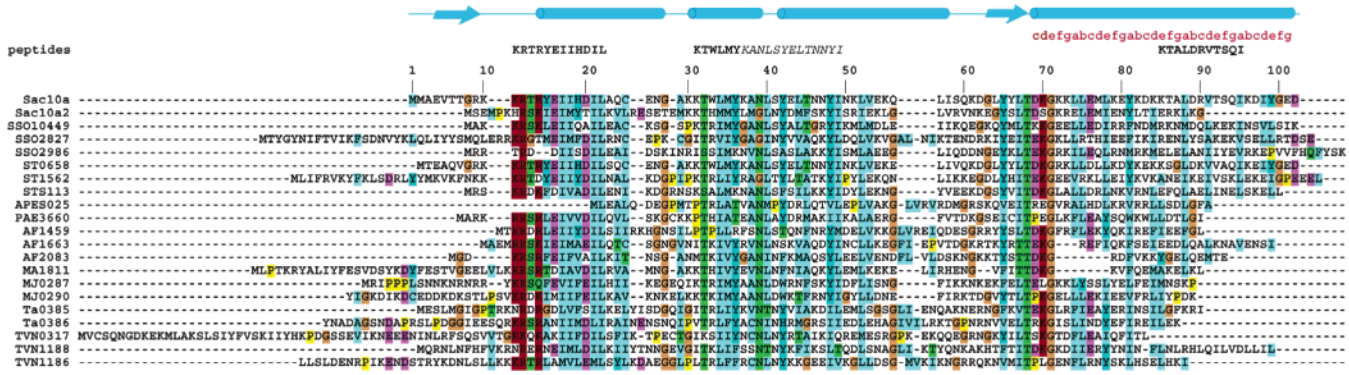

FIGURE 1: The primary sequence of Sac10a and sequence alignments of COG3432 proteins. Sequences of four tryptic peptide fragments used to locate the Sac10a gene are indicated above the Sac10a protein sequence that was derived from the gene. The second (KTWLMY) and third (KANLSYELTNNYI) fragments are contiguous and are distinguished by bold and italic type. Predicted α-helix and *β*-sheet regions in Sac10a are delineated with cylinders and arrows, respectively, above the sequence alignment, and the coiled coil heptad repeat sequence is indicated in red text. Sequence alignments of members of COG3432 (labeled with the gene locus tags listed in Table 2) were obtained using ClustalX. SSO10449 is the locus tag for Sso10a (*19*). Only residues 111−211 are shown for MJ0290, and 121−229 for Ta0386. Sequence similarity is indicated with color highlighting using the default ClustalX parameters (red, 60% K and R; purple, 50% D and E; dark blue, 60% aromatic; light blue, 60% hydrophobic; green, 80% hydrophilic uncharged; all P are yellow and G are orange).

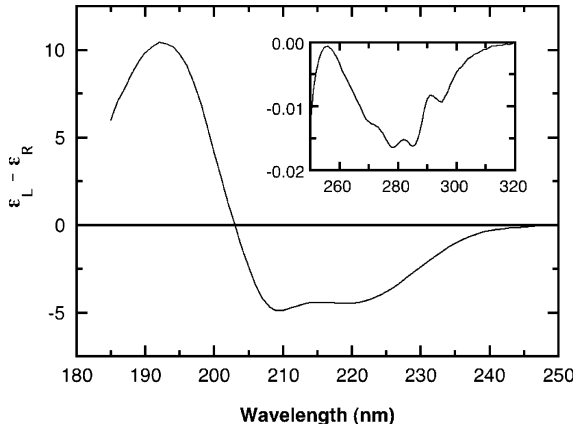

FIGURE 2: CD spectrum of Sac10a at pH 7.0 and 20 °C (0.01 M KH_2PO_4, 0.15 M NaCl, 1 mm path length). The near UV spectrum (with expanded scale) is shown in the inset (1 cm path length).

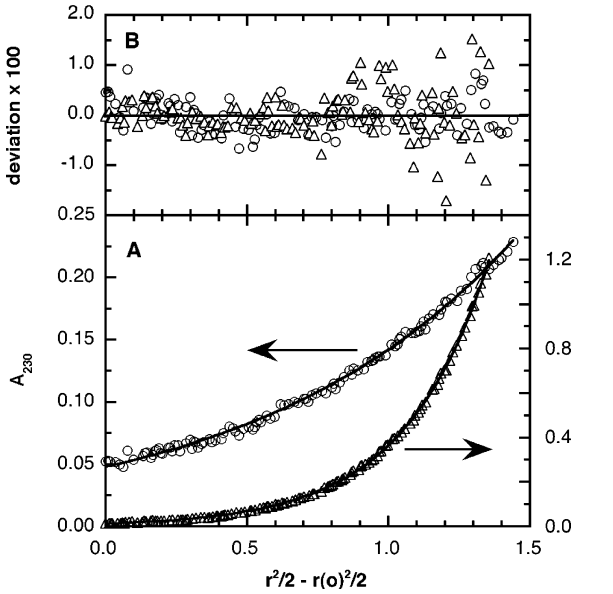

FIGURE 3: (A) Equilibrium centrifugation of Sac10a at pH 7 and 20 °C (0.01 M KH_2PO_4, 0.15 M NaCl). Solid lines result from a global fit of 9 equilibrium profiles (with differing concentrations and rotor speeds) assuming a single molecular species. Only two data sets from the two extreme conditions are shown for clarity: 0.02 mg/mL at 20 K rpm (○) and 0.07 mg/mL at 35 K rpm (△). (B) Residuals of the fits were randomly distributed. The abscissa is offset by the first data point to facilitate comparison.

than 4 heptad repeats with an expected length of about 45 Å. The heptad repeat pattern (Figure 1, red text above the Sac10a sequence) indicates that the expected position of hydrophobic residues at the "*a*" and "*d*" positions in the heptad repeat is occasionally violated, with lysine at the first and third "*d*" positions. However, this is not surprising given that naturally occurring coiled coils can deviate significantly from ideality (*36*) (For an extensive statistical survey of coiled coils and amino acid usage, see http://www.biols.susx.ac.uk/Biochem/Woolfson/html/coiledcoils/). No tendency for formation of a leucine zipper was predicted by the program 2ZIP (*24*).

Analysis of the far UV CD spectrum of native Sac10a (Figure 2) confirmed that the protein contains significant α-helix (49 ± 1%), with 14 ± 3% *β*-sheet, 16 ± 1% *β*-turn, and 21 ± 3% other or unordered structure. No significant differences in CD were detected in phosphate buffer (pH 7.0), with and without 0.15 M NaCl, or in 1 mM glycine (pH 5.6) with no added salt. No concentration dependence in the far UV CD intensity was detected from 0.01 to 1.0 mg/mL in either glycine or phosphate buffer. The near UV CD spectrum (Figure 2, inset) was also invariant under these conditions, indicating no change in interactions surrounding the aromatic side chains.

Oligomerization. Gel filtration chromatography of the native protein showed a single peak in 0.15 M NaCl (pH 7.0) indicating an apparent molecular weight of approximately 17000, and suggesting that the protein existed as a dimer in solution. Sedimentation equilibrium data (9 profiles) provided a molecular weight of 22760 ± 460 in 0.15 M NaCl at pH 7.0 (Figure 3). There was no significant improvement in the quality of the data analysis using a monomer ↔ dimer equilibrium model, indicating that there was no detectable monomer at physiological salt concentrations and neutral pH. A statistical analysis placed a lower limit on the dimer association constant of 10^8 M^{-1} in 0.15 M NaCl at pH 7.

At lower pH, and especially at low salt concentrations, a monomer ↔ dimer association model was necessary to fit the sedimentation equilibrium data. Figure 4 shows the pH

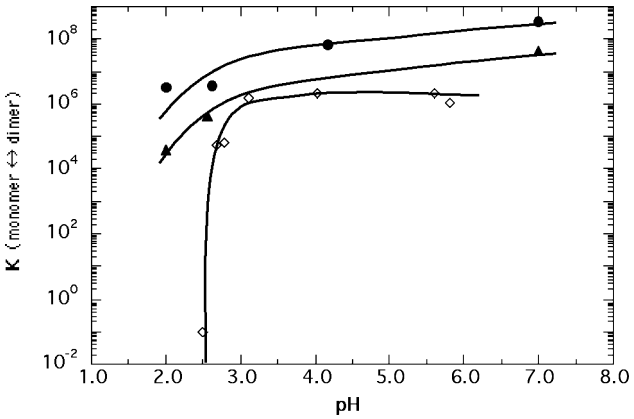

FIGURE 4: Dimerization association constants determined by fitting individual equilibrium centrifugation profiles obtained at various pH values in 0.01 M KH_2PO_4, 0.15 M NaCl (●); 0.01 M KH_2PO_4, no salt (▼); 1 mM glycine (◇). Protein concentration was 0.05 mg/mL. Centrifugation was at 20 °C, 35K rpm. The results for Sac10a in 0.15 M NaCl at pH 7 represent a lower limit to K_{eq} since an equally good fit can be obtained by fitting the data as a single species with twice the molecular weight.

dependence of the monomer ↔ dimer association constant in the absence of salt (1 mM glycine) and in phosphate buffer, with and without 0.15 M KCl. In 0.01 M phosphate buffer, the dimerization association constant was at least 10 times greater in 0.15 M NaCl than in the absence of added NaCl, most likely indicating the importance of hydrophobic interactions at the dimer interface. At the higher salt concentration, the equilibrium constant decreased upon decreasing the pH from 7 to 2 by nearly 1000-fold. In the absence of salt (1 mM glycine), the affinity dropped precipitously below pH 2.7. The sedimentation equilibrium data indicated that native Sac10a exists predominantly as a monomer at 1 μM total protein in 1 mM glycine at pH 2.5. Equilibrium ultracentrifugation demonstrated that Sac10a remains monomeric under these conditions in the absence of salt from 5° to 40 °C (data not shown). The temperature dependence of the CD (see below) indicated that the monomer was fully folded at pH 2.5 in the absence of salt at 10 °C.

Hydrodynamic Properties. The sedimentation coefficient ($s_{20,w}$) of Sac10a determined by sedimentation velocity experiments at pH 7.0 was 2.02 ± 0.04, and was independent of both protein and salt concentration (Figure 5). The ratio of the frictional coefficient to that of a sphere with the equivalent dimer volume, f/f_o, was 1.41, indicating that the shape of the Sac10a dimer deviates significantly from a sphere. Fits of the data to various model shapes consistently indicated that the protein is highly asymmetric, with extreme axial ratios of about 9:1 for an oblate ellipsoid and 7.6:1 for a prolate ellipsoid. Less extreme axial ratios would result for ellipsoids intermediate between the two limits (*37*), but the asymmetry remains. The asymmetry of the Sac10a dimer explains the lower than expected molecular weight observed by gel filtration. The asymmetry is also consistent with the prediction of an extended coiled coil involving the C-terminal third of the sequence.

Sedimentation velocity experiments indicated that the protein is monomeric (at 50 μM total protein) in 1 mM glycine at pH 2.5 with a sedimentation coefficient of 1.69 S (Figure 5). The frictional coefficient ratio f/f_o was 1.06,

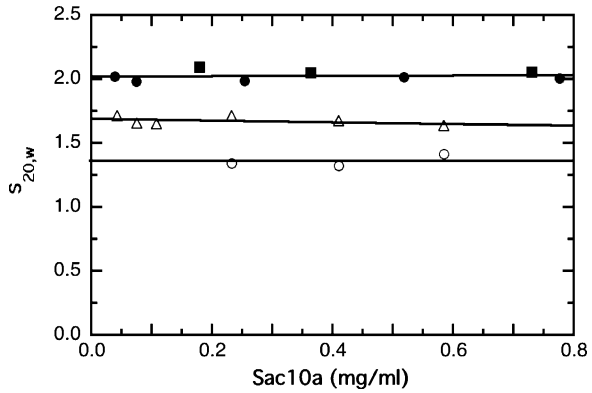

FIGURE 5: Concentration dependence of the sedimentation coefficient, $s_{20,w}$, for dimeric Sac10a at pH 7.0 (0.01 M KH_2PO_4) and 20 °C with 0.0 M NaCl (■) and 0.15 M NaCl (●). The sedimentation coefficient is also shown under conditions that favor the monomer (1 mM glycine, pH 2.5) at 10 °C (△) and 40 °C (○).

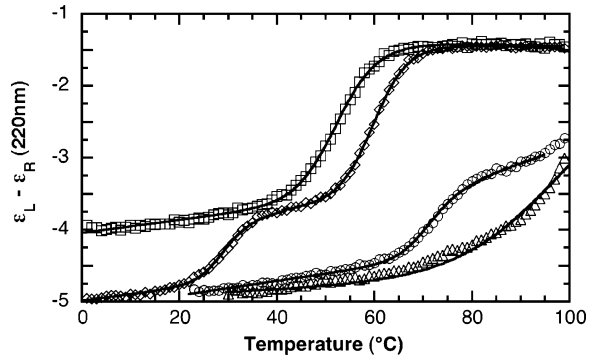

FIGURE 6: Thermal denaturation measured by CD (1.0 cm path length cuvette) of Sac10a (0.01 mg/mL) in 0.01 M KH_2PO_4, 0.15 M NaCl, pH 7.0 (△); 1 mM glycine, pH 5.6 (○); 1 mM glycine, pH 2.5 (◇); 1 mM glycine, pH 2.0 (□). Solid lines show the results of fitting the data to a two-state N ↔ U monomer unfolding model. Systematic deviations are observed in the fit of pH 7 data, as expected for a dimer ↔ monomer unfolding transition. The pH 2.5 data were fit assuming two independent N ↔ U transitions.

indicating that the monomer is globular and not as extended or asymmetric as observed with the dimer. The monomer (in 1 mM glycine, pH 2.5) and dimer (0.15 M NaCl, phosphate buffer, pH 7) had identical near and far UV CD spectra (data not shown), indicating little change in secondary structure upon dissociation, and no detectable change in the local environments of tyrosine and tryptophan side chains.

Hydrophobic Surface Area. Changes in exposure of hydrophobic groups upon dissociation of the dimer was investigated using ANS fluorescence. The fluorescence of ANS is enhanced by association of the dye with hydrophobic surface areas in proteins (*38*). A 14-fold enhancement of ANS fluorescence was observed in the presence of Sac10a upon decreasing the pH to 2.5 (data not shown). The lack of any change in the near UV CD spectrum of Sac10a suggests that the increased binding of ANS is due not to formation of a molten globule but rather to exposure of hydrophobic surface by disruption of the dimer interface.

Protein Stability. Thermal unfolding of Sac10a was monitored using both CD and DSC. At pH 7 in 0.15 M KCl the high thermal stability of the dimeric protein prevented the observation of a complete unfolding transition using CD since the T_m was above 100 °C (Figure 6). The stability of the protein decreased as the pH was lowered in the absence

of salt under conditions more favorable for the monomer, and two transitions became apparent in the temperature dependence of the CD. At pH 5.6 the first transition occurred at 71 °C, with a second occurring above 100 °C. At pH 2.5 the two transitions were well-defined with a T_m of 29.6 °C and a ΔH of 50.3 kcal/mol for the first transition, and a T_m of 59.6 °C and a ΔH of 63.4 kcal/mol for the second. At pH 2.0 the first transition was not observed, while the second occurred at 53 °C. The lowest temperature transition observed in the absence of salt at pH 2.5 and 5.6 was irreversible, which may explain the inability to detect the lower temperature transition at pH 2 (i.e. the transition had occurred during equilibration of the sample). The second transition was reversible under all conditions examined. The solid lines in Figure 6 are fits of the data to simple two-state monomer unfolding reactions, with excellent fitting observed for the pH 2.0 and 2.5 data. The first transition of the pH 5.6 data shows a small deviation from the monomer unfolding model. The unfolding of Sac10a clearly is not fit by a monomer unfolding model at pH 7.0 in 0.15 M NaCl. There was insufficient CD data at pH 7.0 to justify fitting the data to a dimer unfolding model.

The temperature dependence of the secondary structure composition obtained from analysis of CD spectra indicated two distinct thermal transitions involving α-helices at pH 2.5 under conditions favoring the monomer (Figure 7). About 15% of the loss in helix content accompanied the first transition. Only one transition was apparent in the unordered structure component, which increased coincident with the higher temperature transition showing loss of α-helical structure, but the unordered state was not as well-defined by CD data and small transitions may be obscured by noise. Singular value decomposition of CD spectra (18 in all) as a function of temperature demonstrated the existence of only 2 significant (>99%) orthogonal components, representative of folded and unfolded protein. The temperature dependence of the SVD coefficients paralleled the changes seen in the CD intensity (data not shown), and thus there was no spectroscopic evidence for an alternative protein fold induced by temperature.

Scanning calorimetry was used to monitor the unfolding of the more stable dimeric form of Sac10a due to the ability to scan above 100 °C. In 0.15 M NaCl and pH 7 a single endotherm was observed with a T_m of about 108 °C at a protein concentration of 0.41 mg/mL (Figure 8). There was a 20% decrease in the endotherm area following an initial scan to the midpoint temperature, and therefore unfolding was not totally reversible under these conditions. The DSC endotherm was asymmetric, as expected for unfolding of a dimer. The data were fit well by a model involving unfolding of a dimer to two monomeric chains, with a $\Delta H°$ of 239 kcal/mol and a $T°$ of 122 °C. The calorimetric to van't Hoff enthalpy ratio was 0.30, less than the value of 0.5 expected for the 2-state unfolding of a dimer, possibly due to contributions from irreversibility.

DSC of Sac10a under conditions favoring the monomer (pH 2.5 and 1 mM glycine) showed primarily a single major endotherm with a T_m of 62.6 °C and a ΔH of 58.9 kcal/mol (data not shown), consistent with the higher temperature transition observed in CD thermal melts under similar conditions. A second, weak asymmetric transition was observed at lower temperature (25−45 °C) which could not be fit as a simple two-state transition. The intensity of the lower temperature transition was less than 20% of that expected from the van't Hoff enthalpy obtained from CD thermal melting experiments. The higher protein concentrations required for the DSC presumably result in dimer formation or intermolecular interactions between monomers which perturb the lower temperature transition. CD would appear to be the preferred method for studying thermal unfolding of the monomer because of the ability to work at lower (e.g. micromolar) concentrations favoring the monomer.

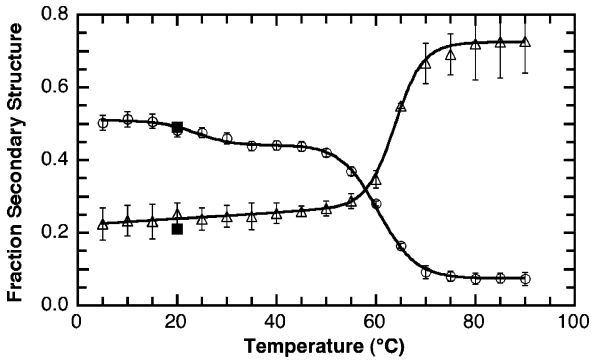

FIGURE 7: Temperature dependence of the fraction of α-helix (○) and unordered (△) secondary structure in Sac10a determined by CD (1 mM glycine, pH 2.5, 0.03 mg/mL protein, 1.0 mm path length). Solid lines show the results of fitting the data to two-state transitions. The α-helix component was fit to two independent transitions. For comparison, the results from an analysis of Sac10a in 0.01 M KH_2PO_4, 0.15 M NaCl, pH 7.0, 20 °C are shown as solid squares.

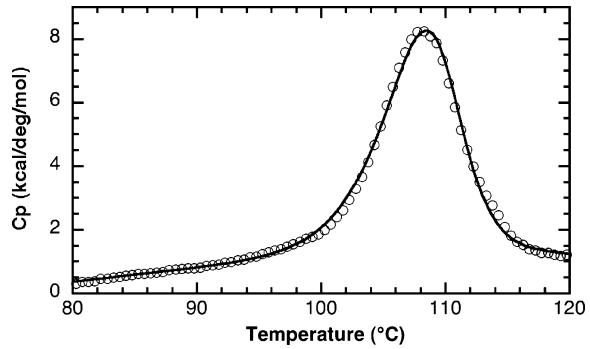

FIGURE 8: DSC of Sac10a at pH 7 (0.41 mg/mL protein, 0.01 M KH_2PO_4, 0.15 M NaCl). For clarity, only every fifth data point is plotted (○). Fitting to a model assuming dimer unfolding to two random coil chains (solid line) gave $\Delta H° = 239$ kcal/mol, $T° = 121.6$ °C, and $\Delta H_{cal}/\Delta H_{vh} = 0.30$.

DNA Binding. The tryptophan fluorescence of Sac10a increased 2-fold in a reverse titration of the protein with increasing concentrations of *E. coli* double-stranded DNA in 0.01 M phosphate buffer at pH 7 (Figure 9A). The DNA binding site size was about 3 base pairs per Sac10a dimer with a dissociation constant of 10^{-6} M (see Table 1). A similar binding site size (3.6 to 4.5) was found for the synthetic polynucleotides poly(dA)-poly(dT), poly(dG)-poly(dC), and poly(dGdC), with slightly greater fluorescence enhancements (about 2.5-fold) and modest increases in affinity to approximately 10^{-7} M. The alternating copolymer poly(dAdT) exhibited the weakest affinity and also had a somewhat larger binding site size of 5.6 base pairs. The fluorescence enhancement was greatest with this DNA giving a 4-fold increase in tryptophan fluorescence upon binding.

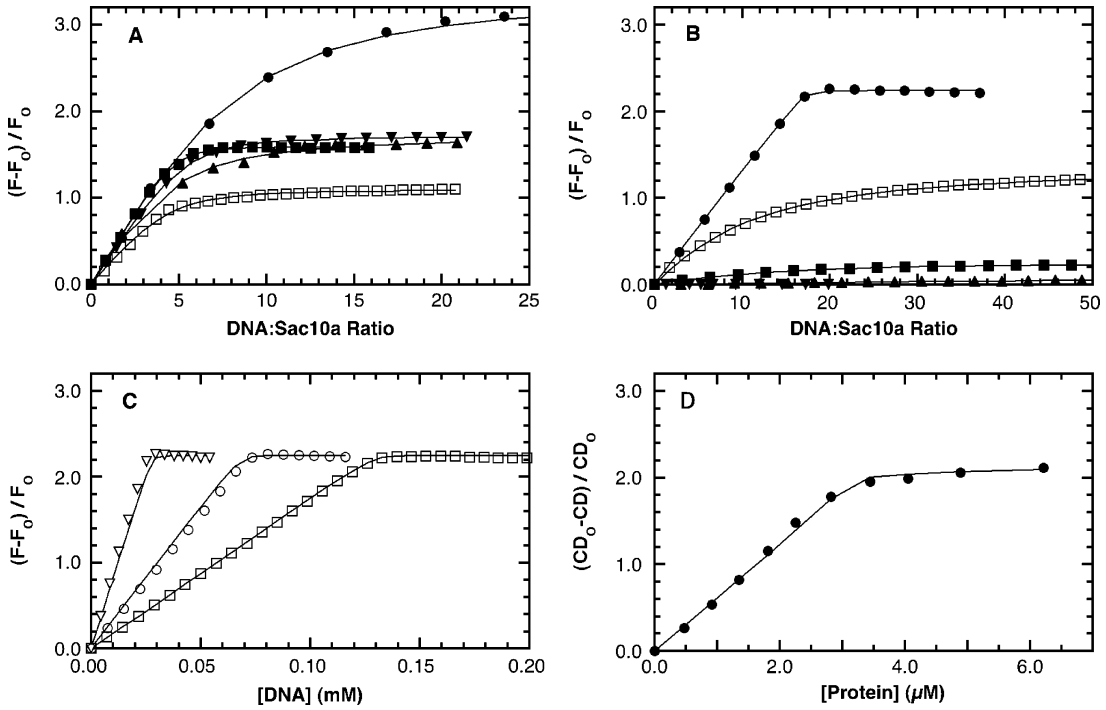

FIGURE 9: Protein/DNA titrations monitored by fluorescence and circular dichroism. (A) Reverse titrations of Sac10a monitored by intrinsic protein fluorescence at pH 7 (0.01 M KH_2PO_4). Sac10a (1.2 μM) was titrated with *E. coli* DNA (□), poly(dA)-poly(dT) (▲), poly(dAdT) (●), poly(dG)-poly(dC) (■), and poly(dGdC) (▼). Data for poly(dAdT) extends-to-a ratio of 40:1 nucleotides/protein. (B) As in A with 0.15 M NaCl. (C) Reverse titrations of Sac10a (1.5 (▽), 3.0 (○), and 7.4 (□) μM) with poly(dAdT). (D) A forward titration of poly(dAdT) (62 μM) by Sac10a followed by CD. All fluorescence data were corrected for dilution and inner filer effects, and CD data were corrected by subtracting the protein contribution at 280 nm. Solid lines are fits of the data to the McGhee−von Hipple model (see Table 1). In C and D the solid lines result from a global fit to all 4 titration curves with $K_D = 9.1 \times 10^{-10}$ and $n = 20.8$ base pairs per dimer.

Table 1: DNA-Binding Parameters of Sac10a[a]

	0 M NaCl			0.15 M NaCl		
DNA	K_D (M)	n (bp)	Q_{max}	K_D (M)	n (bp)	Q_{max}
E. *coli* DNA	1.0×10^{-6} (±0.1)	3.3 (±0.2)	1.0 (±0.1)	1.0×10^{-5} (±0.2)	6.1 (±1.3)	1.4 (±0.02)
poly d(G)-poly d(C)	7.3×10^{-8} (±4.5)	4.1 (±0.2)	1.5 (±0.1)	–[b]	–	0.05
poly d(GC)	1.4×10^{-7} (±1.6)	4.5 (±0.4)	1.5 (±0.3)	–	–	<0.01
poly d(A)-poly d(T)	8.3×10^{-7} (±5.2)	3.6 (±0.7)	1.6 (±0.1)	5.0×10^{-6} (±3.3)	7.2 (±1.4)	0.28 (±0.02)
poly d(AT)	2.5×10^{-6} (±1.7)	5.6 (±0.1)	3.1 (±0.4)	4.8×10^{-10} (±3.5)	16.7 (±1.1)	2.24 (±0.01)

[a] Binding parameters determined by fitting reverse titrations monitored by tryptophan fluorescence with the McGhee−von Hipple lattice model (*34*). The binding site size, n, is the number of base pairs of DNA occluded per Sac10a dimer. Q_{max} is the normalized maximum fluorescence intensity $Q_{max} = (F_{max} - F_o)/F_o$. The mean values and ±1 sd were determined from fitting 2 to 3 independent data sets obtained at different protein concentrations ranging from 1.2 to 7.0 μM. [b] Values which could not be determined due to weak binding or no observed change in fluorescence are indicated with a dash.

Increased salt concentration (0.15 M NaCl) decreased the binding affinity for all of the DNA sequences tested except poly(dAdT) (Figure 9B). The affinity of Sac10a for *E. coli* DNA decreased by an order of magnitude, and the binding site size increased by about 2-fold. A similar affinity and binding site size were observed for poly(dA)-poly(dT), but there was only a 28% increase in tryptophan fluorescence. Smaller fluorescence changes were observed with the two GC containing polynucleotides, and it was not possible to measure binding for these DNA sequences in 0.15 M NaCl.

The affinity of Sac10a for poly(dAdT) increased dramatically with increased salt concentration. Reverse titrations followed by fluorescence were fit well by the noncooperative McGhee−von Hippel model with a dissociation constant of 4.8 (±3.5) $\times 10^{-10}$ M in 0.15 M NaCl at pH 7, i.e. an increase in affinity by over 4 orders of magnitude at higher salt. Inclusion of cooperativity in the model was not required to fit the data. A global fit (Figure 9C,D) of reverse titrations of Sac10a by poly(dAdT) at three different protein concentrations, as well as a forward titration monitored by CD, indicated a dissociation constant of 1×10^{-9} M^{-1} in 0.15 M NaCl. The binding site size obtained for poly(dAdT) in 0.15 M NaCl was 17 base pairs per dimer, significantly larger than observed at lower salt and with other DNA sequences.

Figure 10 shows a forward titration of poly(dAdT)) with Sac10a monitored by CD in 0.15 M NaCl. The spectra are scaled to the molar CD of the DNA component. The increase in the CD at short wavelengths (<240 nm) reflects increasing contributions from protein, while the changes at long wavelengths are dominated by DNA and reflect alterations in the structure of DNA upon Sac10a binding. As seen in the inset of Figure 10, binding of Sac10a led to an inversion of the CD spectrum of poly(dAdT) indicating a significant distortion of the DNA structure consistent with overwinding of the duplex. Less drastic spectral changes were observed with *E. coli* DNA in 0.15 M NaCl (data not shown), with a

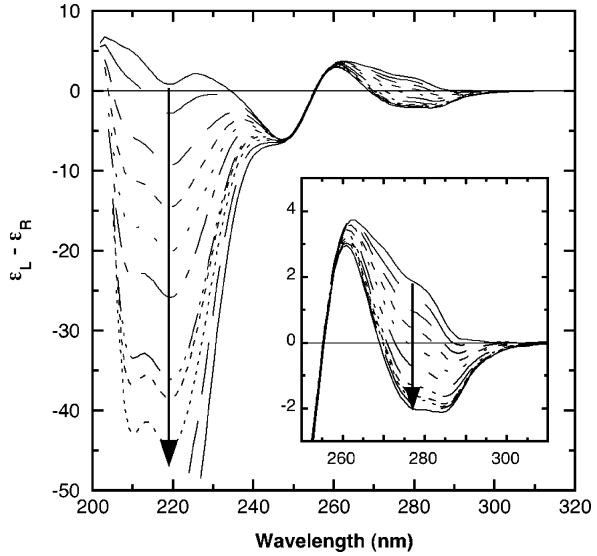

FIGURE 10: Perturbation of the CD spectrum of poly(dAdT) with increasing concentrations of Sac10a. Spectra are scaled to the molar concentration of DNA nucleotides. The initial DNA concentration was 62 μM nucleotides. The Sac10a:DNA ratios were 0, 0.007, 0.015, 0.022, 0.030, 0.037, 0.047, 0.057, 0.068, 0.082, 0.11 (monomer/nucleotide). The arrow indicates the effect of increasing protein concentration. Inset shows the long wavelength region.

decrease in the intensity of the positive long wavelength band by a factor of 2. In the absence of NaCl, only slight decreases were observed in the long wavelength positive CD band of DNA.

DISCUSSION

Sac10a is one of the more plentiful small, basic DNA-binding proteins expressed in *S. acidocaldarius* that have been classified as chromatin proteins (*13*). It shows no sequence homology to the other members of the group (e.g. Sac7d, Alba, Sso10b2). The tight binding observed for poly(dAdT) at physiological salt concentrations is remarkable and is reminiscent of TBP from *Pyrococcus woesei* (*39*). Sac10a is the only member of the group of small DNA-binding proteins from *Sulfolobus* which have demonstrated a pronounced DNA sequence specificity. Whether or not the preferred sequence is alternating AT remains to be determined. Changes in the CD spectrum of poly(dAdT) indicate a significant decrease and inversion of the CD spectra in the near UV region. This is opposite to that observed for Sac7d, which has been shown to kink and unwind DNA (*5*). The CD data for Sac10a are therefore consistent with overwinding of poly(dAdT), which supports the earlier electron micrograph images of Sac10a−DNA complexes which indicated "supertwisting" of circular DNA (*18*).

Structure prediction based on the sequence of Sac10a using 3D-pssm (*40*) indicates that the N-terminal two-thirds of the sequence adopts a classic winged helix fold, a common DNA-binding motif in bacterial and eukaryotic transcription factors (*41*−*43*) as well as linker histones (*44*, *45*). It is intriguing to note that the winged helix motif has been argued to be the most common HTH fold in Archaea (*46*). Structure prediction reveals potential similarities to MotA (1bja, the bacterial phage T4 transcription factor (*47*)) and ModE (1b9m, an *E. coli* molybdate dependent transcription regulator (*48*)). ModE and MotA are especially interesting in that not only are they dimers with winged helix DNA-binding domains but the dimerization motif is an antiparallel coiled coil, which positions the DNA-binding domains at opposite ends of the coiled coil with a separation of approximately 45 Å. Such an arrangement in Sac10a would be consistent with the asymmetry required by the hydrodynamic measurements as well as the 17 bp DNA-binding site size. This may also explain the interesting observation that some crystalline forms of the homologous Sso10a from *S. solfataricus* gave a fiberlike X-ray diffraction pattern (*19*). More detailed NMR and X-ray crystal structural studies of Sso10a will be presented elsewhere, but we note that the structure of Sso10a contains a dimer of winged helices separated by an antiparallel coiled coil rod (*49*).

Sac10a contains a single tryptophan (W33) which is responsible for the observed increase in fluorescence intensity with DNA binding. Modeling and comparison to Sso10a structural data indicate that W33 in Sac10a is most likely solvent exposed and located on the surface of the first helix of the helix−turn−helix motif of the winged helix. The increase in fluorescence is therefore likely to be associated with changes in solvent exposure upon binding to DNA rather than to changes in protein structure. It is interesting that Sac7d also contains a single, solvent exposed tryptophan which becomes buried at the protein−DNA interface, and yet its fluorescence is quenched by binding (*8*, *9*).

The dimer of Sac10a can be dissociated into monomers with decreasing pH in the absence of salt, presumably by dissociation of the coiled coil. It is interesting that the disassociated monomers under these conditions appear to be folded to the same extent as the dimer due to the lack of significant changes in both the near and far UV CD spectra. However, dimer dissociation does result in increased exposure of hydrophobic groups as well as a rearrangement of the protein tertiary structure, as revealed by ANS binding and sedimentation, respectively. If dimerization occurs through formation of coiled coil, there must be a rearrangement of the protein tertiary structure upon dissociation of the anisotropic dimer into more globular monomers. Alternative packing of the exposed extended helix may occur either by folding of the helix back onto itself or by interactions with the HTH domain.

Since Sac10a at a concentration of 1 μM and pH 2.5 remains monomeric from 5° to 40 °C, we conclude that neither of the melting transitions observed in the far UV CD correspond to dimer dissociation. Rather it appears that both transitions represent cooperative unfolding of independent domains in Sac10a, possibly the winged helix DNA-binding domain and the coiled coil. That the first transition observed by CD is due to partial unfolding is supported by an increase in the dimensions of the Sac10a monomer as indicated by a decrease in $s_{20,w}$ from 1.69 to 1.37 S (see Figure 5). A better understanding of the thermal melting transitions will require determination of the structures of both the monomer and dimer.

The Sac10a protein sequence is similar to a number of hypothetical protein sequences coded in archaeal genomes. Most notable is the similarity with the conserved domain COG3432 (*50*) (Figure 1), which is currently composed of 16 hypothetical proteins coded by sequences in the NCBI Conserved Domain Database from both crenarchaeota and euryarchaeota genomes (Table 2). The sequence similarities

Table 2: Comparison of Sac10a to Proteins Containing the COG3432 Conserved Domain

organism	protein[a]	no. of residues	p*I*[b]	sequence similarity to Sac10a[c]	coiled coil length[d]
Crenarchaeota					
Sulfolobus acidocaldarius	Sac10a	101	9.21	–	33 (D68–E100)
	Sac10a2	89	9.33	36/60	14 (E75–K88)[e]
Sulfolobus solfataricus	**SSO10449 (Sso10a)**	95	9.71	41/65	36 (M59–I94)
	SSO2827	130	9.15	27/51	28 (L100–T127)
	SSO2986	100	8.89	34/56	28 (E56–I83)
Sulfolobus tokadaii	ST0658	101	9.06	82/92	36 (L66–E101)
	ST1562	121	9.38	36/63	38 (D79–G116)
	STS113	94	7.88	34/60	29 (L66–L94)
Aeropyrum pernix	**APES025**	77	9.96	nd	nd
Pyrobaculum aerophilum	**PAE3660**	85	9.47	36/54	nd
Euryarchaeota					
Archaeoglobus fulgidus	**AF1459**	90	9.57	34/54	nd
	AF1663	96	7.73	30/56	31 (R65–I95)
	AF2083	83	7.91	32/59	nd
Methanosarcina acetivorans	**MA1811**	106	8.78	45/65	nd
Methanococcus jannaschii	**MJ0287**	97	9.84	36/56	nd
	MJ0290	212	6.06 (8.88)	42/59	nd
Thermoplasma acidophilum	**Ta0385**	99	9.52	nd	nd
	Ta0386	229	6.85 (8.12)	nd	28 (L60–I87)
Thermoplasma volcanium	**TVN0317**	134	9.36	34/51	nd
	TVN1186	232	9.21	27/54	28 (E132–L159)
	TVN1188	107	9.82	nd	nd

[a] Proteins currently listed in the COG3432 conserved domain database (//www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=COG3432&version=) are in bold type. [b] p*I* values are calculated from the amino acid composition. The p*I* values for MJ0290 and Ta0386 in parentheses are for the truncated sequences shown in Figure 1. [c] Sequence similarity is indicated by the percent identity followed by percent conservative substitutions obtained from a BLAST alignment of each sequence with Sac10a. "nd" indicates that no significant similarity was detected. [d] Coiled coil length predicted by COILS is indicated by the number of amino acids followed by the first and last residues in parentheses. "nd" indicates that no significant coiled coil was predicted. [e] Predicted with a window of 14 residues.

shown in Figure 1 indicate that both Sac10a and Sac10a2, as well as three *Sulfolobus tokodaii* proteins (STO658, ST1562, STS113), can be classified as additional members of the COG3432 group. COG3432 is referred to as a predicted transcriptional regulator, but until now none of the proteins have been characterized. The majority of the proteins are basic with p*I* values around 9. The two exceptions are MJ0290 and Ta0396, which are approximately twice as large as the other proteins, but these are also basic in the COG3432 domain.

Sequence alignments of Sac10a with COG3432 proteins show four regions of significant similarity indicating those portions of the sequence that are most likely important in protein function and structure (Figure 1). The three regions of highest similarity correspond to segments K11–C25, T32–Q56, and Y65–G71 in the Sac10a sequence. The level of similarity in the first is remarkable, with the cluster of positive charge suggesting that this region may be important in DNA binding. The 100% sequence identity observed at T68 and G71 at the beginning of the predicted coiled coil helix is also noteworthy. In addition to the similarities, there are notable sequences of variable length at both termini, as well as in the vicinity of the second proposed β-strand, suggesting a loop of variable length. The latter is the expected locus of the β-hairpin "wing" based on sequence comparisons to the winged helix proteins MotA (1bja) and ModE (1b9m). The sequence similarities in the C-terminus indicates that the proposed coiled coil domain in Sac10a is conserved throughout the group, although the length of the coiled coil is variable. The C-terminal sequences of some of the proteins with shorter C-terminal helices showed little or no tendency to form coiled coil using the program COILS (Table 2), although the sequence alignments clearly imply the existence of short coiled coil regions in these proteins (Figure 1). The length of the predicted coiled coil in Sac10a2 is nearly half that in Sac10a. If the structures indeed contain an antiparallel coiled coil as discussed above in analogy to ModE and MotA, the different lengths would result in different separations of the DNA-binding domains. The natural twist in the coiled coil will also lead to different relative orientations of the DNA-binding sites.

When Sac10a was initially characterized along with other small, basic DNA-binding proteins from *Sulfolobus*, it was considered a likely chromatin protein that was important in DNA packaging as evidenced by its ability to fully coat and alter the structure of closed circular DNA (*13*, *18*). The high expression level of Sac10a would support this. The demonstration that the protein is a dimer of subunits which are homologous to a predicted Archaea transcription regulator domain (COG3432) indicates that it may also play a more specific regulatory role. The similarity of the predicted structural motifs of Sac10a to the MerR family of bacterial transcription activators is especially intriguing (*51*). However, the winged helix motif is important not only in transcription factors, which control the expression of specific genes, but also in H1 linker histones, which are important in nucleosome compaction in eukaryotes (*41*, *52*–*54*). These are a comparatively diverse group of proteins which do not adopt the classic histone fold, but contain an N-terminal winged helix domain followed by a lysine rich C-terminal helical domain. We note that eight of the 33 residues in the C-terminal helix of Sac10a are lysine, and the computed p*I* is 9.06. Although the Archaea contain true histones homologous to eukaryotic histones H2A, H2B, H3, and H4, to our knowledge no H1

histone homologue has been described in the Archaea (*55*, *56*).

ACKNOWLEDGMENT

The authors would like to acknowledge the generous assistance of Dr. Roger Garrett (University of Copenhagen) in searching the *S. acidocaldarius* genomic database.

REFERENCES

1. Grote, M., Dijk, J., and Reinhardt, R. (1986) Ribosomal and DNA binding proteins of the thermoacidophilic archaebacterium Sulfolobus acidocaldarius, *Biochim. Biophys. Acta 873*, 405−13.
2. White, M. F., and Bell, S. D. (2002) Holding it together: chromatin in the Archaea, *Trends Genet. 18*, 621−6.
3. Edmondson, S. P., Qiu, L., and Shriver, J. W. (1995) Solution Structure of the DNA-Binding Protein Sac7d from the Hyperthermophile *Sulfolobus acidocaldarius*, *Biochemistry 34*, 13289−304.
4. McCrary, B. S., Edmondson, S. P., and Shriver, J. W. (1996) Hyperthermophile Protein Folding Thermodynamics: Differential Scanning Calorimetry and Chemical Denaturation of Sac7d, *J. Mol. Biol. 264*, 784−805.
5. 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*, 202−5.
6. Krueger, J. K., McCrary, B. S., Wang, A. H., Shriver, J. W., Trewhella, J., and Edmondson, S. P. (1999) The solution structure of the Sac7d/DNA complex: a small-angle X-ray scattering study, *Biochemistry 38*, 10247−55.
7. Edmondson, S. P., and Shriver, J. W. (2001) DNA binding proteins Sac7d and Sso7d from Sulfolobus, *Methods Enzymol.334*, 129−45.
8. McAfee, J., Edmondson, S., Datta, P., Shriver, J., and Gupta, R. (1995) Gene cloning, sequencing, expression, and characterization of the Sac7 DNA-binding proteins from the extremely thermophilic archaeon Sulfolobus acidocaldarius, *Biochemistry 34*, 10063−77.
9. McAfee, J. G., Edmondson, S., Zegar, I., and Shriver, J. W. (1996) Equilibrium DNA Binding of Sac7d Protein from the Hyperthermophile Sulfolobus acidocaldarius: Fluorescence and Circular Dichroism Studies, *Biochemistry 35*, 4034−45.
10. McCrary, B. S., Bedell, J., Edmondson, S. P., and Shriver, J. W. (1998) Linkage of Protonation and Anion Binding to the Folding of Sac7d, *J. Mol. Biol. 276*, 203−24.
11. Lundback, T., Hansson, H., Knapp, S., Ladenstein, R., and Hard, T. (1998) Thermodynamic characterization of non-sequence-specific DNA-binding by the Sso7d protein from Sulfolobus solfataricus, *J. Mol. Biol. 276*, 775−86.
12. Shriver, J. W., Peters, W. B., Szary, N., Clark, A. T., and Edmondson, S. P. (2001) Calorimetric analyses of hyperthermophile proteins, *Methods Enzymol. 334*, 389−422.
13. Dijk, J., and Reinhardt, R. (1986) in *Bacterial Chromatin* (Gualerzi, C., and Pon, C., Eds.) pp 185−218, Springer-Verlag, Berlin.
14. 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*, 3929−33.
15. 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*, 148−51.
16. Wardleworth, B. N., Russell, R. J., Bell, S. D., Taylor, G. L., and White, M. F. (2002) Structure of Alba: an archaeal chromatin protein modulated by acetylation, *EMBO J. 21*, 4654−62.
17. Chou, C. C., Lin, T. W., Chen, C. Y., and Wang, A. H. (2003) Crystal structure of the hyperthermophilic archaeal DNA-binding protein Sso10b2 at a resolution of 1.85 Angstroms, *J. Bacteriol. 185*, 4066−73.
18. 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*, 3715−21.
19. Teale, M., Kahsai, M., Singh, S. K., Edmondson, S. P., Gupta, R., Shriver, J. W., and Meehan, E. (2003) Cloning, expression, crystallization, and preliminary X-ray analysis of the DNA-binding protein Sso10a from Sulfolobus solfataricus, *Acta Crystallogr., Sect. D: Biol. Crystallogr. D59*, 1320−2.
20. Zillig, W. (1993) Confusion in the assignments of Sulfolobus sequences to Sulfolobus species, *Nucleic Acids Res. 21*, 5273.
21. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res, 25*, 4876−82.
22. Baldi, P., Brunak, S., Frasconi, P., Soda, G., and Pollastri, G. (1999) Exploiting the past and the future in protein secondary structure prediction, *Bioinformatics 15*, 937−46.
23. Lupas, A. (1996) Prediction and analysis of coiled-coil structures, *Methods Enzymol. 266*, 513−25.
24. Bornberg-Bauer, E., Rivals, E., and Vingron, M. (1998) Computational approaches to identify leucine zippers, *Nucleic Acids Res. 26*, 2740−6.
25. Johnson, B. B., Dahl, K. S., I. Tinoco, J., Ivanov, V. I., and Zhurkin, V. B. (1981) Correlations between deoxynucleic acid structural parameters and calculated circular dichroism spectra, *Biochemistry 20*, 73−8.
26. van Holde, K. E., and Weischet, W. O. (1978) Boundary Analysis of Sedimentation Velocity Experiments with Monodisperse and Paucidisperse Solutes, *Biopolymers 17*, 1387−1403.
27. Philo, J. (1997) An Improved Function for Fitting Sedimentation Velocity Data for Low-Molecular-Weight Solutes, *Biophys. J. 72*, 435−44.
28. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from CD spectra: Comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set, *Anal. Biochem. 282*, 252−60.
29. Shriver, J., Peters, W., Szary, N., Clark, A., and Edmondson, S. (2001) Calorimetry of Hyperthermophile Proteins, *Metheds. Enzymol. 334*, 389−422.
30. Li, W.-T., Grayling, R., Sandman, K., Edmondson, S., Shriver, J. W., and Reeve, J. N. (1998) Thermodynamic Stability of Archaeal Histones, *Biochemistry 37*, 10563−72.
31. Myers, J., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding, *Protein Sci. 4*, 2138−48.
32. Bedell, J. L., McCrary, B. S., Edmondson, S. P., and Shriver, J. W. (2000) The acid-induced folded state of Sac7d is the native state, *Protein Sci. 9*, 1878−88.
33. Lohman, T., and Mascotti, D. P. (1992) Nonspecific ligand-DNA equilibrium binding parameters determined by fluorescence methods, *Methods. Enzymol. 212*, 424−58.
34. McGhee, J., and von Hippel, P. (1974) Theoretical aspects of DNA-protein interactions: Co-operative and non-cooperative binding of large ligands to a one-dimensional homogeneous lattice, *J. Mol. Biol. 86*, 469−89.
35. Gill, S., and von Hippel, P. (1989) Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data, *Anal. Biochem. 182*, 319−26.
36. Lu, S. M., and Hodges, R. S. (2004) Defining the minimum size of a hydrophobic cluster in two-stranded alpha-helical coiled-coils: effects on protein stability, *Protein Sci. 13*, 714−26.
37. Waxman, E., Laws, W. R., Laue, T. M., Nemerson, Y., and Ross, J. B. (1993) Human factor VIIa and its complex with soluble tissue factor: evaluation of asymmetry and conformational dynamics by ultracentrifugation and fluorescence anisotropy decay methods, *Biochemistry 32*, 3005−12.
38. Stryer, L. (1965) The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe for nonpolar sites, *J. Mol. Biol. 13*, 482−95.
39. O'Brien, R., DeDecker, B., Fleming, K. G., Sigler, P. B., and Ladbury, J. E. (1998) The effects of salt on the TATA binding protein-DNA interaction from a hyperthermophilic archaeon, *J. Mol. Biol. 279*, 117−25.
40. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM, *J. Mol. Biol. 299*, 499−520.
41. Brennan, R. G. (1993) The winged-helix DNA-binding motif: another helix-turn-helix takeoff, *Cell 74*, 773−6.
42. Kenney, L. J. (2002) Structure/function relationships in OmpR and other winged-helix transcription factors, *Curr. Opin. Microbiol. 5*, 135−41.
43. Latchman, D. S. (2004) *Eukaryotic Transcription Factors*, 4th ed., Elsevier, New York.

44. Gajiwala, K. S., Chen, H., Cornille, F., Roques, B. P., Reith, W., Mach, B., and Burley, S. K. (2000) Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding, *Nature 403*, 916−21.
45. Ono, K., Kusano, O., Shimotakahara, S., Shimizu, M., Yamazaki, T., and Shindo, H. (2003) The linker histone homolog Hho1p from Saccharomyces cerevisiae represents a winged helix-turn-helix fold as determined by NMR spectroscopy, *Nucleic Acids Res. 31*, 7199−207.
46. Aravind, L., and Koonin, E. V. (1999) DNA-binding proteins and evolution of transcription regulation in the archaea, *Nucleic Acids Res. 27*, 4658−70.
47. Finnin, M. S., Cicero, M. P., Davies, C., Porter, S. J., White, S. W., and Kreuzer, K. N. (1997) The activation domain of the MotA transcription factor from bacteriophage T4, *EMBO J. 16*, 1992−2003.
48. Hall, D. R., Gourley, D. G., Leonard, G. A., Duke, E. M., Anderson, L. A., Boxer, D. H., and Hunter, W. N. (1999) The high-resolution crystal structure of the molybdate-dependent transcriptional regulator (ModE) from Escherichia coli: a novel combination of domain folds, *EMBO J. 18*, 1435−46.
49. Chen, L., Chen, L.-R., Zhou, X., Yang, Y., Kahsai, M., Clark, A., Edmondson, S., Liu, Z.-J., Rose, J. P., Wang, B. C., Meehan, E., and Shriver, J. W. (2004) The Hyperthermophile Protein Sso10a is a Dimer of Winged Helix DNA-Binding Domains Linked by an Antiparallel Coiled Coil Rod, *J. Mol. Biol. 341*, 73−91.
50. Marchler-Bauer, A., Anderson, J. B., DeWeese-Scott, C., Fedorova, N. D., Geer, L. Y., He, S., Hurwitz, D. I., Jackson, J. D., Jacobs, A. R., Lanczycki, C. J., Liebert, C. A., Liu, C., Madej, T., Marchler, G. H., Mazumder, R., Nikolskaya, A. N., Panchenko, A. R., Rao, B. S., Shoemaker, B. A., Simonyan, V., Song, J. S., Thiessen, P. A., Vasudevan, S., Wang, Y., Yamashita, R. A., Yin, J. J., and Bryant, S. H. (2003) CDD: a curated Entrez database of conserved domain alignments, *Nucleic Acids Res. 31*, 383−7.
51. Brown, N. L., Stoyanov, J. V., Kidd, S. P., and Hobman, J. L. (2003) The MerR family of transcriptional regulators, *FEMS Microbiol. Rev. 27*, 145−63.
52. Wolffe, A. P. (1997) Histone H1, *Int. J. Biochem. Cell Biol. 29*, 1463−6.
53. Horn, J. R., Brandts, J. F., and Murphy, K. P. (2002) van't Hoff and Calorimetric Enthalpies II: Effects of Linked Equilibria, *Biochemistry 41*, 7501−7.
54. Vignali, M., and Workman, J. L. (1998) Location and function of linker histones, *Nat. Struct. Biol. 5*, 1025−8.
55. Reeve, J. N. (2003) Archaeal chromatin and transcription, *Mol. Microbiol. 48*, 587−98.
56. Kasinsky, H. E., Lewis, J. D., Dacks, J. B., and Ausio, J. (2001) Origin of H1 linker histones, *FASEB J. 15*, 34−42.

BI0491752