Role of a Surface Tryptophan in Defining the Structure, Stability, and DNA Binding of the Hyperthermophile Protein Sac7d[†,‡]

Jennifer L. Bedell, Stephen P. Edmondson,* and John W. Shriver*

Laboratory for Structural Biology, Departments of Chemistry and Biological Sciences, Graduate Program in Biotechnology and Bioengineering, University of Alabama, Huntsville, Alabama 35899

Received October 8, 2004; Revised Manuscript Received November 4, 2004

ABSTRACT: Sac7d is a small, chromatin protein from *Sulfolobus acidocaldarius* which induces a sharp kink in DNA with intercalation of valine and methionine side chains. The crystal structure of the protein−DNA complex indicates that a surface tryptophan (W24) plays a key role in DNA binding by hydrogen bonding to the DNA at the kink site. We show here that substitution of the solvent-exposed tryptophan with alanine (W24A) led to a significant loss in not only DNA binding affinity but also protein stability. The W24A substitution proved to be one of the most destabilizing surface substitutions in Sac7d. A global linkage analysis of the pH and salt dependence of stability indicated that the protein stability surface (ΔG vs temperature, pH, and salt concentration) was lowered overall by 2 kcal/mol (from 0 to 100 °C, pH 0 to 7, and 0 to 0.3 M KCl). The lower free energy of unfolding could not be attributed to significant structural perturbations of surface electrostatic interactions. Residual dipolar coupling of partially aligned protein and the NMR solution structure of W24A confirmed that the surface substitution resulted in no significant change in structure. Stabilization of this hyperthermophile protein and its DNA complex by a surface cluster of hydrophobic residues involving W24 and the two intercalating side chains is discussed.

Sac7d[1] is a member of the Sul7 family of basic 7 kDa chromatin proteins found in the crenarchaeon *Sulfolobus*. In addition to Sac7d from *Sulfolobus acidocaldarius*, examples include Sso7d from *Sulfolobus solfataricus*, Ssh7 from *Sulfolobus shibatae*, and Sto7 from *Sulfolobus tokodaii* (*1*−*5*). While the function of these proteins has not been described, they appear to be unique to *Sulfolobus*, a thermoacidophile which lacks the histones that have been characterized in a number of other archaea (*6*). The Sul7 proteins are basic, monomeric proteins which are present in significant amounts *in vivo* and bind to DNA non-sequence-specifically with moderate affinity (*2*). It is commonly believed that they play an architectural role in packaging and stabilizing genomic DNA at the high growth temperatures of these hyperthermophiles (viz. 75−85 °C).

Sac7d adopts an OB fold and may represent an ancient ancestor of the chromo domain found in many chromosomal proteins (*7*). The protein is highly basic with an especially high charge density (30 charged groups in 66 residues). DNA binding occurs by interaction of a three stranded β-sheet with the minor groove of DNA (*8*). A hydrophobic cluster on the surface of the sheet is composed of W24, V26, and M29. The cluster is conserved in all Sul7 proteins. W24 is the only tryptophan in Sac7d, and the intrinsic fluorescence is quenched by ∼90% upon DNA binding, indicating that it is intimately involved with DNA binding or associated structural changes. Crystal structures of DNA complexes with Sac7d and Sso7d indicate that these proteins significantly bend DNA by inducing a sharp kink primarily at a singe base pair step with intercalation of V26 and M29 side chains (*8*, *9*). In the seven crystal structures of Sul7−DNA complexes determined to date (Protein Data Bank entries 1azp, 1azq, 1ca5, 1ca6, 1c8c, 1bnz, and 1bf4), a hydrogen bond between W24 and the DNA appears to enforce the sharp kink by anchoring the base pair adjacent to the kink site (Figure 1).

Comparison of the NMR solution structure of free Sac7d (*10*) with that of the crystal structure of the Sac7d−DNA complex (*8*) demonstrates that there is little change in the protein upon DNA binding. Most interestingly, the position of the W24 side chain in the free and DNA-bound structures is nearly identical. The χ_1 angle for W24 in the NMR structure is 70°, compared to 64.5° in the crystal structure of the DNA complex [with C^{γ} trans to H^{α}, this is the **p** rotomer in the Richardson nomenclature (*11*), g^- in that of Janin et al. (*12*), and g^+ in the nomenclature of Markley et al. (*13*)]. This is not the most preferred χ_1 rotomer for tryptophan, but it places the side chain in an ideal position for hydrogen bonding to DNA. By "pre-positioning" the side chain in this orientation prior to binding, an unfavorable entropic contribution to binding is prevented.

As part of a study of the interactions at the Sac7d−DNA interface which are responsible for bending DNA, we have

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

[‡] Coordinates for the W24A mutant of Sac7d have been submitted to the Protein Data Bank (entry 1XX8).

* To whom correspondence should be addressed: Department of Chemistry, Materials Science Building, University of Alabama, Huntsville, AL 35899. Phone: (256) 824-2477. Fax: (256) 824-6349. E-mail: shriverj@uah.edu, edmonds@uah.edu.

[1] Abbreviations: DSC, differential scanning calorimetry; D_a, axial component of the residual dipolar coupling molecular alignment tensor; *R*, rhombic component of the residual dipolar coupling molecular alignment tensor; RDC, residual dipolar coupling; Sac7d, Sul7 protein isolated from *Sulfolobus acidocaldarius*; Sul7, family of 7 kDa DNA binding proteins in *Sulfolobus*; W24A, tryptophan 24 to alanine site-directed mutant; W24F, tryptophan 24 to phenylalanine site-directed mutant; W24T, tryptophan 24 to threonine site-directed mutant.

10.1021/bi047823b CCC: $30.25 © 2005 American Chemical Society
Published on Web 12/31/2004

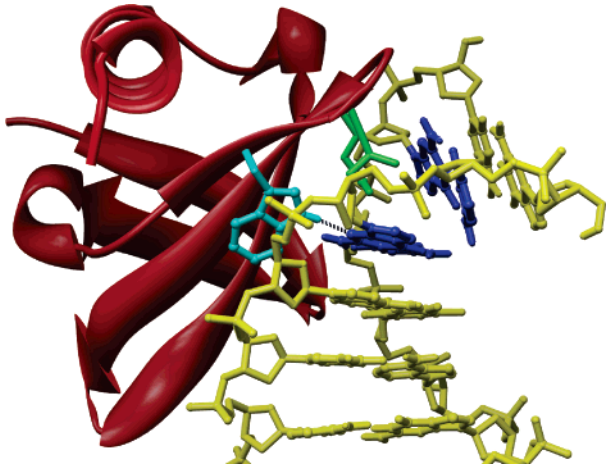

FIGURE 1: Structure of the Sac7d−GCGATCGC complex (PDB entry 1azp) showing W24 (cyan) and the intercalating V26 and M29 (green) side chains at the kink site. The C2-G3 step with a roll angle of 66° [calculated with FREEHELIX (*59*)] is colored blue. Hydrogen bonding between the indole NH group of W24 and N3 of G3 is indicated with the black dashed line.

performed an alanine scan of the protein residues thought to be important in DNA binding. Truncation of the intercalating V26 and M29 led to a small decrease in the T_m of the protein which was consistent with a loss of the hydrophobic interactions with W24 (*14*). Removal of the intercalating side chains led to a small decrease in DNA binding affinity and a decreased ability to bend DNA. We extend these studies here with alanine substitution of W24 and show that the substitution has only a minor effect on the structure of the protein. Surprisingly, we find that the surface tryptophan not only is important in defining DNA binding affinity but also plays an important role in defining the stability of the protein.

MATERIALS AND METHODS

Proteins. The Sac7d gene in the pET-3b expression vector was expressed in *Escherichia coli* BL21(DE3)pLys, and the recombinant protein was purified as previously described (*15*). Site-specific mutagenesis of tryptophan 24 was performed using the Clontech Transformer Site-Directed Mutagenesis Kit. Recombinant proteins were expressed and purified in a manner similar to that for Sac7d, and the identity of the protein was confirmed by changes in UV absorbance, fluorescence, and NMR sequential assignments (see below). ^{15}N-labeled proteins were made by expressing the protein in minimal media supplemented with $^{15}NH_4Cl$.

UV Absorbance and Fluorescence. Protein concentrations were determined from UV absorption spectra. The UV extinction coefficient for Sac7d was 1.1 M^{-1} cm^{-1} at 280 nm (*15*), and that for the W24A, W24T, and W24F proteins was determined to be 0.34 M^{-1} cm^{-1} based on the amino acid content of the protein (*16*, *17*). DNA concentrations were determined using an extinction coefficient of 8400 M^{-1} cm^{-1} at 260 nm for poly(dGdC) (*18*). Wavelength and absorbance accuracies of the spectrophotometer were checked using potassium chromate (*19*). Fluorescence data were collected on an SLM 8000C spectrofluorimeter as previously described (*20*).

DNA Binding Assay. DNA binding of recombinant proteins lacking W24 was monitored indirectly in a competitive assay in which the binding of native Sac7d was assessed in the presence of varying concentrations of nonfluorescent recombinant protein. Reverse titrations of Sac7d/recombinant protein mixtures were performed with increasing concentrations of poly(dGdC). The extent of Sac7d binding was indicated by fluorescence quenching with excitation at 295 nm and emission at 350 nm. Three to five reverse titrations were conducted with the same concentration of Sac7d (typically 2×10^{-6} M). Data were fit using the McGhee−von Hippel model (*21*) assuming noncooperative binding (*20*) to obtain the binding constant, site size, and maximal quenching. Excellent fitting of the data was obtained by assuming identical binding site sizes for Sac7d and competitors. No attempt was made to fit the data with different site sizes for the two proteins.

Circular Dichroism. CD spectra were measured on an AVIV 62DS spectropolarimeter as described elsewhere (*15*). All wavelength scans were collected at 20 ± 0.2 °C, except when using 0.01 cm cuvettes when the sample was maintained at room temperature. Spectra were collected from 330 to 180 nm (or until the dynode voltage reached 500 V) in 1 nm intervals with a 10 s/nm averaging time. Water baselines were subtracted from all protein spectra. CD wavelength spectra were smoothed as described by Savitsky and Golay (*22*). The CD was calibrated using *d*-camphor-10-sulfonic acid. The molar CD per peptide bond was determined using standard procedures (*23*) along with the UV extinction coefficient reported above.

All thermal unfolding experiments followed by CD were monitored at 205 nm at 1 °C temperature increments, using a 30 s averaging time after a 1.5 min equilibration time. The temperature was maintained within ±0.2 °C by a computer-controlled cuvette jacket.

Differential Scanning Calorimetry. Differential scanning calorimetry was performed on an N-DSC II calorimeter (Calorimetry Sciences Corp., Provo, UT) as previously described (*24*) with a scan rate of 1 °C/min. Excess heat capacity data were analyzed using in-house software to determine midpoint temperatures as well as calorimetric and van't Hoff enthalpies (*24*).

Protein Stability Data Analysis. Circular dichroism spectral intensities as a function of temperature, pH, and salt concentration were fit globally by nonlinear regression to a model in which protein folding was linked to two chloride binding and two carboxyl protonation reactions as described previously (*25*). Fitting of simulated data indicated that this minimal model (although oversimplified) was capable of accurately defining (fitting) the unfolding free energy surface to within approximately 0.1 kcal/mol (*25*).

A model which links two ionization and two ligand binding reactions to protein folding contains at least 19 independent reactions (*25*), each of which is characterized by three independent thermodynamic parameters: an equilibrium constant [or equivalently a temperature at which the equilibrium constant equals 1 (i.e., a T_m)], ΔH, and ΔC_p. As described for Sac7d (*25*), the number of parameters was reduced to 11: the ΔC_p of the intrinsic unfolding reaction, pK_u for the carboxyls in the unfolded protein which are affected by folding at low and high salt concentrations, pK_n for the same carboxyls in the native folded protein at low and high salt concentrations, pK_1 for these carboxyls in the folded protein with bound chloride at low and high salt concentrations, two chloride binding constants at low salt concentrations (K_{Cl} and K'_{Cl}), a single chloride binding

constant at high salt concentrations (K_{Cl}), and the enthalpy and heat capacity of chloride binding.

The following data sets were fit simultaneously: (1) a series of eight CD thermal melts (100 data points each) as a function of pH in the absence of added salt ("low salt"), (2) a series of seven CD thermal melts (100 data points each) as a function of pH in 0.3 M KCl ("high salt"), (3) variation of the extent of folding from pH 0 to 8 obtained from CD at 20 °C at low salt (26 data points) and 0.3 M KCl (38 data points), and (4) variation of the extent of folding with salt concentration followed by CD (31 data points) at 25 °C and pH 2. Chloride concentrations were converted to activities using the parametric equations of Pitzer (*26*, *27*). The sum of the squares of the residuals for each data set was weighted by the root-mean-square noise level of the respective data set. Parameter optimization was accomplished by iteratively applying adaptive grid refinement global optimization (Loehle Enterprises, Naperville, IL), grid search (*28*), and simplex optimization (*29*) until no further improvement of the reduced χ^2 was obtained. Adaptive grid refinement has the advantage of locating multiple minima in a χ^2 surface and contracting a grid around each simultaneously. No evidence of multiple minima was obtained in the fitting of the data. The extent to which the 11 model parameters could be defined by the experimental data was determined by statistical measures of the precision of each parameter defined by the variation required to increase the reduced χ^2 by 1 (*28*). A mapping of the χ^2 surface by varying each parameter demonstrated that in some cases the depth of the minimum in the reduced χ^2 surface in that parameter space was less than 1, and an upper and/or lower limit for such parameters could not be defined. In such cases, the precision limit is reported to be infinite, indicating a lack of definition of the parameter by the data (e.g., the lower limit for pK_n).

The temperature dependence of ΔG, ΔH, and ΔS was calculated using T_m, $\Delta H(T_m)$, and ΔC_p with the following equations:

$$\Delta H = \Delta H(T_m) + \Delta C_p(T - T_m) \quad (1)$$

$$\Delta S = \Delta S(T_m) + \Delta C_p \ln(T/T_m) \quad (2)$$

$$\Delta G = \Delta H\left(1 - \frac{T}{T_m}\right) + \Delta C_p\left[T - T_m - T\ln\left(\frac{T}{T_m}\right)\right] \quad (3)$$

where $\Delta S(T_m) = \Delta H(T_m)/T_m$. The errors in ΔH, ΔS, and ΔG were defined by the square root of the respective variances calculated using standard propagation of error equations and assuming negligible covariances (*25*).

Nuclear Magnetic Resonance. NMR spectra were collected on Varian (Palo Alto, CA) INOVA 500 and 800 MHz NMR spectrometers with 5 mM protein samples in a 10% D_2O/90% H_2O mixture at 30 °C. The pH of the sample was adjusted to 4.5 with HCl using a Radiometer glass electrode without correction for the deuterium isotope effect. The protein sample used for amide hydrogen exchange was dissolved in 99% D_2O with 0.3 M KCl.

NMR data were processed using Varian VNMR software, NMRPipe (*30*), FELIX (Accelerys, San Diego, CA), and NMRView (*31*). Two-dimensional (2D) DQF-COSY, TOCSY, and NOESY spectra were collected using standard procedures as described for Sac7d with a 6000 Hz spectral bandwidth in both dimensions (*10*, *32*). Data sets typically consisted of 256 t_1 increments, with 32 FIDs per increment and 1024 complex data points in the t_2 acquisition time domain. The data were zero-filled to 2048 points in both dimensions and apodized with the Lorentzian−Gaussian window function. $^3J_{HNH\alpha}$ coupling constants were determined from DQF-COSY spectra with increased resolution using 256 t_1 increments with 8192 complex data points in t_2. $^1H-^{15}N$ HSQC spectra were collected on ^{15}N-enriched samples in 0.7 mL of a 90% H_2O/10% D_2O mixture with gradient selection and minimal perturbation of the water (*33*, *34*). ^1H chemical shifts are referenced relative to the appropriate shift of the water line, which had been previously measured at 5, 35, and 55 °C relative to TSP in separate experiments without protein. ^{15}N chemical shifts were calculated from the relative frequencies (Ξ) (*13*) of DSS (^1H) and liquid ammonia (^{15}N). W24A $^1H-^{15}N$ HSQC spectral peak assignments were made by comparison to the Sac7d HSQC assignments (*32*) and confirmed using a combination of three-dimensional (3D) NOESY-HSQC and TOCSY-HSQC data (*34*). Proton assignments were made using 2D DQF-COSY, TOCSY, and NOESY spectra.

$^{15}N-^1H$ residual dipolar couplings were measured using a 2D $^1H-^{15}N$ HSQC IPAP pulse sequence from L. Kay's laboratory (http://abragam.med.utoronto.ca/software.html) with 422 and 256 complex points in the ^1H and ^{15}N dimensions, respectively. The data were apodized with a shifted sine bell squared function, linear predicted in the ^{15}N dimension, and zero filled to 512 × 1024 points. Couplings were measured using in-house scripts written for NMRView (*31*). The proteins were aligned with *n*-alkyl-poly(ethylene glycol)/hexanol media (5% C12E5, $r = 0.85$) and gave a ^2H splitting of 24 Hz at 30 °C (*35*). The magnitudes of the axial (D_a) and rhombic (R) components of the alignment tensor used for initial structure refinement were determined from the distribution of the $^{15}N-^1H$ dipolar couplings to be 8 and 0.4 Hz, respectively (*36*).

Initial structure refinement and automated assignment of ambiguous NOEs were performed using ARIA 1.2 (*37*) with 370 NOE volumes. Additional restraints for ARIA included 48 $^3J_{HNH\alpha}$ coupling constants, 13 hydrogen bonds, and 59 $^{15}N-^1H$ residual dipolar coupling measurements. The final distance restraints assigned by ARIA consisted of 262 unambiguous restraints, of which 160 were medium- and long-range restraints. Refined values of the axial (D_a) and rhombic (R) components of the alignment tensor were calculated from the ARIA structure to be 7.8 and 0.41 Hz, respectively. The final set of restraints was used to generate a set of 100 structures using Cartesian simulated annealing with CNS 1.1 (*38*) at an initial temperature of 10 000 K. The 13 structures with the fewest restraint violations and best stereochemistry were selected for a brief refinement with a full electrostatic potential in the presence of water, and the results presented here are based on these models. A single structure for comparison was generated by energy minimization of the average of the ensemble of best structures (PDB entry 1XX8).

Solvent Accessible Surface Areas. Extents of solvent exposure of polar and nonpolar surface areas were calculated using Naccess with a 1.4 Å water probe (*39*). Accessible surface areas for unfolded proteins were calculated using the standard accessibilities in Naccess for an AXA tripeptide.

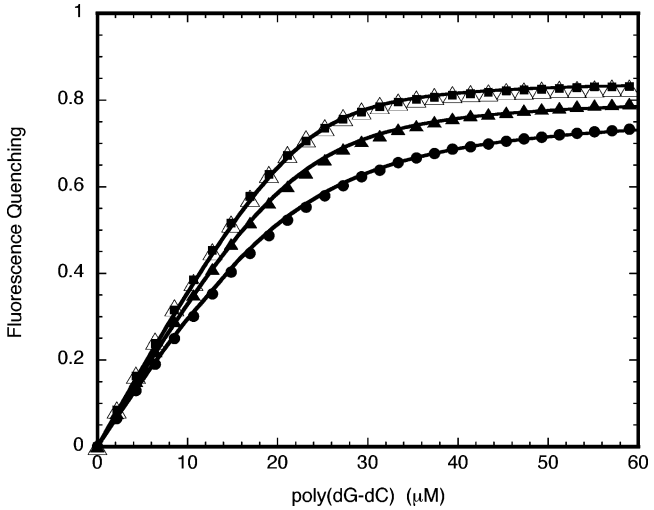

FIGURE 2: Reverse titrations of Sac7d in the presence of various concentrations of W24A. Intrinsic fluorescence quenching of Sac7d by poly(dGdC) was monitored at 20 °C in 10 mM K_2HPO_4 (pH 6.8) and 50 mM KCl. Titrations were performed with 2×10^{-6} M Sac7d and 0 (△), 2×10^{-6} (■), 1.4×10^{-5} (▲), and 3.0×10^{-5} M W24A (●). Fitting of the data with the noncooperative McGhee−von Hippel model is indicated by the solid curves.

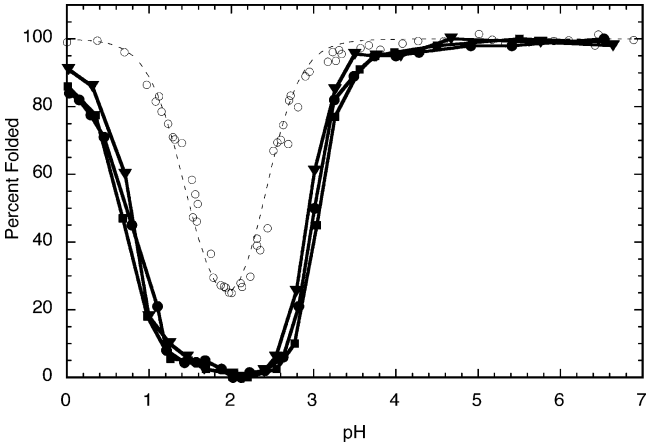

FIGURE 3: Extent of folding of Sac7d and W24 mutant proteins as a function of pH at 20 °C under low-salt conditions (0.001 M glycine). Folding of W24A (●), W24T (▼), and W24F (■) is compared to that of native Sac7d (○) (*42*). The extent of folding was obtained from the observed CD at 205 nm ($\Delta\epsilon^{205}_{obs}$) according to the relation $(\Delta\epsilon^{205}_{obs} - \Delta\epsilon^{205}_{U})/(\Delta\epsilon^{205}_{N} - \Delta\epsilon^{205}_{U})$, where $\Delta\epsilon^{205}_{N}$ is the maximal CD observed at 205 nm at 20 °C and pH 7 and $\Delta\epsilon^{205}_{U}$ is the CD of the unfolded state at 20 °C obtained by extrapolating the high-temperature baseline in a thermal unfolding at pH 2 to 20 °C (*42*).

RESULTS

UV, Fluorescence, and CD Spectral Changes. Substitution of the single tryptophan (W24) in Sac7d with alanine, threonine, or phenylalanine (W24A, W24T, and W24F, respectively) led to a reduction in UV absorbance with ϵ_{278} decreased from 1.1 to 0.34 M^{-1} cm^{-1}. The far-UV CD spectra (180−260 nm) of W24A, W24T, and W24F were essentially identical and differed slightly from that of Sac7d (data not shown). Analysis of the spectra by the variable selection method of Johnson (*40*) indicated that the β-sheet content remained the same as that of Sac7d (24%), but the level of α-helical structure was reduced to 18% (compared to 25% in native Sac7d).

DNA Binding. Substitution of W24 with alanine resulted in a significant loss of DNA binding affinity. DNA binding was monitored indirectly by assessing the ability of W24A to compete with native Sac7d in binding to poly(dGdC). Figure 2 shows poly(dGdC)-induced quenching of the intrinsic tryptophan fluorescence of native Sac7d in the presence of three concentrations of W24A. Alanine substitution of the surface tryptophan dramatically reduced the binding affinity of the protein so that the binding isotherm observed for native Sac7d in the presence of equimolar W24A (2×10^{-6} M) was essentially indistinguishable from that of Sac7d alone. A 10-fold excess of W24A led to a significant reduction in the level of binding by Sac7d (Figure 2), and fitting of the data yielded dissociation constants for binding to poly(dGdC) of $(4.5 \pm 0.5) \times 10^{-5}$ M^{-1} for W24A and $(1.5 \pm 0.5) \times 10^{-7}$ M^{-1} for native Sac7d in 10 mM KH_2PO_4 buffer (pH 6.8) at 20 °C, with identical size sizes for both proteins of 4.5 ± 0.2 bp. Threonine substitution (W24T) was slightly less detrimental with a dissociation constant of $(2.5 \pm 0.5) \times 10^{-5}$ M^{-1}, and phenylalanine was the least disruptive with a K_d of 3×10^{-6} M^{-1} observed for W24F.

Thermal Stability. DSC of W24A in 0.3 M KCl and 10 mM acetate (pH 7) indicated that unfolding of the mutant protein was reversible and exhibited a T_m of 81.1 °C and a ΔH_{vh} of 46.8 kcal/mol. The T_m and ΔH_{vh} were significantly lower than those observed for Sac7d (T_m of 90.1 °C and ΔH_{vh} of 58.5 kcal/mol), indicating that the surface tryptophan played a significant role in stabilizing Sac7d under these conditions. The decrease in T_m observed for W24A was one of the largest observed in a complete alanine scan of Sac7d surface amino acids, including all five aspartates, seven glutamates, 14 lysines, and four arginines (J. Bedell, N. Szary, A. Clark, S. Edmondson, and J. Shriver, unpublished results). DSC and CD thermal melting experiments indicated negligible differences in the stabilities of W24A, W24T, and W24F (data not shown).

In an analysis of DSC data, it is standard practice to analyze the variation of the enthalpy of unfolding with T_m to obtain a ΔC_p of unfolding using the Kirchhoff relation (*41*) by varying the pH over a range necessary to achieve a significant variation in T_m (e.g., pH <4). Such an approach has been shown to be inaccurate with Sac7d due to chloride binding and acid-induced folding below pH 4 (*42*). This was further complicated in recombinant W24A by the low thermodynamic stability of the protein (even though the T_m was high at pH 7) and the inability to obtain completely folded protein below pH 4 at 25 °C (see below). The calorimetric enthalpies of unfolding obtained under such conditions do not reflect the molar enthalpies. Therefore, DSC was used primarily to determine the T_m and ΔH for unfolding of W24A at high pH in 0.3 M KCl. DSC scans were collected below pH 4 to verify predictions made from parameters obtained from fitting the CD experimental data. Protein aggregation prevented DSC experiments below pH 1.5 for W24A.

Salt and pH Dependence of Stability. CD was used to monitor the extent of W24A unfolding as a function of pH, salt, and temperature. As for Sac7d (*42*), decreasing the pH at 20 °C led to significant unfolding of W24A as indicated by changes in the far-UV CD spectrum (Figure 3). The extent of unfolding was much greater than that observed for the native protein. In low salt (0.001 M glycine), W24A began to unfold below pH 4, and the CD decreased, reaching an essentially flat minimum between pH 1.5 and 2.5. The far-

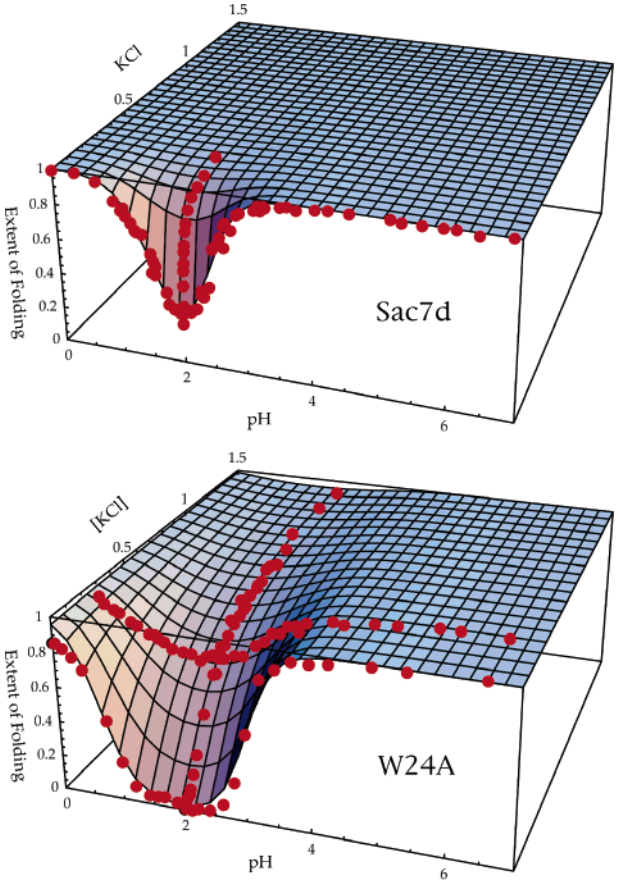

FIGURE 4: Extent of folding of W24A (bottom panel) as a function of pH and KCl concentration at 20 °C. Data defining the extent of folding (red circles) were obtained from the CD at 205 nm as described in the legend of Figure 3 in 0.001 M glycine. The overlayed surface is defined by the parameters in Table 1 obtained from a global fitting of the complete CD and DSC data as described in the text. A comparison to similar data for Sac7d (top panel) (*42*) demonstrates the dramatic changes resulting from the W24A substitution which lead to significantly greater unfolding at low pH as well as a lower binding affinity for chloride.

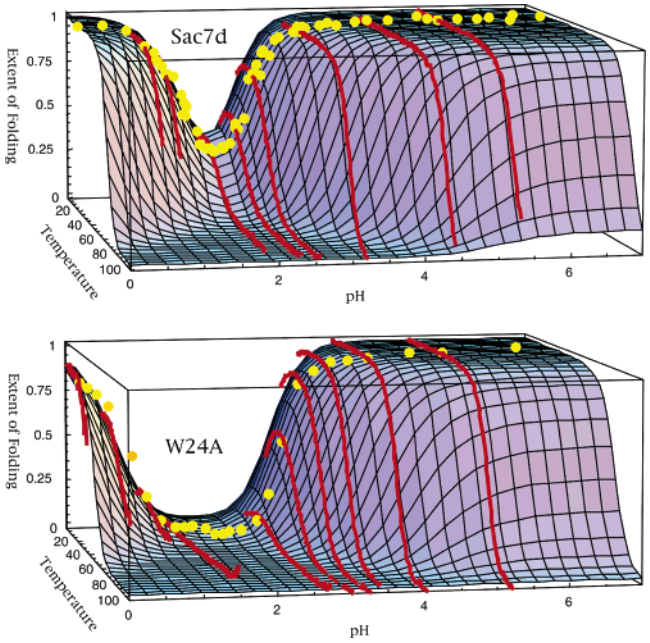

FIGURE 5: pH dependence of the thermal unfolding of W24A (bottom panel) compared to that of Sac7d (top panel). Solid curves show thermal unfolding at various pH values in low salt (0.001 M glycine), and yellow circles show the pH dependence of folding at 20 °C (0 M KCl data from Figure 3). The extent of folding was obtained from CD at 205 nm as described in the legend of Figure 3. The overlayed surface is defined by the fitted parameters obtained from a global fitting of all of the CD and DSC data. Sac7d data are from ref *42*.

UV CD spectra of the protein between pH 1.5 and 2.5 were similar to that observed for the thermally unfolded protein at higher pH. Little or no thermal transition was observed in the CD between pH 1.5 and 2.5. Decreasing the pH below 1.5 led to an increase in CD, indicating acid-induced protein folding. The extent of refolding in W24A was less than that observed for Sac7d [Figure 3 (○)], which was completely refolded to the native state at pH 0 (*32*). At higher salt (e.g., 0.3 M KCl), less unfolding of Sac7d was observed below pH 4, with a minimum of 60% unfolded near pH 2. The extent of unfolding of W24T and W24F as a function of pH at 20 °C was essentially identical to that observed for W24A (Figure 3), consistent with the similar thermal stabilities observed for the three proteins by DSC.

An increasing salt (KCl) concentration at pH 2.0 revealed that a chloride concentration above 1.5 M was required to completely refold W24A at this pH (Figure 4). This is consistent with the inability to completely refold W24A at pH 0, given that the chloride activity at pH 0 is 1 M. We have previously shown that salt-induced refolding of Sac7d at low pH is due to anion binding, and the nature of the cation had little effect. Sulfate and perchlorate were significantly better than chloride at promoting folding of W24A at pH 2 (data not shown), with half-maximal folding achieved with ~0.5 mM SO_4^{2-} or ClO_4^-, while 0.2 M Cl^- was required to achieve a similar effect. This is similar to that observed for the effectiveness of anion-induced folding of native Sac7d and other proteins (*32*, *43*).

Thermal unfolding of W24A as a function of pH and salt concentration was monitored by CD (Figure 5). At low salt (1 mM glycine), the midpoint of the melting transition (T_m) was pH-dependent, and decreased down to pH 2.4 where the protein was essentially completely unfolded. Below pH 1.5, a transition could be observed, and the T_m increased as the pH was decreased to 0. Stabilization of the folded state at 0.3 M KCl permitted observation of a thermal transition over the entire pH range (data not shown). Cold denaturation was evident in the thermal melt data at low pH, where less than 100% folded protein was present at the beginning of the scan. The ability to observe a maximum in the thermal melting scans was important in accurately defining a ΔC_p for unfolding (see below).

Stability Data Analysis. The stability of Sac7d and W24A is controlled in part by electrostatic interactions. This is reflected in the pH and salt dependence of the stability which is dictated by the linkage of anion binding and protonation to protein folding. The steepness of the dependence of folding on pH and salt required at least two anion (chloride) binding sites and the protonation of at least two acidic groups linked to protein folding. A minimal model containing two anion binding sites and two acidic groups was used to parametrize the folding free energy surface as a function of temperature, pH, and salt concentration (*25*). More than two anion binding reactions and two acid ionizations would also be consistent with the data, but were not required. The fitted parameters are similar to those obtained for Sac7d (Table 1), and the

Table 1: Thermodynamic Parameters Characterizing the Temperature, pH and Salt Dependence of the Stability of W24A Compared to That of Native Sac7d[a]

	Sac7d	W24A
	Intrinsic Unfolding Reaction	
T_m	90.7 ± 0.1	81.1 ± 0.1
$\Delta H(T_m)$	58.5 ± 1.0	46.8 ± 1.0
ΔC_p	711 ± 19	719 ± 11
	Chloride Binding Constants and Associated Heats and Heat Capacity Changes	
pK_{Cl} (0.0 M KCl)	−0.14 ± 0.06	0.32 ± 0.14
pK'_{Cl} (0.0 M KCl)	1.84 ± 0.07	1.05 ± 0.2
pK_{Cl} (0.3 M KCl)	−0.14 ± 0.5	−0.12 ± 0.06
ΔH (cal/mol)	−676 ± 1315	−538 ± 1400
ΔC_p (cal deg^{-1} mol^{-1})	−101 ± 122	−80 ± 135
	p*K* Values of Linked Acidic Groups in the Unfolded Protein	
pK_u (0.0 M KCl)	4.67 ± 0.15	4.43 ± 0.04
pK_u (0.3 M KCl)	4.45 ± 0.19	4.47 ± 0.03
	p*K* Values of Linked Acidic Groups in the Folded Protein	
pK_n	0.01 (−∞ → 1.2)	0.12 (−∞ → 1.5)
	p*K* Values of the Linked Acidic Groups in the Folded Protein with Bound Chloride	
pK_l (0.0 M KCl)	2.98 ± 0.03	2.73 ± 0.09
pK_l (0.3 M KCl)	3.15 ± 0.6	3.43 ± 0.03

[a] Parameters are described in Materials and Methods as well as in Clark et al. (*25*). The T_m and $\Delta H(T_m)$ values for the intrinsic unfolding reaction were fixed in a global fit to the experimental values measured from DSC at pH 7 (above the pH range where linkage of carboxyl and chloride binding contribute). Chloride binding constants are reported as pK_{Cl}, i.e., the log of the binding constant. All protonation and chloride binding p*K*, ΔH, and ΔC_p values are defined at 25 °C. Errors [except for those of T_m and $\Delta H(T_m)$] are the precision limits determined by the variation required to increase the reduced χ^2 by 1.0 (*28*). The error in T_m and $\Delta H(T_m)$ is the estimated experimental error from repeated DSC measurements.

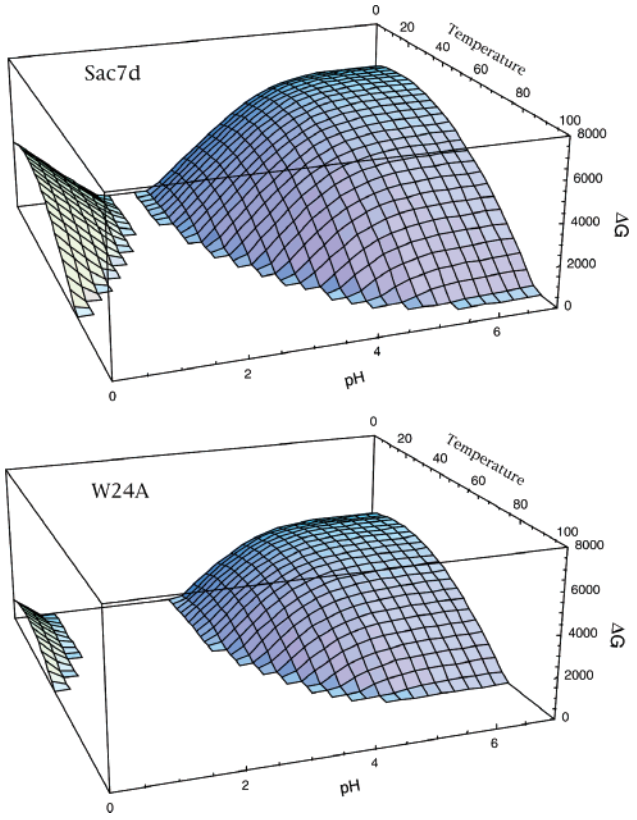

FIGURE 6: Temperature and pH dependence of the unfolding free energy of W24A compared to that of native Sac7d. The free energy of unfolding of W24A was defined by the fitted parameters in Table 1 obtained from a fitting of the pH, salt, and temperature dependence of the extent of folding as described in the text. The surfaces differ primarily by a 2 kcal/mol decrease in the folding free energy of W24A compared to that of Sac7d.

quality of the fitting is demonstrated with overlays of the fitted surfaces onto experimental data (Figures 4 and 5).

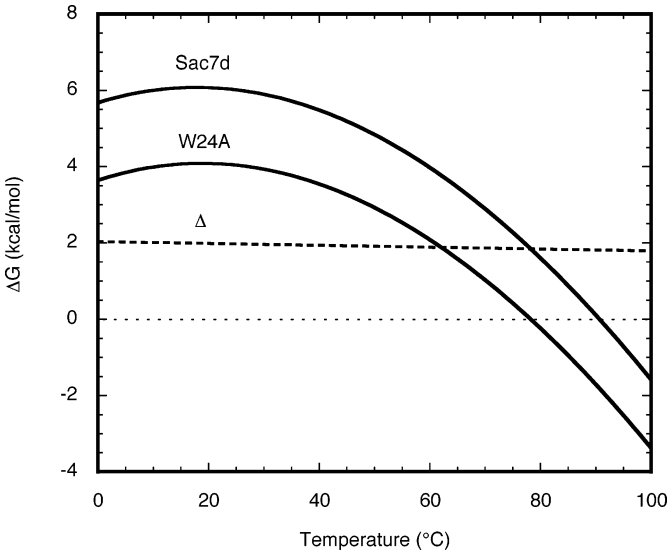

FIGURE 7: Protein stability curves for Sac7d and W24A at pH 7 in low salt. The free energy of unfolding is defined by the linkage model with the fitted parameters in Table 1. The difference (Δ) between the two curves is indicated by the dashed line, with an average $\Delta\Delta G$ of unfolding of 1.9 ± 0.1) kcal/mol over the entire temperature range.

The accuracy of the fitted parameters in defining the temperature, salt, and pH dependence of the free energy and enthalpy of unfolding W24A was indicated by the ability to account for the difficulty in observing a DSC endotherm near pH 2 in 0.3 M KCl. Thermal melts followed by CD indicated that 60% of the protein is folded at 25 °C under these conditions, and a thermal unfolding transition can be observed by CD. However, the fitted parameters demonstrate that the observed $\Delta H°$ of unfolding decreases significantly with pH making it difficult to observe a transition by DSC near pH 2.

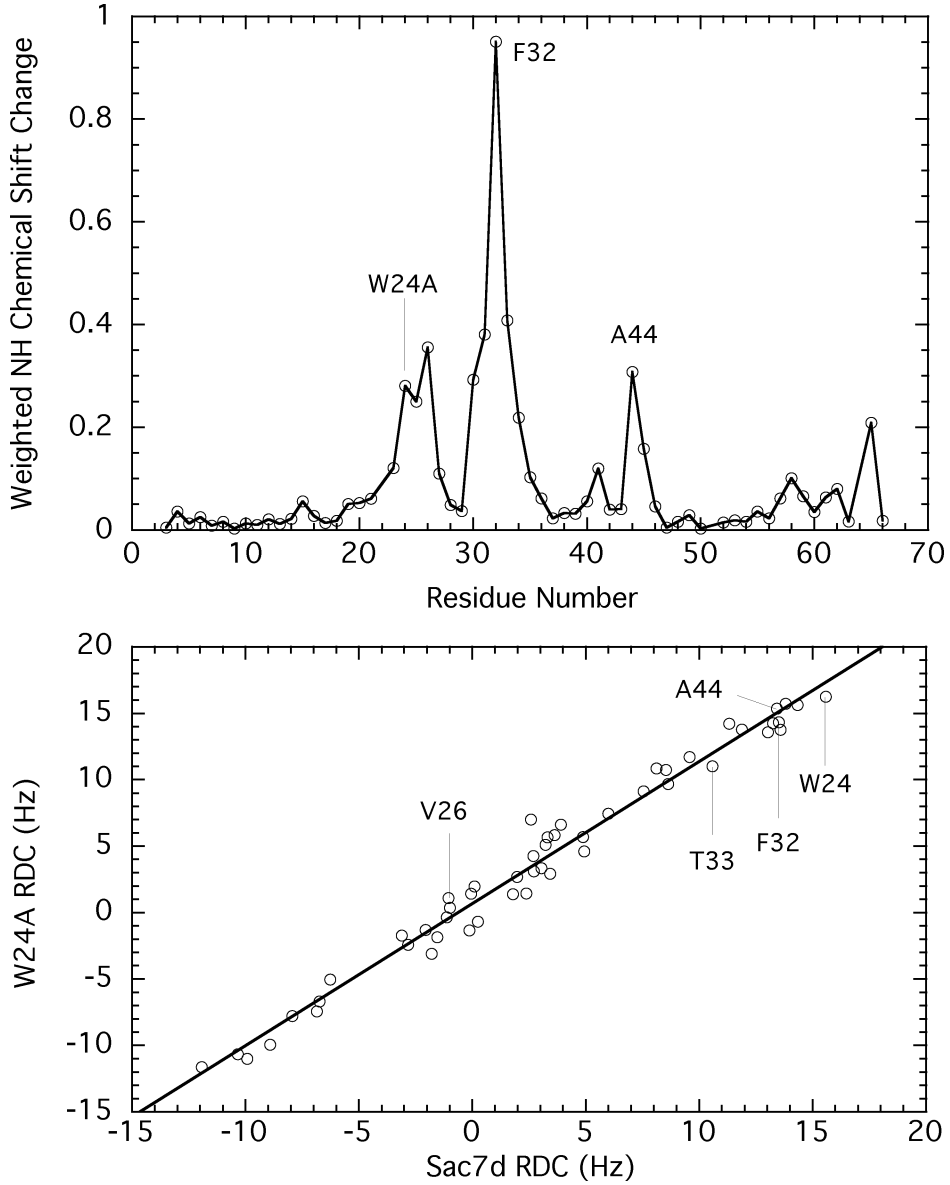

FIGURE 8: Changes in NMR $^{1}H-^{15}N$ HSQC chemical shifts and residual dipolar couplings due to the W24A substitution. The change in chemical shift resulting from the W24A mutation is indicated in the top panel by the absolute value of the vector length of the displacement of the NMR peak with appropriate weighting of the ^{1}H and ^{15}N chemical shifts to account for the difference in magnitude of the two chemical shift scales using the relation $\Delta\delta = \sqrt{(\Delta\delta_{HN})^2 + (0.15\Delta\delta_N)^2}$ (*60*). The residual dipolar coupling for each residue in W24A is plotted against that in native Sac7d in the bottom panel. Selected dipolar couplings are labeled for those residues showing the largest changes in NH chemical shifts. A dipolar coupling for M29 could not be measured because of spectral overlap. The absence of a significant difference between the RDCs of Sac7d and W24A indicates that the changes in chemical shift in the top panel can be attributed to local susceptibility differences and ring current effects due to removal of the aromatic tryptophan side chain.

The free energy surface describing the stability of W24A as a function of temperature, pH, and salt is similar to that observed for Sac7d, except for an overall decrease of ~2 kcal/mol (Figure 6). Cross sections through the surfaces at pH 7 indicate that the stability of W24A is 1.9 ± 0.1 kcal/mol less than that of native Sac7d (Figure 7). The observed free energy of unfolding for Sac7d is 6.2 kcal/mol at pH 7 and 20 °C (in 0.3 M KCl), whereas for W24A, it is 4.2 kcal/mol. The temperatures of maximum stability at pH 7 are essentially identical for the two proteins: 18.9 °C for W24A compared to 17.9 °C for Sac7d. Variation of the salt concentration at pH 7 and 20 °C had little effect on the difference in stability, with a difference of 2.1 ± 0.2 kcal/mol from 0 to 0.3 M KCl (data not shown). The difference in anion binding affinity at low pH is simply a manifestation of the difference in energies of the folded states of the two proteins. Similarly, variation of the pH had little effect, with a difference of 1.9 ± 0.4 kcal/mol from pH 0 to 7. The negligible changes in the pH and salt dependence of Sac7d as a result of the W24A substitution indicate that the loss of the surface tryptophan does not significantly perturb surrounding electrostatic interactions important in defining the stability of Sac7d. We conclude that the 2 kcal/mol decrease in free energy due to the W24A substitution is largely due to the expected changes in hydrophobic, van der Waals, and/or hydrogen bonding interactions.

The decrease in unfolding free energy observed for W24A is primarily due to a decrease in the ΔH of unfolding, with the fitted heat of unfolding decreasing from 7.7 kcal/mol at 20 °C (0.3 M KCl at pH 7) to 5.0 kcal/mol as a result of the W24A substitution, and no significant change in ΔC_p. $T\Delta S$ for unfolding becomes slightly less favorable, decreasing

Table 2: NMR Solution Structure Refinement Statistics

	no.	rmsd
NOE restraints		
short-range NOEs	102	
medium-range NOEs	102	–
long-range NOEs	58	–
total NOEs (unambiguous)	262	0.026 Å
ambiguous NOEs	77	0.009 Å
other restraints		
hydrogen bonds	13	0.047 Å
$^3J_{HNH\alpha}$ coupling constants	48	1.03 Hz
dipolar couplings	59	0.65 Hz
Ramachandran plot (%)		
most favored	69	
additional allowed	23	
generously allowed	6	
disallowed	2[b]	
comparison of best 13 structures		
backbone, using secondary regions[a]		0.71 Å
backbone, all residues		1.34 Å
all heavy atoms		1.84 Å

[a] Residues 3−8, 11−34, and 41−59. [b] Backbone dihedral angles were observed in disallowed regions of a Ramachandran plot for residues N37 (in the loop between the second and third strands of the β-sheet) and K52 (at the beginning of the C-terminal helix).

from 1.5 to 0.8 kcal/mol under these conditions.

NMR Structure. The relatively small size (66 residues), wide chemical shift dispersion, and high solubility made ^{1}H chemical shift assignments of W24A straightforward using 2D NMR spectra with conventional methods. These were confirmed with ^{1}H−^{15}N HSQC, 3D ^{1}H−^{15}N HSQC-NOESY, and HSQC-TOCSY data. ^{1}H−^{15}N HSQC chemical shift differences due to the W24A substitution occurred in residues in the three-stranded β-sheet adjacent to W24A, with the largest change occurring for the NH group of F32 in the adjacent strand (Figure 8, top panel**)**. The NH group of F32 is hydrogen bonded to the carbonyl of G43 in the third strand of the sheet. The χ_1 angle for W24 places the aromatic side chain over the β-sheet so that deletion of this group is expected to cause significant susceptibility changes in the underlying NH groups. The lack of a significant structural change due to the W24A substitution is indicated by the ^{1}H−^{15}N residual dipolar couplings of W24A, which are virtually identical to those measured for Sac7d with a correlation coefficient of 0.99 (Figure 8, bottom panel). RDCs for those NH groups which displayed the largest changes in chemical shifts in the HSQC (Figure 8, top panel) exhibited a negligible change, within the experimental error of 1.0 Hz. We note that the RDCs indicate that the orientation of the NH vector for V26 differs significantly from that of the other NH vectors in the β-sheet. This is due to a noticeable curvature of the sheet at the tight turn between the first and second strands. This most likely plays a role in properly positioning the V24 and M29 side chains for intercalation.

To confirm the above implications that the W24A substitution had little effect on the structure of Sac7d, the solution structure of W24A was defined using ARIA/CNS with RDCs as well as NOE, dihedral angle, and hydrogen bond constraints. The NMR structure was well-defined with a backbone (C′, C$^\alpha$, and N) rmsd over all residues of 1.34 Å (0.71 Å over backbone atoms involved in ordered secondary structure). A summary of the refinement statistics is presented in Table 2. Comparison of the W24A structure with the NMR structure of Sac7d (PDB entry 1sap) confirms that little change in backbone atom positions results from the W24A substitution (Figure 9). Except for W24 in native Sac7d (*10*), the majority of the side chain positions are not well defined in the NMR structures. A backbone rmsd of 1.1 Å between the Sac7d and W24A structures was calculated, which is comparable to the level of precision in the individual structures and indicates that they are identical within experimental error. The rmsd over the DNA-binding three-stranded β-sheet (excluding the unstructured loop between D35 and T40) was 0.81 Å.

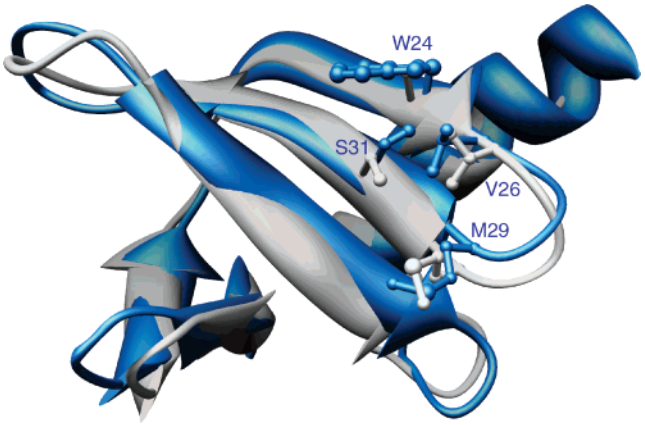

FIGURE 9: Overlay of the NMR solution structure of W24A (white) onto the NMR structure of native Sac7d (blue) demonstrating that little change in the backbone conformation results from the W24A substitution. The structures are the energy-minimized average structures of the ensemble of best structures. Structural alignment was performed by minimizing the differences between backbone atoms of the two structures over regions of well-defined secondary structure (specifically, C′, C$^\alpha$, and N atoms in residues 3−8, 11−34, and 41−59). Side chains for W24, V26, M29, and S31 are indicated to demonstrate the proximity of the residues in the folded protein. Except for W24 (*10*), the side chain rotomers are not well defined by the NMR data. The χ_1 for W24 in the native Sac7d solution structure (PDB entry 1asp) is similar to that in the crystal structure of the DNA complex (PDB entry 1azp).

DISCUSSION

We have shown that the solvent-exposed tryptophan which exists on the DNA binding surface of Sac7d plays an important role in defining the stability of the Sac7d−DNA complex. The results are consistent with expectations based on the crystal structures of Sac7d−DNA complexes which demonstrates that the surface tryptophan W24 packs against the sugar−phosphate backbone and hydrogen bonds to a nucleotide base at the intercalation site. Surprisingly, mutagenesis of W24 results in a significant reduction in the stability of the protein, which could potentially result from structural changes that might be responsible for the loss in DNA binding affinity. A decrease in thermal stability of the homologous Sso7d has been reported as a result of a similar W23A substitution in that protein (*44*). We have therefore performed a careful analysis of changes in both the stability and structure resulting from the W24A substitution in Sac7d. Analysis of the dependence of the stability on pH and salt concentration has permitted an accurate measure of the protein stability surface which demonstrates that the W24A substitution results in a global lowering of the stability curve of the protein. The global linkage analysis indicates that the change in stability cannot be explained by the structural perturbation of surrounding electrostatic interactions for this

highly charged, hyperthermophile protein. This is confirmed by residual dipolar couplings and a refinement of the NMR structure.

Structure of W24A. The structure of W24A was shown to be virtually identical to that of native Sac7d. A slight movement of V26 away from residue 24 by ~1.5 Å indicates that surface hydrophobic interactions between W24 and V26 (and possibly M29) may exist in the native protein, leading to a slight curvature of the β-sheet. However, the magnitude of this difference is small and borders on the precision limit of the NMR structure. The inability to detect amide protons in the C-terminal helix after the protein was dissolved in D_2O indicated that the helix was less stable in W24A than in Sac7d, where slowly exchanging amide protons were observed from residue 53 to 66. Thus, the C-terminal helix has a greater tendency to be disordered in W24A than in Sac7d. This explains the difference in the far-UV CD spectra of W24A and Sac7d at pH 7.0 which indicated that the percentage of α-helix decreased from 25 to 18% in W24A. We attribute the decrease in helical stability in the NMR data and loss of helical content reflected in the CD to the decrease in global stability of W24A and not to a local destabilization. Hydrogen exchange experiments (M. Kahsai and J. Shriver, unpublished results) indicate that the C-terminal helix of Sac7d is less stable than the DNA binding interface defined by the three-stranded β-sheet. A decrease in global stability will therefore increase the probability of local "breathing motions" and hydrogen exchange in the C-terminal helix without changing the structure of the folded state.

DNA Binding Affinity. The removal of the tryptophan side chain resulted in a 300-fold decrease in DNA affinity. Given the lack of significant structural changes in the protein DNA binding site, we attribute this to the loss of specific interactions between the tryptophan and the DNA, including hydrophobic interactions as well as the potential for hydrogen bonding as indicated in the crystal structure. In the crystal structures (PDB entry 1azp), the aromatic side chain of W24 packs against the deoxyriboses of G3 and A4 of the DNA molecule [d(GCGATCGC)$_2$], and the NH group of the tryptophan indole side chain makes a hydrogen bond to N3 of the G3 base (Figure 1). In addition, the lowered stability of W24A and possible increased disorder in the C-terminal helix can lead to protein folding upon DNA binding and an entropy penalty which will be reflected in a lower binding affinity. The fact that W24A, W24F, and W24T demonstrated different binding affinities with essentially the same loss in protein stability would indicate that the predominant factor leading to the loss in affinity was a change at the interface and not protein stability. Notably, the 20-fold lower affinity observed for W24F compared to the 300-fold loss observed for W24A would indicate that hydrophobic packing at the protein−DNA interface outside of the intercalation site is an important part of the binding interaction for this nonspecific DNA binding protein.

Stability of W24A. Crystal structures of a number of Sac7d− and Sso7d−DNA complexes indicate that the single, solvent-exposed tryptophan in Sac7d plays a central role in DNA binding (*8*, *9*, *45*, *46*). Positioning of the W24 side chain in the **p** rotomer ($\chi_1 = 70°$) for optimal DNA interactions exists in the free protein prior to binding. The **p** rotomer with C^γ *trans* to H^α is observed for ~20% of all tryptophans located in β-sheets (*11*, *12*). The conformer population distribution would indicate that this rotomer should be destabilizing. Prepositioning the side chain in a disallowed conformer for proper alignment at the DNA interface can be expected to require specific interactions in the native protein with surrounding residues. It is therefore surprising that removal of the bulky aromatic side chain of tryptophan should lead to little change in structure and one of the largest decreases in stability that has been observed in Sac7d due to single-amino acid substitutions at the surface.

Tryptophan is the least common amino acid in small proteins (*47*). It is also the largest and is generally considered one of the most hydrophobic of the 20 naturally occurring amino acids. A comparison of its distribution between the surface and interior of small proteins (~8 kDa) indicated a nearly 4-fold greater tendency to be located in the interior, which was equal to or greater than that observed for phenylalanine, valine, leucine, isoleucine, and methionine (3.2, 2.1, 4.1, 2.5, and 1.9-fold differences, respectively) (*48*).

Although the hydrophobic effect plays an important role in driving protein folding by burying apolar surface area, the presence of a hydrophobic residue on the surface is not necessarily destabilizing (*49*−*51*). Since the stability of the native protein fold represents a difference in free energy between the folded and unfolded states, destabilization can result from either a lower native state free energy or an increase in the free energy of the unfolded structure. If a surface hydrophobic residue has similar solvent exposure in the native and unfolded states, solvent accessibility should have little effect on the stability of the folded state (*52*). Destabilization can result from steric interactions with neighboring groups in the folded state, "hyper-exposure" in the folded state (*53*), or from preferential residual structure in the unfolded chain leading to the "reverse hydrophobic" effect (*49*). In contrast, stabilization of the folded state by surface hydrophobes could result from formation of surface hydrophobic clusters (*50*). Such an effect appears to be the basis for the increase in stability resulting from increased hydrophobicity in the neutral protease of *Bacillus subtilis* (*54*). Aromatic clusters on the surface of thermophile proteins have been documented and may contribute to enhanced stability (*55*).

The formation of a solvent-exposed hydrophobic cluster involving W24, S31, V26, and M29 is an obvious possible explanation for the importance of W24 in stabilizing Sac7d. Substitution of W24 with an alanine in the NMR structure of Sac7d (PDB entry 1sap) results in an increase in the nonpolar surface area of 31 Å^2 due to the increased exposure of the neighboring hydrophobic side chains.[2] However, comparing the surface area of the folded protein to that calculated from residues in an unfolded chain indicates that unfolding Sac7d leads to an exposure of nonpolar surface area (by 150 Å^2) greater than that calculated for the W24A substitution. Thus, on this basis, the native protein would be expected to be less stable than W24A.

Preliminary experimental evidence for the importance of a surface hydrophobic cluster comes from a demonstration

[2] This change is calculated by making the substitution in a single structure (PDB entry 1sap) rather than comparing the accessible surface areas calculated from the NMR structures of Sac7d and W24A because the differences are small and sensitive to side chain positions, many of which are not precisely defined by the NMR data.

of the importance of hydrophobicity of V26 and M29. Alanine substitution of V26 and M29 leads to a loss in stability comparable in magnitude to that caused by the W24A substitution as reflected by a decrease in T_m from 90 to 81 °C (W. Peters, S. P. Edmondson, and J. Shriver, unpublished results). However, perturbation of a surface hydrophobic cluster cannot explain all of the effects of mutating W24 on the stability of Sac7d. Substitution of W24 with alanine, threonine, and phenylalanine led to essentially identical losses in stability. In addition, increasing the hydrophobicity of V26 and M29 with double isoleucine, leucine, or phenylalanine substitutions resulted in either no change or a slight decrease in stability (T_m values of 90, 88, and 87 °C, respectively). Finally, the substitution of alanine for W24 in the double V26A/M29A mutant led to a further loss in stability, which was greater than that observed for the single W24A substitution in native Sac7d (T_m decreased from 81 to 68 °C).

This would indicate that the stabilizing contribution made by W24 cannot be attributed to a general nonspecific hydrophobic effect, but rather, effects specific to tryptophan must play a role. The β-sheet propensity of tryptophan can be eliminated since both threonine and phenylalanine have similar or higher propensities (*56*, *57*). Tryptophan may also form cation$-\pi$ interactions, but there is no evidence in the NMR structure (or the X-ray structures) for such an interaction. A more likely specific contribution might come from hydrogen bonding of the S31 hydroxyl to the aromatic rings of W24 (*58*). The distance between the hydroxyl oxygen and the NE1 atom of W24 is 3 Å. However, a survey of known protein structures indicates that this interaction is not prevalent, and therefore, it is expected to confer only a small advantage (*58*). A detailed study of the influence of S31 on the structure, stability, and DNA binding of Sac7d will be presented elsewhere.

REFERENCES

1. Dijk, J., and Reinhardt, R. (1986) in *Bacterial Chromatin* (Gualerzi, C., and Pon, C., Eds.) Springer-Verlag, Berlin.
2. Edmondson, S. P., and Shriver, J. W. (2001) DNA binding proteins Sac7d and Sso7d from *Sulfolobus*, *Methods Enzymol. 334*, 129−145.
3. Mai, V. Q., Chen, X., Hong, R., and Huang, L. (1998) Small abundant DNA binding proteins from the thermoacidophilic archaeon *Sulfolobus shibatae* constrain negative DNA supercoils, *J. Bacteriol. 180*, 2560−2563.
4. Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y., Yoshizawa, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Masuda, S., Yanagii, M., Nishimura, M., Yamagishi, A., Oshima, T., and Kikuchi, H. (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7, *DNA Res. 8*, 123−140.
5. White, M. F., and Bell, S. D. (2002) Holding it together: Chromatin in the Archaea, *Trends Genet. 18*, 621−626.
6. Reeve, J. N. (2003) Archaeal chromatin and transcription, *Mol. Microbiol. 48*, 587−598.
7. Eissenberg, J. C. (2001) Molecular biology of the chromo domain: An ancient chromatin module comes of age, *Gene 275*, 19−29.
8. Robinson, H., Gao, Y. G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (1998) The hyperthermophile chromosomal protein Sac7d sharply kinks DNA, *Nature 392*, 202−205.
9. Gao, Y. G., Su, S. Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (1998) The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA, *Nat. Struct. Biol. 5*, 782−786.
10. 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*, 13289−13304.
11. Lovell, S. C., Word, J. M., Richardson, J. S., and Richardson, D. C. (2000) The penultimate rotamer library, *Proteins 40*, 389−408.
12. Janin, J., and Wodak, S. (1978) Conformation of amino acid side chains in proteins, *J. Mol. Biol. 125*, 357−386.
13. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wuthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids, *J. Mol. Biol. 280*, 933−952.
14. Peters, W. B., Edmondson, S. P., and Shriver, J. W. (2004) Thermodynamics of DNA binding and distortion by the hyperthermophile chromatin protein Sac7d, *J. Mol. Biol. 343*, 339−360.
15. McAfee, J. G., Edmondson, S. P., Datta, P. K., Shriver, J. W., and Gupta, R. (1995) Gene cloning, expression, and characterization, of the Sac7 proteins from the hyperthermophile *Sulfolobus acidocaldarius*, *Biochemistry 34*, 10063−10077.
16. Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins, *Biochemistry 6*, 1948−1954.
17. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319−326.
18. Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E., and Cantor, C. R. (1970) Physicochemical studies on polydeoxyribonucleotides containing defined repeating nucleotide sequences, *J. Mol. Biol. 54*, 465−497.
19. Gordon, A., and Ford, R. (1972) *The Chemist's Companion: A Handbook of Practical Data, Techniques, and References*, John Wiley, New York.
20. McAfee, J. G., Edmondson, S. P., 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*, 4034−4045.
21. McGhee, J. D., and von Hippel, P. H. (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*, 469−489.
22. Savitsky, A., and Golay, M. J. E. (1964) Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem. 36*, 1627.
23. Johnson, W. C. J. (1984) *Food Analysis Principles and Techniques, Volume 2, Physiochemical Techniques*, Marcel Dekker, Inc., New York.
24. Shriver, J., Peters, W., Szary, N., Clark, A., and Edmondson, S. (2001) Calorimetry of hyperthermophile proteins, *Methods Enzymol. 334*, 389−422.
25. Clark, A. T., McCrary, B. S., Edmondson, S. P., and Shriver, J. W. (2004) Thermodynamics of core hydrophobicity and packing in the hyperthermophile proteins Sac7d and Sso7d, *Biochemistry 43*, 2840−2853.
26. Pitzer, K. S., and Mayorga, G. (1973) Thermodynamics of electrolytes. II Activity and osomotic coefficients for strong electrolytes with one or both ions univalent, *J. Phys. Chem. 77*, 2300−2308.
27. Pitzer, K. S. (1973) Thermodynamics of electrolytes. I. Theoretical basis and general equations, *J. Phys. Chem. 77*, 268−277.
28. Bevington, P. R., and Robinson, D. K. (1992) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.
29. Press, W., Flannery, B., Teukolsky, S., and Vetterling, W. (1989) *Numerical Recipes: The Art of Scientific Computing (Fortran Version)*, Cambridge University Press, Cambridge, U.K.
30. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
31. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol. Biol. 278*, 313−352.
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*, 1878−1888.

33. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc. 114*, 10663−10665.
34. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques, *J. Biomol. NMR 4*, 845−858.
35. Ruckert, M., and Otting, G. (2000) Alignment of biological macromolecules in novel nonionic liquid crystlline media for NMR experiments, *J. Am. Chem. Soc. 122*, 7793−7797.
36. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information, *J. Magn. Reson. 133*, 216−221.
37. Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) Automated assignment of ambiguous nuclear Overhauser effects with ARIA, *Methods Enzymol. 339*, 71−90.
38. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
39. Hubbard, S. J., Campbell, S. F., and Thornton, J. M. (1991) Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors, *J. Mol. Biol. 220*, 507−530.
40. Manavalan, P., and Johnson, W. C. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra, *Anal. Biochem. 167*, 76−85.
41. Lewis, G. N., Randall, M., Pitzer, K. S., and Brewer, L. (1961) *Thermodynamics*, McGraw-Hill, New York.
42. 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*, 203−224.
43. Goto, Y., Takahashi, N., and Fink, A. (1990) Mechanism of acid-induced folding of proteins, *Biochemistry 29*, 3480−3488.
44. Catanzano, F., Graziano, G., Fusi, P., Tortora, P., and Barone, G. (1998) Differential scanning calorimetry study of the thermodynamic stability of some mutants of Sso7d from *Sulfolobus solfataricus*, *Biochemistry 37*, 10493−10498.
45. Su, S., Gao, Y. G., Robinson, H., Liaw, Y. C., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (2000) Crystal structures of the chromosomal proteins Sso7d/Sac7d bound to DNA containing T-G mismatched base-pairs, *J. Mol. Biol. 303*, 395−403.
46. Ko, T. P., Chu, H. M., Chen, C. Y., Chou, C. C., and Wang, A. H. (2004) Structures of the hyperthermophilic chromosomal protein Sac7d in complex with DNA decamers, *Acta Crystallogr. D60*, 1381−1387.
47. White, S. H. (1992) Amino acid preferences of small proteins: Implications for protein stability and evolution, *J. Mol. Biol. 227*, 991−995.
48. Miller, S., Janin, J., Lesk, A., and Chothia, C. (1987) Interior and surface of monomeric proteins, *J. Mol. Biol. 196*, 641−656.
49. Pakula, A. A., and Sauer, R. T. (1990) Reverse hydrophobic effects relieved by amino-acid substitutions at a protein surface, *Nature 344*, 363−364.
50. Tisi, L. C., and Evans, P. A. (1995) Conserved structural features on protein surfaces: Small exterior hydrophobic clusters, *J. Mol. Biol. 249*, 251−258.
51. Krowarsch, D., and Otlewski, J. (2001) Amino-acid substitutions at the fully exposed P1 site of bovine pancreatic trypsin inhibitor affect its stability, *Protein Sci. 10*, 715−724.
52. Schwehm, J. M., Kristyanne, E. S., Biggers, C. C., and Stites, W. E. (1998) Stability effects of increasing the hydrophobicity of solvent-exposed side chains in staphylococcal nuclease, *Biochemistry 37*, 6939−6948.
53. Tamura, A., and Sturtevant, J. M. (1995) A thermodynamic study of mutant forms of *Streptomyces* subtilisin inhibitor. III. Replacements of a hyper-exposed residue, Met73, *J. Mol. Biol. 249*, 646−653.
54. Frigerio, F., Margarit, I., Nogarotto, R., de Filippis, V., and Grandi, G. (1996) Cumulative stabilizing effects of hydrophobic interactions on the surface of the neutral protease from *Bacillus subtilis*, *Protein Eng. 9*, 439−445.
55. Kannan, N., and Vishveshwara, S. (2000) Aromatic clusters: A determinant of thermal stability of thermophilic proteins, *Protein Eng. 13*, 753−761.
56. Kim, C. A., and Berg, J. M. (1993) Thermodynamic β-sheet propensities measured using a zinc-finger host peptide, *Nature 362*, 267−270.
57. Smith, C. K., Withka, J. M., and Regan, L. (1994) A thermodynamic scale for the β-sheet forming tendencies of the amino acids, *Biochemistry 33*, 5510−5517.
58. Samanta, U., Pal, D., and Chakrabarti, P. (2000) Environment of tryptophan side chains in proteins, *Proteins 38*, 288−300.
59. Dickerson, R. E., and Chiu, T. K. (1997) Helix bending as a factor in protein/DNA recognition, *Biopolymers 44*, 361−403.
60. Tugarinov, V., and Kay, L. E. (2003) Quantitative NMR studies of high molecular weight proteins: Application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G, *J. Mol. Biol. 327*, 1121−1133.

BI047823B