Energetics of Sequence-Specific Protein−DNA Association: Conformational Stability of the DNA Binding Domain of Integrase Tn916 and Its Cognate DNA Duplex[†]

Stoyan Milev,[‡] Alemayehu A. Gorfe,[‡] Andrey Karshikoff,[§] Robert T. Clubb,[||] Hans Rudolf Bosshard,[‡] and Ilian Jelesarov*,[‡]

Institute of Biochemistry, University of Zürich, Winterthurerstrasse 190m Room 44 L42, CH-8057 Zürich, Switzerland, Center for Structural Biochemistry, Department of Biosciences, NOVUM, Karolinska Institutet, S-141 57 Huddinge, Sweden, and Department of Chemistry and Biochemistry and UCLA-DOE Laboratory of Structural Biology and Genetics, University of California, 405 Hilgard Avenue, Los Angeles, California 90095

Received October 1, 2002; Revised Manuscript Received January 16, 2003

ABSTRACT: Sequence-specific DNA recognition by bacterial integrase Tn916 involves structural rearrangements of both the protein and the DNA duplex. Energetic contributions from changes of conformation, thermal motions and soft vibrational modi of the protein, the DNA, and the complex significantly influence the energetic profile of protein−DNA association. Understanding the energetics of such a complicated system requires not only a detailed calorimetric investigation of the association reaction but also of the components in isolation. Here we report on the conformational stability of the integrase Tn916 DNA binding domain and its cognate 13 base pair target DNA duplex. Using a combination of temperature and denaturant induced unfolding experiments, we find that the 74-residue DNA binding domain is compact and unfolds cooperatively with only small deviation from two-state behavior. Scanning calorimetry reveals an increase of the heat capacity of the native protein attributable to increased thermal fluctuations. From the combined calorimetric and spectroscopic experiments, the parameters of protein unfolding are $T_m = 43.8 \pm 0.3$ °C, $\Delta H_m = 255 \pm 18$ kJ mol^{-1}, $\Delta S_m = 0.80 \pm 0.06$ kJ mol^{-1}, and $\Delta C_p = 5.0 \pm 0.8$ kJ K^{-1} mol^{-1}. The DNA target duplex displays a thermodynamic signature typical of short oligonucleotide duplexes: significant heat absorption due to end fraying and twisting precedes cooperative unfolding and dissociation. The parameters for DNA unfolding and dissociation are $\Delta H_m = 335 \pm 4$ kJ mol^{-1} and $\Delta C_p = 2.7 \pm 0.9$ kJ K^{-1} mol^{-1}. The results reported here have been instrumental in interpreting the thermodynamic features of the association reaction of the integrase with its 13 base pair target DNA duplex reported in the accompanying paper [Milev et al. (2003) *Biochemistry 42*, 3481−3491].

Antibiotic resistance of bacteria is spread by promiscuous conjugative transposons (*1*). One of the most thoroughly studied representatives of the family is the Tn916 element carrying resistance to tetracycline. Excision of the Tn916 transposon requires the formation of a complicated nucleoprotein complex similar to the λ-phage "intasome" (*2*). In the synaptic complex, the C-terminal catalytic domain of the transposon-encoded integrase is brought to the cleavage site by the N-terminal DNA binding domain, which binds to repeated sequences within the transposon arm.

The solution structure of the N-terminal DNA binding domain complexed to a 13 bp DNA duplex has been solved by NMR spectroscopy (*3*). Unlike most other major groove binders, the Tn916 integrase recognizes its DNA target by positioning the face of a three-stranded β-sheet into the major groove. The conformation of the protein is changed in the complex when compared to its free native form (*4*).[1] Surprisingly, the protein appears more disordered in the DNA-bound state (*3*). The DNA target site keeps a B-form conformation in the bound state except that the major groove is widened in the middle of the duplex.

In the accompanying paper, we describe the energetics of the association reaction of the DNA binding domain with the 13 bp target DNA duplex (*5*). The reaction exhibits a large negative heat capacity change, ΔC_p, which is itself temperature-dependent and cannot be accounted for by the amount of polar and nonpolar surface buried at the complex interface, as would be the case for a rigid body type of association. A steep increase of the heat capacity of the complex was observed before the thermal transition into the unfolded protein and dissociated DNA strands. This behavior could only be interpreted with the help of the thermodynamic parameters of unfolding of the free protein and the duplex DNA described in the present paper (*5*).

[†] This work was supported in part by the Swiss National Science Foundation.

* Corresponding author. Telephone: +41 1 655 5547. Fax: +41 1 635 6805. E-mail: iljel@bioc.unizh.ch

[‡] University of Zürich.

[§] Karolinska Institutet.

[||] University of California, Los Angeles.

[1] To facilitate reading, the N-terminal fragment 2−74 of integrase Tn916 and the 13 bp duplex target DNA are called "protein" and "DNA", respectively.

10.1021/bi026936x CCC: $25.00
Published on Web 03/08/2003

MATERIALS AND METHODS

Materials. Experiments were conducted in standard buffer composed of 50 mM Na-phosphate, 100 mM NaCl pH 6, ionic strength 0.16 M. The pH of samples containing urea or GdmCl[2] was adjusted after addition of the denaturant. Urea and GdmCl concentrations were determined by refractometry. All chemicals were of analytical grade and were used without further purification. Overexpression and purification of the N-terminal 74-residue DNA binding domain containing a Cys57Ala mutation was performed as described (*5*). Single-stranded oligonucleotides were purchased from Metabion GmbH (Martinsried) and were purified by HPLC. For duplex preparation, equimolar amounts of the two complementary strands were mixed and annealed by heating to 70 °C and slowly cooling to room temperature. Concentrations were determined from A_{260} after dialysis and complete digestion by phosphodiesterase I (Sigma).

Methods. CD measurements were carried out on a Jasco J-715 instrument and fluorescence measurements on a Perkin-Elmer LS 50B luminescence spectrophotometer as described (*5*). In thermal unfolding experiments by CD the continuous heating rate was 0.5 or 1 deg min^{-1}. In thermal unfolding experiments by fluorescence, the temperature of the Peltier-thermostated cuvette holder was increased stepwise with 5 min equilibration time between steps. The temperature was controlled by a sensor in physical contact with the sample solution. Denaturant induced conformational changes were monitored from the position of the fluorescence emission maximum induced by excitation at 295 nm, after 2−5 μM protein had been incubated for 12−15 h at the desired denaturant concentration. DSC experiments were performed on a VP-DSC calorimeter (MicroCal Inc) as described (*5*).

Analysis of Spectroscopic Data. Analysis of heat-induced and denaturant-induced unfolding curves followed the formalism describing a simple two-state transition between the folded and unfolded state, N and U, respectively. At each temperature or denaturant concentration, the observed signal *S*, representing either the molar ellipticity or the shift of the fluorescence emission maximum, is given by

$$S = f_U S_U + (1 - f_U) S_N \quad (1)$$

where f_U is the fraction of unfolded molecule and S_U and S_N refer to the folded and unfolded state, respectively. S_U and S_N were assumed to be linear functions of temperature or denaturant concentration of the general form $S_\sigma = S_{\sigma,0} + \alpha_\sigma T$ or $S_\sigma = S_{\sigma,0} + \alpha_\sigma$[denaturant], with σ indicating the N or U state. For a given total concentration $[N]_0$, the equilibrium unfolding constant K_U is defined by

$$K_U = \frac{f_U^n [N]_0^{n-1}}{1 - f_U} \prod_i^k m_i^{m_i} \quad (2)$$

where m_i is the stoichiometric coefficient of an unfolding subunit (if unfolding is coupled to subunit dissociation) and $n = \sum_i^k m_i$ is the order of the reaction. For the monomeric transition of the protein, $K_U = (f_U/1 - f_U)$. For the duplex DNA, $K_U = (f_U^2[N]_0/1 - f_U)$. The unfolding enthalpy ΔH_m at the transition midpoint T_m was obtained from the temperature dependence of K_U according to

$$K_U(T) = K_U(T_m) \exp\left\{\frac{\Delta H_m}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) - \frac{\Delta C_p}{RT}\left[T - T_m - T\ln\left(\frac{T}{T_m}\right)\right]\right\} \quad (3)$$

R is the gas constant and ΔC_p is the heat capacity change. Equations 1−3 can be combined to analyze thermal unfolding curves by nonlinear curve optimization. Since T_m, ΔH_m, and ΔC_p are strongly interdependent, the statistical significance of the values extracted from the analysis of a single trace is low. However, when ΔC_p is small compared to ΔH_m, the second term in the curly brackets of eq 3 can be neglected.

In a different treatment, S_N and S_U were defined from the linear pre-transitional and post-transitional parts of the melting trace to obtain a plot of $f_U = (S - S_N/S_U - S_N)$ against *T*. From such plots, ΔH_m was calculated as (*6*)

$$\Delta H_m = (2n + 2) R T_m^2 \left(\frac{\partial f_U}{\partial T}\right)_{T=T_m} \quad (4)$$

ΔC_p was obtained from plots of ΔH_m versus *T* (Kirchoff plots). To this end, thermal stability was varied by changing the solvent conditions. Knowing T_m, ΔH_m, and ΔC_p, the unfolding free energy ΔG_U was calculated from the Gibbs−Helmholtz equation:

$$\Delta G_U(T) = \Delta H_m(T_m)\cdot\left(1 - \frac{T}{T_m}\right) + \Delta C_p\left[T - T_m - T\ln\left(\frac{T}{T_m}\right)\right] - RT\ln K_U \quad (5)$$

Data from isothermal denaturant titrations were analyzed following the linear extrapolation mode (LEM). The free energy of unfolding in the presence of denaturants, ΔG_U(D), is given by

$$\Delta G_U(D) = -RT\ln K_U(D) = \Delta G_U(W) - m_D[D] \quad (6a)$$

where m_D has units of kJ mol^{-1} M^{-1} and describes the linear dependence of ΔG_U(D) on [D]. ΔG(W) is the free energy of unfolding in the absence of denaturant. Since for a monomolecular unfolding transition K_U is 0 at [D] = $[D]_{1/2}$, we can write

$$\Delta G_U(D) = m_D[D] - m_D[D]_{1/2} \quad (6b)$$

Equations 1, 2, and 5 were combined to obtain a fitting function suitable to perform a nonlinear regression analysis of the unprocessed experimental data resulting in optimized values for m_D, $[D]_{1/2}$, $S_{\sigma,0}$, and α_σ. The free energy of unfolding at zero denaturant concentration was calculated from $\Delta G_U(W) = m_D[D]_{1/2}$.

Analysis of DSC Data. Deconvolution Analysis. Deconvolution of the heat capacity traces followed the formalism described elsewhere (*7*−*9*). In short, the excess enthalpy of a macromolecular system undergoing two-state unfolding,

[2] Abbreviations: ASA, solvent-accessible surface area; bp, base pair; CD, circular dichroism; DSC, differential scanning calorimetry; GdmCl, guanidinium hydrochloride; MD, molecular dynamics.

when defined as an "excess" over its native state, is

$$\langle \Delta H \rangle = f_U \Delta H_{cal} \quad (7)$$

where ΔH_{cal} is the calorimetric molar enthalpy of denaturation of the cooperative unit. The excess heat capacity is obtained by differentiation of eq 7:

$$\langle C_p \rangle = C_{p,N} + f_U(C_{p,U} - C_{p,N}) + \Delta H_{cal}\frac{\partial f_U}{\partial T} = C_{p,N} + f_U \Delta C_p + \Delta H_{cal}\frac{\partial f_U}{\partial T} \quad (8a)$$

ΔC_p is the heat capacity change of unfolding; $C_{p,U}$ and $C_{p,N}$ are the heat capacities of the unfolded and folded state, respectively. $C_{p,N}$ is a linear function of temperature of the form $a + b \times T$; $C_{p,U}$ can be approximated by a weak parabolic function, $c + d \times T + e \times T^2$. The progress of unfolding (fraction unfolded) is linked to the equilibrium constant through eq 2. After explicit differentiation of the last term on the right-hand side of eq 8a, the excess heat capacity is given by

$$\langle C_p \rangle = C_{p,N} + f_U \Delta C_p + \frac{\Delta H_{cal} \Delta H_{vH}}{RT^2} \times \frac{f_U(1 - f_U)}{n - f_U(n-1)} \quad (8b)$$

where ΔH_{vH} is the van't Hoff enthalpy of the process. If ΔH_{vH} is temperature-independent in the narrow temperature range where unfolding takes place, ΔH_{vH} equals $(-R\partial \ln K_U)/\partial(1/T)$ according to eq 3. For cooperative two-state unfolding $\Delta H_{cal} = \Delta H_{vH}$. The combined eqs 2, 3, and 8 were fitted to the experimental heat capacity data to calculate the thermodynamic unfolding parameters characterizing a cooperative unfolding unit. For the general case of i overlapping, formally independent unfolding processes the apparent heat capacity is expressed as

$$\langle C_p \rangle = C_{p,N} + \sum_i^k f_{U,i} \Delta C_{p,i} + \sum_i^k \Delta H_{cal,i} \frac{\partial f_{U,i}}{\partial T} \quad (8c)$$

The van't Hoff enthalpy was calculated from DSC data according to

$$\Delta H_{vH} = (\sqrt{n} + 1) T_{max} \sqrt{R\left(C_{p,max} - \frac{\Delta C_p \sqrt{n}}{\sqrt{n} + 1}\right)} \quad (9)$$

where $C_{p,max}$ is the height of the heat absorption peak (*7*).

The absolute heat capacity C_p was derived from the apparent heat capacity, C_{app}, according to (*10*)

$$C_{app} = (C_p - \upsilon_M) m_M + \text{const} \quad (10)$$

where υ_M is the partial specific volume of the macromolecule and m_M is its mass in the calorimetric cell.[3] The procedure requires the collection of thermograms at several different protein concentrations. The slope of plots of C_{app} vs m_M numerically equals $C_p - \upsilon_M$. υ_M values of the protein, 13 bp DNA duplex, and the complex were computed according to Karshikoff & Ladenstein (*11*). This method reproduces with high precision υ_M of globular proteins. Indeed, the calculated υ_M is within 2% of υ_M estimated from the amino acid sequence according to Makhatadze et al. (*12*) For the calculation, 20 NMR conformers of the complex and 25 NMR conformers of the protein were used and the DNA was modeled in an ideal B-form conformation. Values of υ_M used in eq 10 were 0.561 g cm^{-3} for DNA, 0.715 ± 0.006 g cm^{-3} for protein, and 0.659 ± 0.003 g cm^{-3} for the complex.

Calculation of Solvent-Accessible Surface Area and MD Simulations. ASA was calculated with the help of the program NACCESS using the default set of atomic radii and parameters (*13*). Calculation of ASA of the native protein was performed on the mean structure of the NMR ensemble representing the unbound form of the protein after minimization by standard methods using CHARMM. The unfolded state was modeled by the polypeptide in its fully extended conformation. The unfolding heat capacity was calculated using elementary contributions per Å^2 of polar, nonpolar and aromatic surface described in (*14*−*16*). Molecular dynamics simulations in explicit solvent, using Particle Mesh Ewald summations for long-range electrostatics, were performed for a simulation time of 2 ns (*17*). The program CHARMM with the Charmm27 parameter set was used (*18*, *19*).

RESULTS

The N-terminal 74 residues of the Tn916 integrase, also known as INT-DBD, are sufficient for sequence-specific DNA recognition. The solution structures of the N-terminal domain alone as well as complexed to a 13 bp target DNA duplex (sequences 5′-GAGTAGTAAATTC-3′ and 5′-GAATTTACTACTC-3′) have been solved by NMR spectroscopy (*3*, *4*). The DNA binding domain used in the present study comprised residues 2−74 of the integrase. The missing N-terminal Met does not participate in binding and has no effect on the stability of the complex (not shown). Experiments were performed in the same buffer used for NMR structure determination by Wojciak et al. (*3*).

In the first part of this Results Section, we describe isothermal unfolding of the protein induced by chemical denaturants. Thermal unfolding of the protein monitored by CD and fluorescence spectroscopy is described in the second part, and unfolding observed by differential scanning calorimetry in the third part. In the last part of Results, melting of the duplex DNA as monitored by CD spectroscopy and by DSC is presented.

Conformational Stability of Protein Deduced from Denaturant Unfolding. The protein contains a single tryptophan located at the DNA binding site. The fluorescence emission spectrum has a peak centered at 342 nm. Thermal and chemical denaturation shifts the emission maximum by 12−14 nm to the red and quenches emission (not shown). The shift of the emission maximum, λ_{max}, was used to follow denaturant induced unfolding (Figure 1). Unfolding by urea and GdmCl is well described by a two-state transition from

[3] Note that in the present form eq 10 formally violates the calculus-of-units rules. The reason is that the first term should actually reads $(C_p - \upsilon_M \rho_b C_{p,b}) m_M$, where ρ_b is the density of the solution and $C_{p,b}$ is the heat capacity of the solvent per unit mass. As explained in the original paper (*10*), for most experiments in dilute aqueous solutions, the product $\rho_b C_{p,b} \sim 1$ cal K^{-1} cm^{-3}, and the numerical value of the slope $\Delta C_p/\Delta m_M$ becomes $C_p - \upsilon_M$.

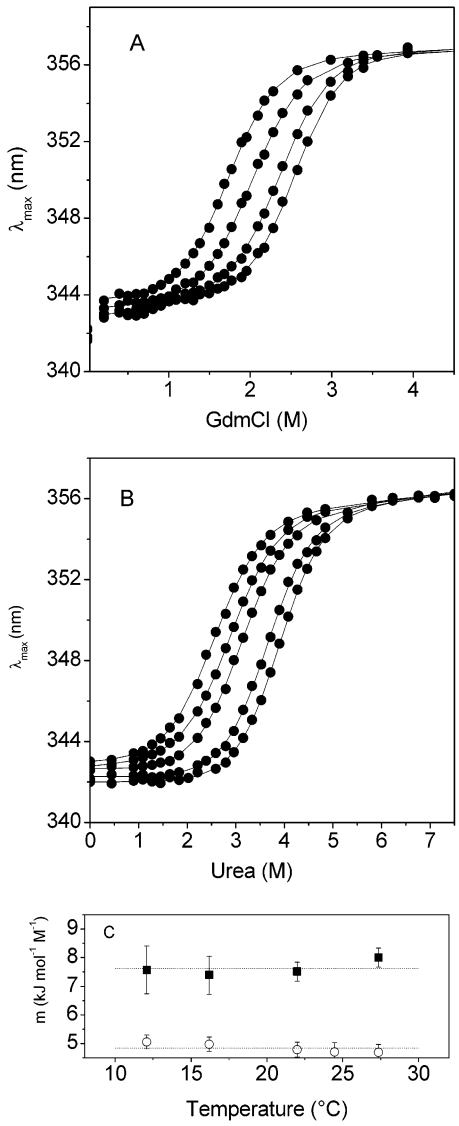

FIGURE 1: Protein unfolding induced by GdmCl (A) and urea (B) monitored by the shift of the tryptophan emission maximum. Symbols, experimental data; lines, best fits from nonlinear regression analysis according to eq 6. Unfolding experiments were performed between 27 and 12 °C (curves from left to right). (C) Slopes m_D from LEM analysis (eq 6) of GdmCl unfolding (filled symbols) and urea unfolding (open symbols). The dashed lines indicate the mean values.

the folded to the unfolded state (eq 6). The concentration of denaturant at the transition midpoint, $[D]_{1/2}$, progressively shifts to lower denaturant concentration as the temperature increases. ΔG_U was calculated by the linear extrapolation method assuming a linear dependence of ΔG_U on the denaturant concentration. Values of ΔG_U from urea and GdmCl unfolding are very similar, implying no significant influence of (solvent screened) charge−charge interaction to the stability of the protein at pH 6 (Table 1 of Supporting Information). Values of m_D for GdmCl unfolding are the same within error between 12 and 27 °C, whereas m_D from urea unfolding is slightly decreasing with increasing temperature (Figure 1C).

Thermal Unfolding of Protein Followed by CD and Fluorescence Spectroscopy. Thermal Unfolding Followed by CD Spectroscopy. Below 240 nm, the CD spectrum of the protein[4] indicates contributions from α-helix, β-sheet, turns, and loops scaled according to the relative content of these structure elements (*20*). In the near-UV region from 260 to 300 nm, there is a well-defined peak centered at 285 nm, which mirrors the asymmetric environment of tyrosine and aromatic residues (Figure 1 of Supporting Information). Thermal unfolding as well as denaturant unfolding decreases the CD signal maximally at 219 nm (change of secondary structure) and 285 nm (change of tertiary structure in the environment of aromatic residues). The two wavelengths were used to follow thermal unfolding (Figure 2). The shape of the melting curves conforms to one major structural transition taking place in a relatively narrow temperature range. Below 30 °C the signal changes linearly at both wavelengths. This behavior is typical for many proteins and probably results from gradual structural changes in the native state. Above 60 °C the signal at 219 nm displays a temperature slope of about 14 deg cm^2 dmol^{-1} K^{-1} while the signal at 285 nm remains constant. The spectral changes agree with a two-state unfolding mechanism (eqs 2 and 3). Values of T_m and ΔH_m deduced from the data at 219 and 285 nm are the same within error (Table 1). The apparent difference between the two nonnormalized melting curves of Figure 2 is due to the very different slopes before and after the main conformational transition. The difference disappears when the data are normalized to fraction of unfolded protein (see Figure 8A of the Discussion).

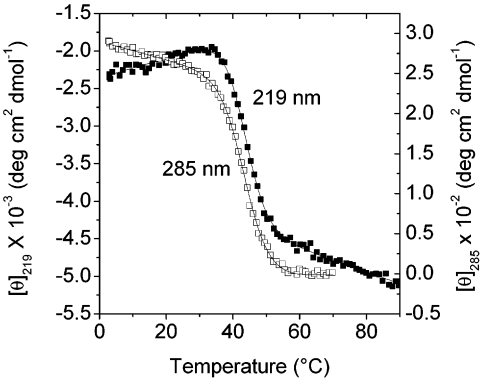

FIGURE 2: Thermal unfolding followed by circular dichroism. The data are presented as the change of the molar ellipticity per residue at 219 nm (filled symbols, left axis) and at 285 nm (open symbols, right axis). Lines are best fits for a two-state transition according to combined eqs 1−3.

Thermal Unfolding Followed by Fluorescence Spectroscopy. Figure 3 shows thermal melting curves represented as the shift of the fluorescence emission maximum after excitation at 295 nm. As for CD melting, the data agree with simple two-state transition from the folded to the unfolded state. Values of T_m and ΔH_m are identical within error to those deduced from the CD melting traces in Figure 2 (Table 1). Taken together, the close correspondence of the midpoint temperatures T_m and the van't Hoff enthalpies ΔH_m deduced from the CD changes at 219 and 285 nm and from the shift of the fluorescence emission maximum very strongly indicates that the disruption of secondary and tertiary structure elements are tightly coupled processes conforming to a simple two-state transition reaction.

Heat Capacity Change of Protein Unfolding. The heat capacity change of unfolding is defined as $\Delta C_p = \partial \Delta H_m /$

[4] See Figure 1 of Supporting Information.

Table 1: Thermodynamic Parameters for Thermal Protein Unfolding Measured by CD, Fluorescence and DSC[a]

	T_m (°C)	ΔH_m (kJ mol^{-1})	ΔC_p (kJ K^{-1} mol^{-1})
		CD Spectroscopy[b]	
	43.4 (219 nm)	258 ± 25 (219 nm)	
	43.7 (285 nm)	230 ± 25 (285 nm)	
		Fluorescence Spectroscopy[c]	
	44.2	275 ± 30	
		Differential Scanning Calorimetry[d]	
single transition	44.1	238 ± 12 (cal)	4.0 ± 0.6 (at 25 °C)
Figure 5B		248 ± 13 (vH)	3.7 ± 0.6 (at T_m)
		vH/cal = 1.04 ± 0.07	
single transition	43.7	255 ± 13 (cal)	4.9 ± 0.6 (at 25 °C)
Figure 5C		236 ± 12 (vH)	5.8 ± 0.6 (at T_m)
		vH/cal = 0.92 ± 0.07	
two transitions	34.8 (fit)	22 ± 4 (fit)	1.9 ± 0.1 (fit)
Figure 5C	44.1 (fit)	243 ± 12 (fit)	2.6 ± 0.5 (fit)
deconvolution analysis	28 ± 5 (fit)	30 ± 10 (fit)	
Figure 5D	43.7 (fit)	245 ± 19 (fit)	
Kirchhoff plot[e]			5.7 ± 0.4
mean value	43.8 ± 0.3 [f]	255 ± 18 [f,g]	5.0 ± 0.8 [g]

[a] Conditions: 50 mM Na−phosphate, 100 mM NaCl, pH 6.0. [b] Maximal possible error from systematic variation of the slopes in the pre-transitional and post-transitional regions of the melting curves. [c] SD from three experiments. [d] Errors are SD of seven experiments. [e] From Figure 4. [f] Mean value for main transition. [g] These values are used in the accompanying paper (*5*).

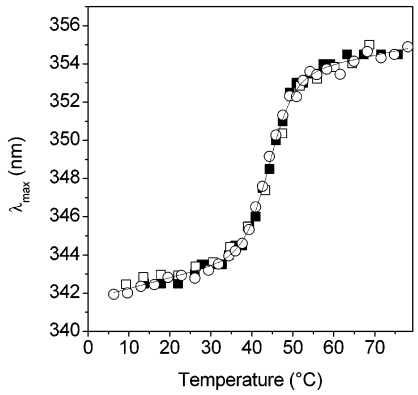

FIGURE 3: Thermal unfolding followed by fluorescence spectroscopy. The data are presented as the shift of the tryptophan fluorescence emission maximum. Different symbols refer to independent experiments. Lines are best fits for a two-state transition according to the combined eqs 1−3.

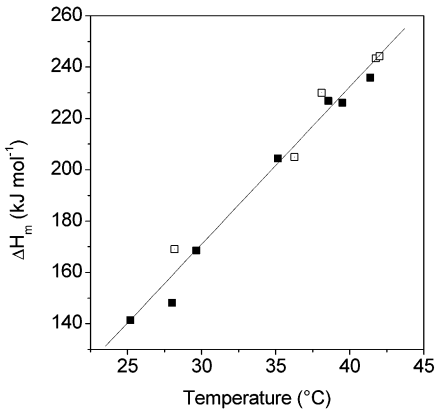

FIGURE 4: Determination of the heat capacity of unfolding, ΔC_p, from spectroscopic data. The slope is a linear best fit for ΔC_p of 5.7 ± 0.5 kJ K^{-1} mol^{-1}. Filled symbols, data collected in the presence of urea; open symbols, data collected in the presence of GdmCl; see the text for details and Table 2 of Supporting Information.

∂T_m. To estimate ΔC_P from the slope of a plot of ΔH_m against T_m (Kirchoff plot), a variation of of T_m and ΔH_m values in a broad temperature range had to be achieved. To this end, thermal unfolding experiments were repeated in the presence of small, nondenaturing amounts of GdmCl or urea. Figure 4 shows the Kirchoff plot obtained with the enlarged set of T_m and ΔH_m values (Table 2 of Supporting Information). The slope of the plot yields ΔC_p of 5.7 ± 0.5 kJ K^{-1} mol^{-1}.

The chosen procedure for obtaining ΔC_p may be flawed since the enthalpy and heat capacity of denaturant binding to proteins are negative and may affect the calculated thermodynamic parameters. The cosolutes may also influence the unfolding mechanism. However, three observations indicate that the procedure is applicable in the present case. First, thermal melting curves recorded in the presence of nondenaturing amounts of urea or GdmCl are monophasic and can be fit by the same two-state transition model (Figure 2 of Supporting Information). Second, the progress of unfolding (f_U) is identical at any given pair of temperature and denaturant concentration when monitored by two different experimental probes, fluorescence and far-UV CD spectroscopy (Figure 3 of Supporting Information). Third, ΔC_p calculated from the data of Figure 4 is in good agreement with ΔC_p deduced from the DSC experiments to be presented now.

Thermal Unfolding of Protein Followed by DSC. Partial Molar Heat Capacity. Figure 5A shows the temperature-dependence of the partial molar heat capacity of the protein. The transition maximum and the shape of the calorimetric trace are independent of the protein concentration and the rate of heating, in agreement with a reversible[5] monomolecular conformational transition. Hence, slow kinetic steps are not affecting the conformational transition. To eliminate possible complications caused by aggregation at high temperatures, DSC experiments were performed with low protein concentrations (typically 0.5 mg ml^{-1}) so that the absolute heat capacity C_p could not be measured precisely. Therefore, C_p was obtained from the change of the apparent heat capacity with protein concentration (eq 10).

[5] See Figure 4 of Supporting Information.

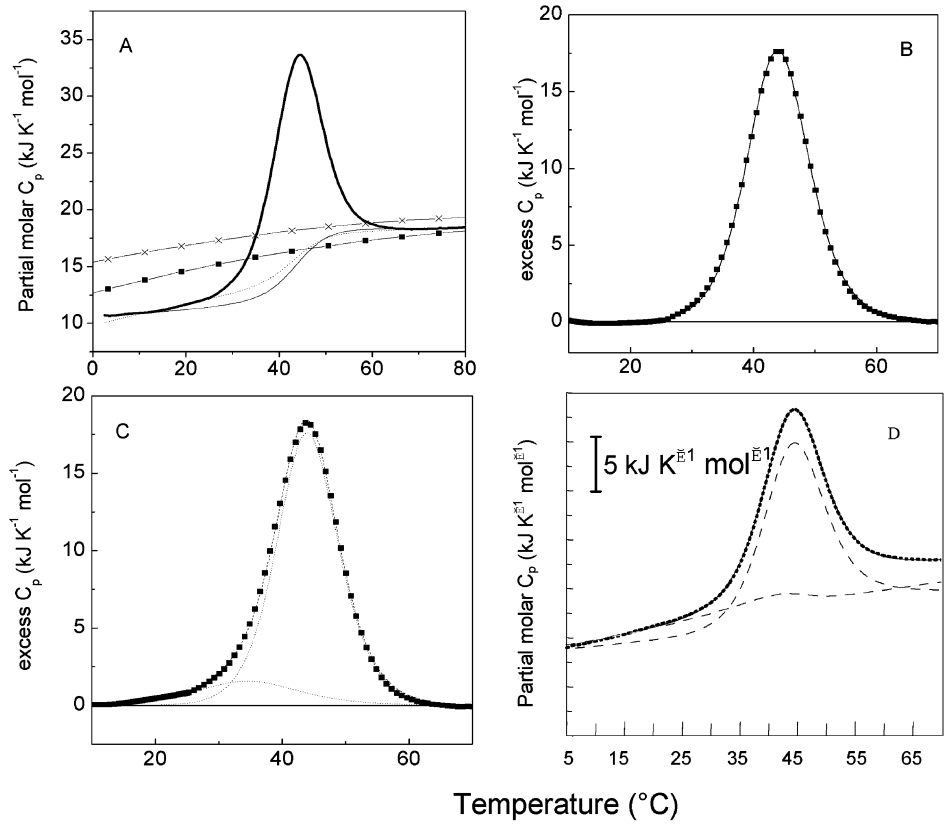

FIGURE 5: Thermal unfolding of the protein observed by DSC. (A) Temperature dependence of the partial molar heat capacity function (heavy line). The intrinsic heat capacity change is modeled in two ways: connection of the heat capacity of the folded state between 5 and 15 °C with the heat capacity of the unfolded state above 65 °C (thin solid line); connection of the heat capacity of the folded state between 15 and 25 °C with the heat capacity of the unfolded state above 65 °C (dotted line). The heat capacity of the denatured state calculated according to the HKF model (−■−) and the MP model (−×−) are also indicated. Experiments were performed with 65 μM protein in standard buffer of pH 6.0. (B) Excess heat capacity function calculated with the help of the dotted "baseline" of panel A. The experimental data (symbols) are best described by a single transition ($i = 1$ in eq 8c). (C) Excess heat capacity function calculated with the help of the solid "baseline" of panel A. The experimental data (symbols) are best described by two transitions ($i = 2$ in eq 8c). (D) Deconvolution of the apparent partial molar heat capacity. The two-component reactions corresponding to a minor and a major transition are indicated by broken lines. The experimental function (solid line, corresponding to the trace of panel A) and the function synthesized using the two component reactions (dotted line) match perfectly. See the text for detailed description of the different analyses of the DSC traces.

The value of C_p at 25 °C is 12.25 ± 0.50 kJ K^{-1} mol^{-1}, corresponding to 1.43 ± 0.12 J K^{-1} per g of protein. This is well within the range of specific heat capacities for other small globular proteins and equal to that of barnase (*15*, *21*). The initial slope of the pre-transitional portion of the specific heat capacity below 15 °C is small, 6−7 × 10^{-3} J K^{-2} g^{-1} but increases above ∼15 °C to 12.5 × 10^{-3} J K^{-2} g^{-1}. This increase may point to a small conformational transition preceding the main unfolding reaction, as will be discussed below.

The heat capacity of the denatured protein is constant above 65 °C and has a value of 18.9 ± 0.7 kJ K^{-1} mol^{-1} (2.13 ± 0.08 J K^{-1} g^{-1}). The heat capacity of a denatured globular protein can be calculated by adding up the partial molar heat capacity of the constituent groups of the protein. Makhatadze and Privalov modeled the heat capacity of the amino acids of a fully solvated and unstructured protein by using data on small organic model compounds (model MP, *22*). Another recent procedure uses data on side chain analogues, poly-glycines and denatured proteins (model HKF, *23*). The expected heat capacity of the infolded protein calculated by models MP and HFK, respectively, have been added to Figure 5A. The experimental heat capacity trace of the protein is enveloped by C_p calculated by models MP and HKF above 65 °C.

Heat Capacity Change from DSC Melting Curve. The difference between the partial heat capacity of the unfolded and the folded state represents the direct calorimetric estimate of the unfolding heat capacity change, ΔC_p^{cal}. Taking the mean of the MP and the HFK model calculations, ΔC_p at 25 °C is 3.9 kJ K^{-1} mol^{-1} (53 J K^{-1} mol residue^{-1}), and 4.1 kJ K^{-1} mol^{-1} at 10 °C (56 J K^{-1} mol residue^{-1}). Considering the uncertainties of the model calculations, the two values of ΔC_p^{cal} are in agreement with ΔC_p of 5.7 ± 0.5 kJ K^{-1} mol^{-1} obtained from the Kirchoff plot of Figure 4.

Evidence for a Minor Conformational Transition Preceding the Main Unfolding Reaction. Calculation of thermodynamic unfolding parameters from DSC data can be done in two ways. The principle assumption is that the observed temperature dependence of the molar heat capacity results from two contributions: the change of the heat capacity of the protein on going from the folded state to the unfolded state, called the "intrinsic" heat capacity change, and the "excess" heat capacity resulting from the enthalpy fluctuation in the transition zone (*7*, *8*). The intrinsic heat capacity is modeled by extrapolating the pre- and post-transitional heat capacities into the transition zone in proportion to the progress of the heat absorption. The resulting "baseline" is then subtracted from the experimental data to obtain the

remaining excess heat capacity peak (Figure 5B,C). This procedure is not trivial in the present case because the initial heat capacity of the folded protein increases nonlinearly before the main melting transition. In other words, we have to decide how to define the "baseline". If we take the heat capacity trace between 15 and 25 °C as the heat capacity of the unfolded protein, the "baseline" marked by the dotted sigmoidal line of Figure 5A results. Calculation of the excess heat capacity by subtracting this baseline from the main trace yields the trace shown in Figure 5B. Alternatively, we take as the heat capacity of the unfolded protein the trace between 5 and 15 °C and subtract the corresponding "baseline" (thin solid line of Figure 5A) from the main trace to obtain the excess heat capacity curve shown in Figure 5C. Though the traces of Figure 5B, C appear similar, close inspection reveals distinct differences. Analysis of the trace of Figure 5B fits perfectly to a single transition (best fit for $i = 1$ in eq 8c) and ΔH_{cal} of 238 ± 12 kJ mol^{-1}. Analysis of the trace of Figure 5C fits best to two transitions (best fit for $i = 2$ in eq 8c), as indicated by the dotted curves with midpoint temperatures of 34.8 °C and 44.1 °C, respectively. The enthalpy change ΔH_m^{cal} defined by the area under the peak of Figure 5C is 255 ±13 kJ mol^{-1}.

A second way to calculate the enthalpy of unfolding is by performing a van't Hoff analysis according to eq 9. In the case of a single transition between the folded and unfolded state, the ratio $\Delta H_m^{vH}/\Delta H_m^{cal}$ has to be 1. Indeed, $\Delta H_m^{vH}/\Delta H_m^{cal}$ is 1.04 ± 0.07 for the excess heat capacity trace of Figure 5B, in accord with a single transition. However, in the case of Figure 5C, $\Delta H_m^{vH}/\Delta H_m^{cal}$ is 0.92 ± 0.07. Although this ratio is perhaps not significantly below unity, it indicates that the unfolding reaction between 15 and 65 °C shown in Figure 5C is not perfectly two-state and is best described by a minor and a major transition.

A more rigorous analysis according to eq 8 makes no assumption about the "baseline" connecting the pre- and post-transitional heat capacity traces but uses a general value for the heat capacity function of a folded protein ($C_{p,N}$ of eq 8). The mean temperature dependence of $C_{p,N}$ of stable small globular proteins is 6.7 ± 1.0 × 10^{-3} J K^{-2} g^{-1} (set of 12 proteins; Table 3 in (*21*)). Analyzing the trace of Figure 5A with this value of $C_{p,N}$ (recalculated for the molecular mass of INT-DBD), one obtains the minimal model shown in Figure 5D, which accurately describes the experimental heat capacity trace by two transitions. The main transition has a T_m of 43.7 ± 0.3 °C and a ΔH_m of 245 ± 19 kJ mol^{-1}. The minor transition has a T_m of 28.1 ± 5.0 °C and a ΔH_m of 29.7 ± 10.4 kJ mol^{-1}. ΔC_p of the main transition is 2.6 ± 0.5 kJ K^{-1} mol^{-1} and of the minor transition 1.9 ± 0.1 kJ K^{-1} mol^{-1}.

Values of T_m, ΔH_m and ΔC_p obtained from the different analyses of the DSC traces in Figure 5 as well as from thermal unfolding monitored by CD and fluorescence spectroscopy in Figures 2 and 3 are put together in Table 1. The thermodynamic parameters deduced from independent calorimetric and spectroscopic experiments agree well. Thermal denaturation of the DNA-binding domain of integrase Tn916 follows the two-state mechanism of unfolding closely but not perfectly. There is firm calorimetric evidence for a temperature-induced, partial conformational change starting about 25 °C below the main cooperative unfolding reaction and proceeding with a relatively low heat absorption, yet with a large heat capacity increase. The mean value of ΔC_p of unfolding is 5.0 ± 0.8 kJ K^{-1} mol^{-1}. This value was used to analyze the thermodynamics of protein−DNA association described in the accompanying paper (*5*).

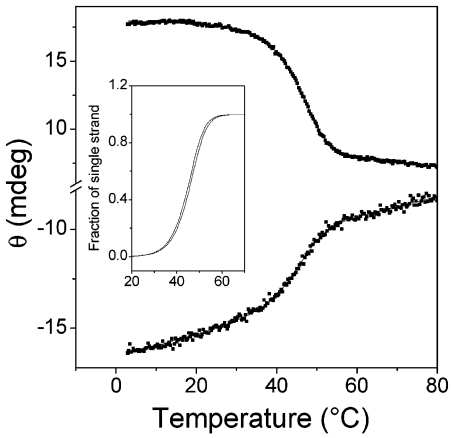

FIGURE 6: Thermal unfolding of the duplex DNA followed by CD. The change of ellipticity of a 65 μM DNA sample was monitored at 251 (lower trace) and 283 nm (upper trace). Continuous lines associated with the corresponding symbols are best nonlinear regression fits according to eqs 1−4. The inset shows the fraction of unfolded duplex calculated from the two experiments.

Conformational Stability of Target DNA Duplex. The temperature-induced unfolding and refolding of the 13 bp target DNA duplex was characterized by CD spectroscopy and DSC. The melting transition was highly reversible, as inferred from the reproducibility of both the heat capacity trace and the ellipticity changes upon cooling and reheating of the sample. The CD spectrum of the DNA is typical for short synthetic DNA duplexes and displays reversible temperature-induced changes (see Figure 2 in Milev et al. (2003) *Biochemistry 42*, 3481−3491). Figure 6 shows thermal melting of the DNA monitored at 251 and 283 nm where the spectral changes are maximal. The data are well described by the van't Hoff formalism for bimolecular two-state unfolding (eqs 1−4). The same enthalpy of unfolding is calculated from the 251 nm and the 283 nm melting trace.

Figure 7A presents the partial molar heat capacity function of DNA melting obtained by DSC. The heat capacity peak develops from a rather low temperature and is asymmetric. The latter feature is typical for unfolding linked to dissociation. Temperature-induced strand dissociation is also evident from the increase of T_m with increasing DNA concentration. Linear extrapolation of the initial and final heat capacities into the transition zone (dotted lines in Figure 7A) shows a very small difference at T_m between the heat capacity of the duplex and the sum of the heat capacities of the unfolded complementary strands. The mean apparent transition enthalpy obtained by integration of the four heat capacity peaks delimited by the dotted "baselines" is 332 ± 4 kJ mol^{-1} (ΔH_{peak} in Table 2). A very similar value is obtained from plotting $1/T_m$ versus the concentration of DNA duplex (Table 2).

Heat Capacity Change of DNA Dissociation and Unfolding. The heat capacity of the duplex increases steeply below the onset of the main transition. The average initial slope is 0.192 ± 0.017 kJ K^{-2} mol^{-1}, corresponding to 25 × 10^{-3} J K^{-2} per g of DNA. The latter value is four times higher than

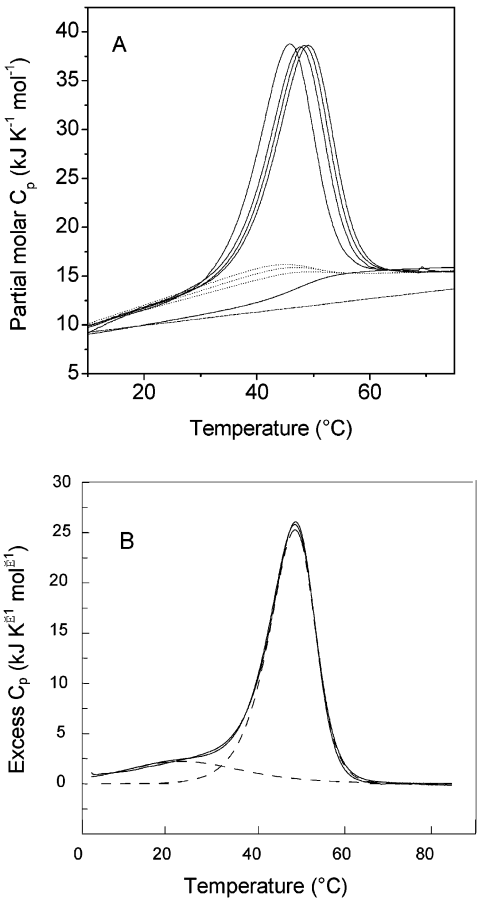

FIGURE 7: Thermal unfolding of the duplex DNA followed by DSC. (A) Melting of 33, 65, 90 and 119 μM duplex DNA (left to right) in standard buffer and at a scanning rate of 1 deg min^{-1}. Melting temperatures are 45.8, 47.6, 48.3, and 49.1 °C, respectively. Dotted lines connecting the pre-transitional and post-transitional traces model the apparent intrinsic heat capacity change. The calculated heat capacity of the folded duplex is indicated by the dash−dotted line (see text about how this line was constructed). (B) Excess heat capacity function calculated from the trace for 65 μM duplex DNA and the thin solid "baseline" of panel A. The low- and high-temperature components indicated by dashed lines correspond, respectively, to a broad pre-transitional accumulation of thermal energy and to cooperative melting accompanied by strand dissociation. Addition of the two components (dotted line) perfectly matches the experimental data (symbols).

Table 2: Enthalpy (kJ mol^{-1}) of Dissociation and Unfolding of the 13 bp DNA Duplex Measured by DSC and by CD[a]

	DSC[b]		CD[c]	
ΔH_{grad}[d]	ΔH_{peak}[e]	ΔH_{peak}[f]	ΔH_{vH}[g]	ΔH_{vH}
74 ± 2	332 ± 4	341	335 ± 2	334

[a] Conditions: 50 mM Na−phosphate, 100 mM NaCl, pH 6.0. [b] Mean of four experiments performed with 33, 65, 90, and 119 μM duplex DNA. [c] From melting monitored by CD, Figure 6. [d] Enthalpy accumulated as a result of the gradual increase of the heat capacity below the main cooperative transition, as obtained by deconvolution analysis of the excess heat capacity function of Figure 7B. [e] Enthalpy of the main cooperative dissociation and unfolding step in Figure 7A. [f] Calculated from the slope of a plot of $1/T_m$ against the duplex DNA concentration (Figure 5 of Supplementary Information). [g] From the analysis of the shape of the heat absorption peak according to eq 9.

the mean value for stable globular proteins of 6×10^{-3} J K^{-2} g^{-1} (*7*). Therefore, the experimentally observed pre-transitional heat capacity may not reflect the intrinsic heat capacity function of the fully folded DNA duplex. If so, the linear extrapolation of the pre-transitional heat capacity into the melting zone cannot be used to estimate the heat capacity change characterizing unfolding and dissociation of the DNA duplex between low temperature and high temperature, say, between 5 and 70 °C. This peculiarity of DNA melting is well documented for short duplexes (*24*, *25*); hence, a good value of ΔC_p is difficult to gain from DSC. An average ΔC_p value of 0.270 ± 0.089 kJ K^{-1} (mol bp)$^{-1}$ for DNA unfolding and dissociation was reported based on precise calorimetric analysis of polynucleotides (*26*); a value of 0.196 kJ K^{-1} (mol bp)$^{-1}$ was obtained from Kirchoff's plots of ΔH_m versus T_m from DNA melting data (*26*, *27*). We have combined the data published in references (*24*, *25*, *28*) and have corrected them for different GC content (Figure 6 of Supporting Information). From this large data set we calculate ΔC_p of 0.21 ± 0.07 kJ K^{-1} (mol bp)$^{-1}$, corresponding to ΔC_p of 2.7 ± 0.9 kJ K^{-1} mol^{-1} for unfolding and dissociation of the 13 bp target DNA duplex. Subtracting 2.7 kJ K^{-1} mol^{-1} from the heat capacity at 65 °C in Figure 7A one obtains the heat capacity of the folded duplex at that temperature. The straight line connecting this point with the heat capacity at the lowest temperature where the duplex is fully folded yields, as a first approximation, the intrinsic temperature dependence of the heat capacity of the folded state (dash−dotted line in Figure 7A). The slope of this function is 69 ± 13 J K^{-2} mol^{-1} or (8.8 ± 1.7) × 10^{-3} J K^{-2} per g of DNA, in good agreement with (10 ± 2) × 10^{-3} J K^{-2} g^{-1} found for duplexes of 10, 12, and 16 base pairs (*25*).

With the help of the duplex heat capacity function so constructed, we can calculate the "baseline" linking the low and the high temperature limbs of the DSC trace (thin sigmoidal line in Figure 7A) and construct the excess heat capacity function shown in Figure 7B. Deconvolution analysis of the excess heat capacity function with the help of eq 8c yields two different, partially overlapping phases: an initial gradual heat absorption, followed by a phase of intense heat absorption associated with the disruption of bulk packing interactions and bimolecular strand dissociation. The first phase is centered around 20 °C and proceeds with absorption of 15−20% of the total heat. The enthalpies of the two processes are listed in Table 2.

Summarizing the analysis of the DNA duplex melting, we note that the enthalpy change of the main transition (ΔH_{peak} of Table 2) is well documented by independent calorimetric and spectroscopic experiments. As is typical of short DNA duplexes, there is a significant premelting phase (ΔH_{grad} of Table 2). Because of this gradual heat absorption below the main transition, ΔC_p could not be measured directly. On the basis of published results for many different DNA duplexes, we calculate ΔC_p of 2.7 ± 0.9 kJ K^{-1} mol^{-1} for the 13 bp duplex DNA. This value was used to analyze the thermodynamics of protein-DNA association described in the accompanying paper (*5*).

DISCUSSION

Conformational Stability of the DNA Binding Domain of Integrase Tn916. The DNA binding domain follows closely but not perfectly the two-state model of unfolding. Figure 8A shows that the heat-induced shift from the native to the unfolded state exhibits the same temperature dependence

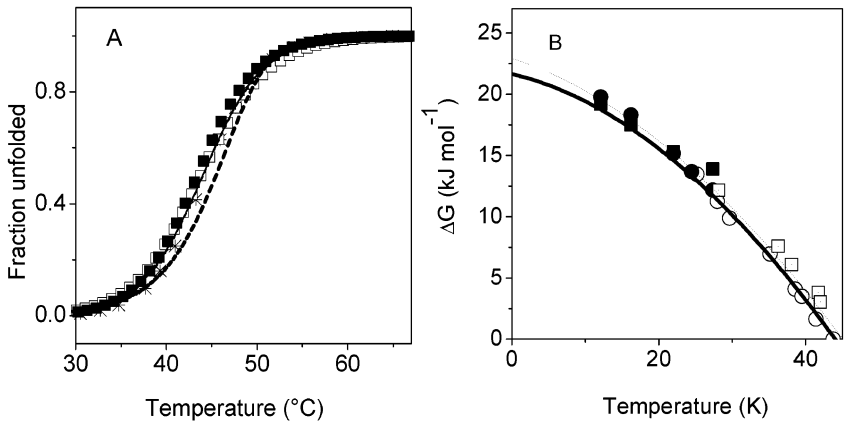

FIGURE 8: Thermal stability of the integrase DNA binding domain. (A) Fraction of unfolded protein as a function of temperature. Symbols refer to thermal unfolding followed by fluorescence (asterisks), CD signal at 219 nm (filled squares), and CD signal at 285 nm (open squares). Solid line, progress of unfolding calculated from DSC data in Figure 7C, including two transitions. Broken line, progress of unfolding calculated from DSC data in Figure 7B, including only a single transition. (B) Thermal stability curves. Symbols: isothermal unfolding by urea (closed circles), isothermal unfolding by GdmCl (filled squares), thermal unfolding in the presence of nondenaturing amounts of urea (open circles, Table 2 of Supporting Information), and thermal unfolding in the presence of nondenaturing amounts of GdmCl (open squares, Table 2 of Supporting Information). The lines illustrate best results from nonlinear least-squares regression analysis according to the Gibbs−Helmholtz eq 5. Heavy line, simulation based on the mean values measured by DSC (Table 1): $T_m = 43.9$ °C, $\Delta H_m = 250$ kJ mol^{-1}, $\Delta C_p = 4.6$ kJ K^{-1} mol^{-1}. Dotte line, combined urea, and GdmCl unfolding data: $T_m = 45$ °C, $\Delta H_m = 267$ kJ mol^{-1}, $\Delta C_p = 4.3$ kJ K^{-1} mol^{-1}.

when monitored by the progress of heat uptake (DSC), the melting of secondary structure elements (CD signal at 219 nm), and the disruption of tertiary packing interactions (CD signal at 285 nm and tryptophan fluorescence emission). Good correspondence between DSC-derived and spectroscopy-derived parameters is achieved if a minor pre-transition with T_m around 30 °C is considered, but not if only the main transition is considered (compare solid with dashed line in Figure 8A). Combination of data from denaturant-induced unfolding and from unfolding by heat give a consistent picture of the energetics of the unfolding process. Figure 8B shows the protein stability curve constructed with the help of eq 5.

The thermodynamic parameters of unfolding are within the range of values for stable globular proteins (*15*). At 43 °C the specific enthalpy of 3.5 ± 0.2 kJ (mol residue)$^{-1}$ and the specific entropy of 11.0 ± 0.7 J K^{-1} (mol residue)$^{-1}$ are among the highest reported for proteins (Table 3 of ref *15*). This applies also to lower temperatures since the specific heat capacity change is similar as for other proteins, 63 ± 11 J K^{-1} (mol residue)$^{-1}$. The unfolding free energy is rather low, about 14 kJ mol^{-1} at 25 °C (Figure 8B), and is comparable to ΔG_U of the SH3 spectrin domain (*29*). It appears that strong packing interactions are largely compensated by unfavorable entropic factors so that the stability of the DNA binding domain is low. Whether low stability is an intrinsic feature of the DNA binding domain of Tn916 integrase or only a feature of the isolated 74-residue DNA binding domain is not known. Interestingly, according to NMR data, the DNA-bound protein is more unstructured than the free protein, which is rather unusual. Low stability implies increased flexibility and facilitates local unfolding, hence, the low stability of the folded structure may have biological meaning (*30*, *31*).

The m_D value for denaturant-induced unfolding at constant temperature as well as the heat capacity change of unfolding are thought to correlate with the surface of protein exposed to solvent upon unfolding (*14*, *15*, *32*). Even if this correlation has no rigorous thermodynamic foundation, it is a useful operational criterion to judge whether unfolding proceeds with a typical exposure of molecular surface. Assuming that the 74-residue DNA binding domain unfolds to a fully extended peptide chain, we calculate m_D values of 4.12 kJ mol^{-1} M^{-1} for urea-induced unfolding and 8.72 kJ mol^{-1} M^{-1} from the amount of exposed surface (*33*). The predictions correlate well with the mean m_D-values of 4.84 ± 0.33 and 7.62 ± 0.84 kJ mol^{-1} M^{-1} of Figure 1C. Also, the mean value of ΔC_p of 5.0 ± 0.8 kJ K^{-1} mol^{-1} (Table 1) agrees well with ΔC_p of 5.5 ± 0.8 kJ K^{-1} mol^{-1} calculated from the change of surface exposure according to different parametrization schemes (*14*−*16*). One may therefore conclude that chemical and thermal denaturation of the DNA binding domain leads to a highly solvated state lacking significant intermolecular contacts. This is further supported by the fact that the heat capacity of the protein in the unfolded state matches the heat capacity predicted for a fully solvated polypeptide chain (Figure 5A).

By inference it follows that the folded DNA binding domain has a compactly folded structure, which is confirmed by packing density calculations performed on 25 NMR conformers. The calculated volume of voids and cavities of 2150 ± 60 and 13 ± 10 Å^3, respectively, are typical of a compact protein (*11*). The specific void volume per atom is 1.80 ± 0.05 Å^3, similar to the median value for 35 proteins of molecular weight below 14 kDa (*11*). The protein is also compact according to calorimetric criteria since the partial specific heat capacity at 25 °C resembles a typical globular domain (*21*). However, starting from ∼15 °C, C_p increases rapidly, indicating heat absorption by a temperature-dependent process. The changes in the environment of aromatic residues and the overall secondary structure also show temperature dependence below 30 °C. Still, the heat capacity increase is not exceptionally steep and does not indicate substantial structural rearrangements and increasing surface exposure between 2 and 30 °C. In the NMR structure taken at 27 °C, close to the onset of the main heat absorption peak,

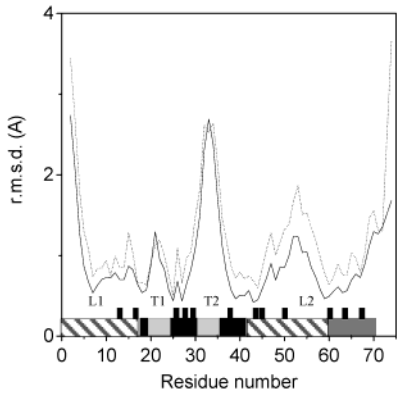

FIGURE 9: Flexibility of the integrase DNA-binding domain from MD simulations. Root-mean-square deviations (rms) of C_α atoms from the average trajectory structure as a function of residue number. Continuous line, simulation at 278 K; broken line, simulation at 298 K. The secondary structure elements are indicated as follows: hatched, loops; black, β-strands; light gray, turns; dark gray, α-helix; upper black bars, residues contributing to the hydrophobic core. For a model of the protein, see Figure 1 of the accompanying paper (*5*).

the secondary structure elements are intact, and the extended loops L1 and L2 are well defined and are anchored to the body of the protein, contributing to the hydrophobic core (see Figure 1 of accompanying paper, *5*). DNA binding affinity is preserved up to 30 °C (*5*). Most likely, the initial slope of the heat capacity function reflects gradual accumulation of thermal motions.

In an attempt to gain information about the flexibility of the protein, we performed 2 ns MD simulations in explicit water. The results are shown in Figure 9. Excluding the chain termini, the highest fluctuations are seen in turns T1 and T2 and loop L2. Turn T2 is pulled away from the α-helix by ~4.5 Å without disrupting of the local turn conformation yet causing destabilization of the α-helix, which is shortened by 2 residues (*3*). The chemical shifts of Lys 54 and Arg 55 of loop L2 change on addition of DNA. Turn T1 is also pulled by ~1.8 Å toward the DNA. There seems to be some built-in flexibility in these regions of the protein. Whether such modular motion takes place in the absence of DNA is not known. On average, the protein is fluctuating more at higher temperature, particularly in its C-terminal half (Figure 9). Taken together, MD simulations make it likely that the increase in heat capacity above 15 °C reflects partial destabilization and increased thermal motions, perhaps preferentially in the C-terminal half.

Conformational Stability of Target DNA Duplex. Current knowledge about the thermodynamic stability of DNA lags behind our much deeper understanding of protein stability for reasons that have been discussed (*24*−*26*). The heat capacity change associated with DNA dissociation and unfolding has long been believed to be negligible because the heat capacities measured immediately before and after the main heat absorption peak of a thermogram are similar. On the other hand, Kirchoff analysis shows that the melting enthalpy is temperature-dependent (*27*). Large ΔC_p values have been measured by isothermal titration calorimetry (*24*, *25*, *34*). A more consistent picture of duplex melting has emerged only recently (*26*). Isolated DNA strands form temperature-sensitive intramolecular structures, which are disrupted upon duplex formation. The heat capacity of short duplexes increases steeply with temperature and becomes approximately equal to the summed heat capacities of the separated strands. As suggested previously, this reflects temperature-dependent noncooperative structural fluctuations of the duplex, e.g., end fraying and twisting as also seen in the present study (Figure 7). Because of this difficulty, we have decided to use a calculated value of 2.7 ± 0.9 kJ K^{-1} mol^{-1} for ΔC_p of our DNA duplex. The calculation is based on a large set of data for different oligonucleotides and considers the GC content. The validity of the calculated heat capacity change is supported by three observations. First, the heat capacity function of the folded duplex derived with the help of ΔC_p of 2.7 kJ K^{-1} mol^{-1} (dash−dotted line of Figure 7A) is identical with the heat capacity function calculated for several other short duplexes (*25*). Second, the enthalpy of the main dissociation and unfolding reaction (ΔH_{peak} of Table 2) changes by less than 10% when ΔC_p is varied within its limit of uncertainty of $\pm$ 0.9 kJ K^{-1} mol^{-1}. Third, extrapolating the total melting enthalpy back to room temperature with the help of ΔC_p of 2.7 kJ K^{-1} mol^{-1}, one obtains ΔH, ΔS and ΔG of duplex formation at 25 °C of 27 kJ $(mol\ bp)^{-1}$, 76 J K^{-1} $(mol\ bp)^{-1}$ and 4.1 kJ $(mol\ bp)^{-1}$, respectively. These values are typical for natural and synthetic DNA molecules. The dissociation constant of the 13 bp duplex DNA increases from 1×10^{-13} M at 5 °C to 1×10^{-8} M at 30 °C. At 30 °C and 30 μM concentration, the duplex is still more than 98% populated.

CONCLUSIONS

The main purpose of the present thermodynamic analysis was to aid in the interpretation of the association reaction between the Tn916 integrase and its cognate target DNA. For one, it was necessary to know the thermal stability and conformational state of the associating molecules in the range of 5−30 °C in which calorimetric and spectroscopic protein-DNA binding experiments were performed (*5*). The analysis shows that the protein and the short target DNA duplex are folded up to 30 °C. Second, the heat capacity profile of the protein and the DNA had to be clarified in order to interpret the heat capacity profile of the protein−DNA complex (*5*). The present DSC experiments demonstrate that the protein accumulates temperature-sensitive structural fluctuations below the onset of the cooperative unfolding transition. The gradual, nonlinear structural changes of the protein leading to a minor transition around 30 °C could be noticed only from the heat capacity trace but not from the spectroscopic data. The thermal features of the protein, the steep pretransitional heat capacity increase of the target DNA, the melting enthalpies of both molecules and their heat capacity changes of unfolding are indispensable to understand the energetics of the protein-DNA association reaction. Apart from its particular purpose, the present work also reveals temperature dependent conformational dynamics and flexibility in a DNA binding domain. These features may be of general importance to sequence-specific DNA recognition in view of the fact that protein−DNA association reactions are often accompanied by remarkable structural rearrangements.

ACKNOWLEDGMENT

We thank Dr. Peter Gehrig for mass spectroscopy analysis.

SUPPORTING INFORMATION AVAILABLE

Table with thermodynamic parameters for denaturant-induced unfolding; figure and table with data for thermal unfolding in the presence of nondenaturing amounts of urea and GdmCl; figure showing the CD spectrum of the protein; figure showing fraction of unfolded protein as a function of temperature and urea concentration; figure showing raw data of DSC experiment; figure showing plot of $1/T_m$ versus natural logarithm of duplex concentration; figure showing ΔC_p for 23 DNA duplexes normalized for their GC content. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Scott, J. R., and Churchward, G. G. (1995) *Annu. Rev. Microbiol. 49*, 367−397.
2. Landy, A. (1989) *Annu. Rev. Biochem. 58*, 913−949.
3. Wojciak, J. M., Connolly, K. M., and Clubb, R. T. (1999) *Nat. Struct. Biol. 6*, 366−373.
4. Connolly, K. M., Wojciak, J. M., and Clubb, R. T. (1998) *Nat. Struct. Biol. 5*, 546−550.
5. Milev, S., Gorfe, A. A., Karshikoff, A., Clubb, R. T., Bosshard, H. R., and Jelesarov, I. (2003) *Biochemistry 42*, 3481−3491.
6. Marky, L. A., and Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.
7. Privalov, P. L., and Potekhin, S. A. (1986) *Methods Enzymol. 131*, 4−51.
8. Freire, E., and Biltonen, R. L. (1978) *Biopolymers 17*, 463−479.
9. Freire, E. (1995) *Methods Enzymol. 259*, 144−169.
10. Kholodenko, V., and Freire, E. (1991) *Anal. Biochem. 270*, 336−338.
11. Karshikoff, A., and Ladenstein, R. (1998) *Protein Eng. 11*, 867−872.
12. Makhatadze, G. I., Medvedkin, V. N., and Privalov, P. L. (1990) *Biopolymers 30*, 1001−1010.
13. Hubbard, S. J., and Thornton, J. M. (1993). *NACCESS*, Computer Program, Department of Biochemistry and Molecular Biology, University College, London.
14. Murphy, K. P., and Freire, E. (1992) *Adv. Protein Chem. 43*, 313−361.
15. Makhatadze, G. I., and Privalov, P. L. (1995) *Adv. Protein Chem. 47*, 307−425.
16. Spolar, R. S., Livingstone, J. R., and Record, M. T., Jr. (1992) *Biochemistry 31*, 3947−55.
17. Darden, T. A., York, D. M., and Pedersen, L. J. (1993) *J. Chem. Phys. 98*, 10089−10092.
18. MacKerell, J., A. D.;, Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K. M., C.; Michnick, S.; Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, I., W. E.;, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. A.-a. (1998) *J. Phys. Chem. B 102*, 3586−3616.
19. Brooks, B. R., Bruccorleri, R. E., Olafson, B. D., David, J. S., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
20. Perczel, A., Park, K., and Fasman, G. D. (1992) *Proteins: Struct. Funct. Genet. 13*, 57−69.
21. Gomez, J., Hilser, V. J., Xie, D., and Freire, E. (1995) *Proteins 22*, 404−412.
22. Privalov, P. L., and Makhatadze, G. I. (1990) *J. Mol. Biol. 213*, 385−391.
23. Amend, J. P., and Helgeson, H. C. (2000) *Biophys. Chem. 84*, 105−136.
24. Holbrook, J. A., Capp, M. W., Saecker, R. M., and Record, M. T., Jr. (1999) *Biochemistry 38*, 8409−8422.
25. Jelesarov, I., Crane-Robinson, C., and Privalov, P. L. (1999) *J. Mol. Biol. 294*, 981−995.
26. Chalikian, T. V., Volker, J., Plum, G. E., and Breslauer, K. J. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 7853−7858.
27. Klump, H. (1990) in *Landolt Börnstein Neue Serie Group VII* (Saenger, W., Ed.) pp 241−256, Biophysik Springer, Berlin.
28. Rouzina, I., and Bloomfield, V. A. (1999) *Biophys. J. 77*, 3242−3251.
29. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L. (1994) *Biochemistry 33*, 2142−2150.
30. Battiste, J. L., Li, R. H., and Woodward, C. (2002) *Biochemistry 41*, 2237−2245.
31. Zavodszky, P., Kardos, J., Svingor, and Petsko, G. A. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 7406−7411.
32. Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
33. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−48.
34. Ladbury, J. E., Sturtevant, J. M., and Leontis, N. B. (1994) *Biochemistry 33*, 6828−6833.

BI026936X