The DNA-binding domain of *Drosophila melanogaster* c-Myb undergoes a multistate denaturation

Anup MADAN, P. K. RADHA, Arvind SRIVASTAVA, Lakshmi C. PADHY and Ramakrishna V. HOSUR

Tata Institute of Fundamental Research, Bombay, India

(Received 18 December 1994/14 February 1995) – EJB 94 1895/2

The DNA-binding domain of *Drosophila* c-Myb protein has been studied using different spectroscopic probes, namely CD, fluorescence, acrylamide quenching and NMR, to determine the structure of some of its sub-domains and their relative stabilities in aqueous solutions. While CD and fluorescence spectroscopy showed that the protein had completely lost its tertiary and secondary structures in approximately 3 M urea, solvent accessibility of the tryptophan residues was still partial, as determined by acrylamide quenching. This suggested the presence of significant amounts of residual structure which persisted until the urea concentration was raised to approximately 6.0 M. Thermal-denaturation experiments also indicated the presence of an intermediate in the unfolding pathway. The experimental data could be fitted assuming a minimum of three states in both modes of denaturation. The thermodynamic parameters for the apparent three-state transition have been determined. From the protein stability curve, we have determined that *Drosophila melanogaster* Myb R123 has maximal stability at 16°C and pH 7.0.

**Keywords.** Myb denaturation; Myb DNA-binding domain; spectroscopic investigations; protein unfolding.

The DNA-binding domain of several myb family proteins [1, 2] are made up of three apparent repeating units (R1-R2-R3) of approximately 51–53 amino acid residues each. In each of the repeats, three conserved tryptophan residues are placed at intervals of 18/19 amino acids [3–5]. The c-myb protein binds to DNA in a sequence-specific manner [6] and various models have been proposed based upon results from Raman spectroscopy [7], site-directed mutagenesis [8] and NMR and molecular dynamics [9, 10]. Out of the three conserved stretches (R1, R2 and R3), R2-R3 has been shown to be critical for DNA binding in the chicken c-Myb protein. Recently, the secondary structures of the R3 repeat of mouse [9] and R2-R3 of chicken [10] have been determined by NMR. The results show the presence of a helix-turn-helix motif, common in many other DNA-binding proteins [11]. We have recently over-expressed the DNA-binding domain (R123) of *Drosophila melanogaster* in *Escherichia coli*. Our experimental results show that this domain forms a stable autonomously folding structural and functional unit. The protein has been found to be largely α helical, the helical content being 58.6%. Out of the nine tryptophan residues, three are solvent accessible (Madan, A., Radha, P. K., Hosur, R. V. and Padhy, L. C., unpublished results).

It has been visualized that, in proteins, the transition from the native state to the denatured state may progressively pass through states with increasing disorder [12]. Characterization of these states requires trapping them [13, 14] as they may appear transiently during the denaturation [15, 16]. However, in some proteins under conditions of partial denaturation, stable states have been shown to exist in an equilibrium [17–24]. In the recent past, there has been an increased interest in these less-ordered states, since they may play important roles in protein folding and stability, transport of proteins across membranes, proteolysis, protein modifications and turn over [23].

We present in this paper the data on the bacterially expressed DNA-binding domain of *Drosophila* c-Myb (Dm-Myb) to show that its unfolding process is more complex than a simple two-state mechanism and that one can detect significant amounts of residual structure under an apparent denatured state of the protein.

**MATERIALS AND METHODS**

**Materials.** Molecular-mass markers, ammonium sulfate, polyethyleneimine, isopropyl β-D-thiogalactopyranoside, 6-(p-toluidino)-2-naphthalene sulfonate (TNS), spectroscopic-grade dimethylsulfoxide, acrylamide, SDS, SP-Sephadex, Mops, Hepes and Mes were purchased from Sigma. DE 52 was from Whatmann and culture media were from Difco. Ultra-pure urea for denaturation experiments was obtained from Schwarz and Mann. Spectroscopic-grade ethylene glycol was obtained from Fluka. All other reagents were of analytical grade. TNS was dissolved in N,N-dimethyl formamide and the concentration of the stock solution was determined to be 170 mM by absorbance measurement using a molar absorption coefficient of 1.89×10⁴ at 317 nm.

**Protein preparation.** The DNA-binding domain of the Dm-Myb was over expressed in *E. coli* and purified as described (Madan, A., Radha, P. K., Hosur, R. V. and Padhy, L. C., unpublished results). The purity of the protein was checked elsewhere by 15% SDS/PAGE [26] and was found to be greater than 95%. Protein concentration was determined by UV absorption, as suggested recently [27].

**Protein unfolding.** CD, fluorescence and NMR spectroscopy were used for monitoring the protein unfolding process.
Fluorescence studies. Fluorescence measurements were made on a Shimadzu RF-540 spectrofluorimeter. The band pass of both the excitation and emission monochromators was 5 nm. All measurements were carried out using 1-cm path-length cuvettes. For collection of the protein fluorescence spectra, the samples were excited at 295 nm and the spectra were corrected for the response of the buffer alone. For denaturation experiments, the fluorescence intensity at 340 nm was monitored as a function of the urea concentration.

The fraction of the tryptophan residues exposed as the protein unfolds was studied using the neutral quencher acrylamide. The changes in the protein fluorescence due to acrylamide quenching at 340 nm at various urea concentrations were monitored, and the data was analyzed according to the Lehrer equation [28]:

$$\frac{F_o}{F_o - F} = \frac{1}{f_o k_{q\infty}[Q]} + \frac{1}{f_o},$$

(1)

where $F_o$ and $F$ are the fluorescence intensities of the protein in the absence and presence of the quencher, respectively. $f_o$ is the fraction of quenchable fluorescence, $k_{q\infty}$ is the rate constant for quenching, $t_o$ is the fluorescence lifetime of the fluorophores in the absence of the quencher.

CD spectroscopy. CD measurements were made on a Jasco J600 spectropolarimeter. Denaturation was monitored by changes in the intensity of the CD band at 222 nm. For each spectrum, ten scans were recorded and averaged. Each spectrum was recorded in 0.2-nm wavelength increment and the signal was acquired for 1.0 s at each wavelength. For thermal denaturation, thermostatically controlled cuvettes were used and the temperature was maintained with the help of a LKB Multitemp II thermostatic circulator.

NMR spectroscopy. NMR spectra were recorded on a Bruker AMX 500 spectrometer operating at 500 MHz for $^1$H. The samples consisted of 5 mM protein in H$_2$O at pH 4.8. Two-dimensional clean total correlated [34] spectra were recorded with 2048 $t_2$ points for each of 512 $t_1$ points, where $t_1$ and $t_2$ are the usual time variables of two-dimensional spectroscopy. The water signal was partially suppressed by presaturation. The spin-lock mixing time was 80 ms in all the cases.

RESULTS

We have used a variety of spectroscopic techniques and strategies to derive information on the unfolding process of Dm-Myb R123. Fig. 1 shows the CD and fluorescence spectra of the native and denatured proteins. The CD spectrum (Fig. 1a) of the native protein shows two bands at 208 nm and 222 nm, characteristic of a helical structure. Urea denaturation results in significant changes in the CD spectrum. Similarly, in the fluorescence emission spectra (Fig. 1b), urea denaturation is accompanied by
Madan et al. (Eur. J. Biochem. 230) 735

Fig. 2. Urea denaturation of Dm-Myb R123. (a) Denaturation profile of 90 μM Dm-Myb R123 in 10 mM Tris/HCl, pH 7.0, with increasing urea concentrations, monitored by changes in ellipticity (Θ) at 222 nm or in fluorescence intensity (●) at 340 nm. Samples were incubated at the desired urea concentration for 30 min prior to measurement. The apparent fractional change (Fapp) at various urea concentration was calculated according to Eqn (2). (b) Fractions of the native (●), intermediate (Θ) and unfolded forms (Δ) of the Dm-Myb R123 as a function of the urea concentration at pH 7.0. The various fractions were calculated from the equilibrium constants determined by the fitting of the urea denaturation curve by a three-state model. (c) Fitting of the transition region of the curve in (a) to a three-state denaturation model. The calculated curve is shown by a solid line; (●) represents the experimental data points.

Table 1. Parameters for the three-state fits. All parameters were determined from three independent experiments employing protein concentrations of 90, 105 and 125 μM, respectively. CD, fluorescence and TNS fluorescence parameters were determined from the urea denaturation curve. The temperature denaturation monitored by CD spectroscopy was at pH 7.0.

<table>
<thead>
<tr>
<th>Method used</th>
<th>ΔG°</th>
<th>ΔH°</th>
<th>ΔS°</th>
<th>Δn1</th>
<th>Δn2</th>
<th>ΔH</th>
<th>ΔS1</th>
<th>ΔS2</th>
<th>t1/2</th>
<th>t2/1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD</td>
<td>1.5 ± 0.2</td>
<td>4.6 ± 0.3</td>
<td>6.1 ± 0.5</td>
<td>35</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluorescence</td>
<td>1.6 ± 0.2</td>
<td>4.7 ± 0.2</td>
<td>6.3 ± 0.4</td>
<td>35</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNS fluorescence</td>
<td>1.2 ± 0.4</td>
<td>5.0 ± 0.3</td>
<td>6.2 ± 0.7</td>
<td>34</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>194.5 ± 10.0</td>
<td>0.60 ± 0.03</td>
<td>136.3 ± 7.0</td>
<td>0.4 ± 0.2</td>
<td>47.5</td>
<td>66.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a red shift from 340 nm to 350 nm and an increase in the fluorescence intensity.

Urea-induced protein denaturation. Urea-induced unfolding of the protein was reversible at all three protein concentrations used, and the profiles obtained by both CD and fluorescence were similar in all the cases. Fig. 2a shows a typical unfolding profile as a plot of Fapp as a function of the urea concentration, where Fapp is given by the equation:

\[ F_{app} = \frac{Y_{obs}}{Y_{u}} - \frac{Y_{n}}{Y_{u}} \]

where \( Y_{obs} \) is the observed value of the parameter (fluorescence or ellipticity) at a given urea concentration. \( Y_{n} \) and \( Y_{u} \) are the values of the parameters for the native and the unfolded proteins, respectively. The values of \( Y_{n} \) and \( Y_{u} \) are obtained by linear extrapolation of the pre-transition and post-transition regions [35]. All such curves corresponding to the three different protein concentrations have been analyzed. Only one such analysis is presented here explicitly.

The typical unfolding curve shown in Fig. 2a represents data points from two independent modes of observation (CD and fluorescence) employing the same protein concentration and using the same range of urea concentration. The profiles of unfolding in both the cases lead to similar results. Additionally, it was possible to retrace the path of unfolding when the protein in 8 M urea was progressively diluted to initiate and propagate the folding process. These results suggest that the folding/unfolding of Dm-MybR123 can be treated as thermodynamically reversible. Inspection of the unfolding curve shows an apparent biphasic transition; a relative slow phase until approximately 2.0 M urea followed by a more rapid phase at 2.0–5.0 M urea. Initial attempts to explain the data on the basis of a two-state model was unsatisfactory. A quantitative analysis of the transition region of the data of Fig. 2a was, however, possible by a three-state model:

\[ N \xrightarrow{\kappa_1} I \xrightarrow{\kappa_2} U \]

where N, I and U represent native, intermediate and unfolded states, respectively. According to this model, an equation for \( F_{app} \) can be derived following [12] and is given by:

\[ F_{app} = f_n K_u \left( 1 + \frac{d_n K_n}{K_u} \right) \]

where \( f_n \) is the fraction of the native species present at the urea concentration at which the spectroscopic parameter is measured and \( d_n \) is given by the equation:

\[ d_n = \frac{Y_n - Y_u}{Y_u - Y_n} \]

where \( Y_u, Y_n \) and \( Y \) are the values of the parameters for the native, the unfolded and the intermediate states, respectively, of the protein at a given urea concentration. \( K_n \) and \( K_u \) are the equilibrium constants which are given according to the 'denaturation binding model' [12], as
Fig. 3. Changes in the TNS fluorescence as a function of the urea concentration. 7.0 µM protein was incubated at the desired urea concentration for 30 min. TNS was added to a final concentration of 25 µM in each sample and the emission spectra were recorded on exciting the sample at 326 nm. The changes in fluorescence intensity were monitored at 435 nm.

Fig. 4. Tryptophan residues exposed, determined by fluorescence quenching by acrylamide. 9 µM Dm-Myb R123 protein in 10 mM Hepes, pH 7.0 was incubated at the desired urea concentration for 30 min prior to any measurement. The changes in the fluorescence intensity at 340 nm were monitored with increasing concentrations of acrylamide at the desired urea concentration. The data were analyzed according to Eqn 1 and, from the intercept, the fraction of tryptophan residues accessible to the acrylamide was calculated for each urea concentration.

\[ K_e = K_{w}^0 (1 + 0.1 \delta) A_n \]  

and

\[ K_e = K_{w}^0 (1 + 0.1 \delta) A_{n0} \]

where \( \delta \) is the activity of urea, \( K_{w}^0 \) is the equilibrium constant in the absence of urea, \( A_n \) is the difference in the number of binding sites for urea between intermediate and native forms of the protein and \( A_{n0} \) is the difference in the number of binding sites for urea between unfolded and intermediate forms of the protein.

Fig. 5. Thermal denaturation of the Dm-Myb R123. Thermal denaturation curves were obtained by monitoring the changes in the CD spectra at 222 nm using a protein concentration of 90 µM in 10 mM potassium acetate, pH 4.0, (●), 10 mM potassium phosphate, pH 5.0, (□), 10 mM potassium phosphate, pH 6.0, (▼) and 10 mM Mops, pH 7.0, (▲). The apparent fractional change \( F_{app} \) at various temperatures was calculated according to Eqn 2. (b) Fitting of the transition region of the experimentally observed thermal denaturation curve (pH 7.0) to a three-state curve (solid line); (●) represents the experimental data points (a).

Molecular changes accompanying urea-induced unfolding. In a previous study, we have documented that the Dm-Myb R123 binds to a single molecule of the environmentally sensitive probe, TNS. It was further shown that the bound TNS was surrounded by four or five tryptophan residues and at least one tyrosine residue [36]. Thus, TNS was bound to a hydrophobic site contributed by more than one of the Myb repeating units as a result of its tertiary structure. We have, therefore, attempted to evaluate the stability of this TNS-binding 'hydrophobic site' as a function of the urea concentration. Since TNS remains fluorescent in only its protein-bound state, the loss of the TNS fluorescence was monitored to indicate unfolding of this sub-domain. The results presented in Fig. 3 show that urea indeed destabilized this hydrophobic sub-domain as its concentration was progressively increased. The rate of loss of TNS fluorescence closely mirrored the appearance of the unfolding intermediate.

These results seem to indicate that nearly 90% of unfolding has occurred at a urea concentration of approximately 3 M (Figs 2 and 3). To further investigate the molecular changes as-
associated with the progressive denaturation, we have monitored the solvent accessibility index of the tryptophan residues as determined from the acrylamide quenching experiments at various urea concentrations. These are presented in Fig. 4. Unlike the cases of CD, fluorescence and TNS, the tryptophan residues are exposed to the solvent rather differently. For instance, at urea concentration approximately 3.0 M, CD, fluorescence and bound TNS all show essentially a complete loss of tertiary and secondary structures, yet the extent of accessibility of tryptophan residues to the quencher remains only at 67%. It was ascertained that this different result from collisional quenching was not an artifact due to the presence of urea by control experiments where measurements of free tryptophan fluorescence quenching in urea solutions of similar concentration ranges showed no tangible effects. Our data, therefore, leads to the conclusion that some structure is still maintained in 3–6 M urea which could not be detected by other techniques. These observation may be interpreted as suggesting that the unfolding transitions detected by quencher accessibility are probably fundamentally different from the unfolding transitions detected by CD, fluorescence and the bound TNS. This contention is further supported by the observation that the transition shown in Fig. 4 is very complex as it could not be quantified by models involving as many as four states. Nonetheless, this transition was found to be reversible over the entire range of urea concentrations, without any hysteresis. Thus, the exposure of tryptophan residues to the quencher documents a finer dissection of the transition process involving several potential intermediate states and extends the range of urea concentration over which the protein actually unfolds.

**Thermal unfolding of Dm-Myb R123.** We monitored the changes in the intensity of the CD band at 222 nm as a function of temperature. Fig. 6 shows the portions of the clean total correlation spectra in H2O at temperatures 16°C (A), 40°C (B) and 80°C (C). All the cross peaks originate from the NH protons and gross identifications are given in (A). Single-letter symbols identify the amino acid types and Greek symbols α, β, γ, etc. identify particular proton types in the amino acids. Particularly important are the Kε and RD, which correspond to the side chain amino protons.

**Fig. 6. Portions of the clean total correlation spectra in H2O at temperatures 16°C (A), 40°C (B) and 80°C (C).** All the cross peaks originate from the NH protons and gross identifications are given in (A). Single-letter symbols identify the amino acid types and Greek symbols α, β, γ, etc. identify particular proton types in the amino acids. Particularly important are the Kε and RD, which correspond to the side chain amino protons.

**Fig. 7. C,H-C,H cross peaks in the clean total correlation spectra of Dm-Myb R123 in H2O at 16°C (A) and 80°C (B).** 12 distinct peaks are identified in the spectrum at 80°C.
The unfolding of Dm-Myb R123 was found to be reversible in each case. Typical unfolding profiles are presented in Fig. 5a. The unfolding profile in the transition zone of Fig. 5a could be adequately explained with the invocation of a three-state model. A typical three-state fit of the experimental data points at pH 7.0 is shown (Fig. 5b) and the various parameters of the fits are listed (Table 1). For this non-linear fitting, we assumed that the enthalpy changes (AH) and the entropy changes (AS) are not explicitly dependent on temperature in the transition region. It is interesting to note that the standard enthalpy changes for the transitions in this analysis are similar to those obtained by Sarai et al. [37] by differential scanning calorimetry on the mouse R123. This similarity (within 20%) to denatured transitions, both AH and AS terms make nearly equal contributions.

We further used NMR spectroscopy to gain some insight into the molecular aspects of unfolding process. Fig. 6 shows portions of two-dimensional clean total correlation [34] spectra at pH 4.8, containing cross peaks belonging to various NH (both backbone and side chain) protons in the protein at three different temperatures, 16, 40 and 80°C. The different regions of the cross peaks have been identified (Fig. 6A). However, for the present, we focus attention on the peaks belonging to the side chains of the lysine and arginine residues which appear distinctly from the other NH protons [38]. We notice that, as the temperature is increased, the lysine peaks disappear (at 40°C) whereas those of arginine residues remain, indicating that lysine residues become accessible to the solvent faster than the arginine residues. At 80°C, both the lysine and the arginine peaks have disappeared completely. It is clear that the lysine interactions are relatively weak compared to arginine interactions and that some of the intermediates in temperature denaturation likely arise from loss of stability contributions from lysine residues.

Comparison of two-dimensional clean total correlation spectral regions [34] of the protein showing C-H-C,H cross peaks of the histidine imidazole rings of Dm-Myb R123 at 16°C and 80°C (Fig. 7) shows clear differences. The state attained at 80°C is such that some order is still present. For example, at 80°C 12 peaks are seen as against six expected for the six histidine residues. This indicates that there are at least two structures in equilibrium with sizeable populations in the thermal 'denatured' state.

**DISCUSSION**

Our results have showed that the unfolding of Dm-Myb R123 is a complex process involving several states. Unfolding in the range 0–3 M urea is characterized by a process involving at least one intermediate which becomes predominant (at approximately 2.0 M urea) in the three-state equilibrium. As far as CD, fluorescence and TNS are concerned, the unfolding process is essentially complete in 3.0 M urea. The loss of helical structure found in experiments was concomitant with the loss of the TNS fluorescence. Thus, it was imperative that the integrity of the TNS-binding site was in direct correlation with the integrity of the helices present in the R123 structure. Since TNS in its bound state probed the integrity of a relatively specific and a small region of the protein, we expected TNS-urea experiments to allow a finer dissection of the unfolding process but this was not seen.

Experiments designed to quantify the accessibility of the tryptophan residues to the neutral quencher acrylamide present a profile of unfolding which is different from the three other modes. In this case, at 3.0 M urea only 67% of the tryptophan residues were accessible to the quencher, while other methods recorded an apparent completion of the unfolding process. It was not until the urea concentration was raised to nearly 6.0 M that the exposure of the tryptophan residues to the quencher reached near completion. The reversibility of this process underscores the potential importance of the path during protein refolding.
Although residual structures for the denatured proteins have been recorded before [41–46], our results provide another compelling example.

In the thermal unfolding of Dm-Myb R123, we again found that more than two states are necessary to explain the transition at each of the four discrete pH values used. From NMR, we observed that some of the residues display order even at fairly high temperatures. Notably, histidine residues are components of an ordered structure surviving at 80°C, testifying the presence of long-range interactions and making a case for these histidine residues scattered among all the three repeats to be their mediators.

Finally, we have evaluated the heat capacity parameters for the Dm-Myb R123 protein from the unfolding transition, which allows the computation of thermodynamic parameters of stability. The results show that the protein has maximal stability at 16°C and at pH 7.0.

The facilities provided by the 500 MHz FT NMR National Facility supported by the Department of Science and Technology, Government of India, and located at T. I. F. R., Bombay are gratefully acknowledged. We thank P. N. Bhavsar for helpful suggestions. LCP thanks the Lady Tata Memorial Trust for the financial support.

REFERENCES

25. Reference deleted.


