Bacterial expression, characterization and DNA binding studies on *Drosophila melanogaster* c-Myb DNA-binding protein

Anup MADAN, Plachikat K. RADHA, Ramakrishna V. HOSUR and Lakshmi C. PADHY
Tata Institute of Fundamental Research, Bombay, India

(Received 4 May 1995) – EJB 95 0708/2

The *Drosophila* Myb homologue retains an evolutionarily conserved typical sequence of three imperfect tandem tryptophan repeat units (R1-R2-R3) of 51–53 amino acids towards its N-terminus as its presumptive DNA binding domain. Using PCR amplification and the T7 expression vector pET 11d, we have overproduced this tryptophan repeat domain of *Drosophila* Myb in *Escherichia coli* and the protein has been purified. Circular dichroic measurements indicate that the protein has a high helical component (58.6%) in its overall structure. The protein is found to recognize the same cognate target sequence TAACGG, as recognized by the vertebrate proteins. The DNA binding properties of the protein have been investigated in detail by fluorescence spectroscopy taking advantage of the large number of tryptophan residues present in the protein. The fluorescence of the native *Drosophila* R123 was quenched when synthetic duplex DNA oligomers were added to the protein. The oligomers containing specific Myb target sites quenched the protein fluorescence to a greater extent than the non-specific DNA. Binding constants of the protein to the targets were also length dependent for smaller oligomers. Experiments with the collisional quencher acrylamide and cysteine modification reagent indicated that the specific and non-specific target sequences interact with the protein differently. In the former case both the buried and the exposed tryptophan residues were affected by DNA binding whereas in the latter only the solvent-exposed residues were involved.

**Keywords:** Myb; DNA-binding domain; PCR; tryptophan role in DNA binding.

The c-Myb protein is a proto-oncogene product and binds to DNA in a sequence-specific manner [1, 2]. The protein functions as a transcriptional regulator and activates a large number of genes corresponding to different proteins [3–10]. The DNA binding activity of the protein is mediated by a region close to its amino terminus [11, 12]. This domain contains three imperfect, tandem repeats of 51–53 amino acid residues and is represented in general as R1-R2-R3 (or R123 in brief). Each of the repeat units contains three Trp residues that are evenly spaced, 18–19 amino acids apart, and are evolutionarily conserved. Besides, the sequences corresponding to the DNA binding domains of the Myb proteins are highly conserved through evolution and have been detected in as widely divergent species as yeast, *Arabidopsis*, *Zea mays*, *Drosophila*, chicken, mouse and man [13, 14]. It is therefore envisioned that the Trp residues play an important role directly or indirectly in the sequence-specific DNA binding. The notion is further supported by the site-directed mutagenesis experiments where replacement of Trp at many positions drastically affected the DNA binding properties of the R123 protein [15].

The three-dimensional structure of R123 protein is not yet known. It has however been seen from NMR structural studies that the R3 repeat folds into a structure having three short helices and a hydrophobic core consisting of three conserved Trp and several other hydrophobic residues [16]. Similarly, compari-

***Correspondence to*** R. V. HosurL. C. Padhy, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay - 400 005, India

**Abbreviations.** Dm-Myb R1239, DNA binding domain of *Drosophila melanogaster* c-Myb homologue; EMSA, electrophoresis mobility shift assay; Nbs, 5,5'-dithiobis(2-nitrobenzoic acid).

**EXPERIMENTAL PROCEDURES**

**Materials.** Ammonium sulfate, polyethyleneimine isopropyl β-D-thiogalactopyranoside, SP-Sephadex, protein molecular mass markers were purchased from Sigma, DEAE-cellulose (DE-52) anion-exchange resin was from Whatman; bacteriologi
cal media components were from Difco. Enzymes used for recombiant DNA were from New England Biolabs and Promega. The DIG Taq DNA sequence kit was from Boehringer Mannheim.
Fig. 1. Strategy used to construct an expression vector pET-1ld (R123) to overexpress the DNA binding domain of the Dm-Myb protein in E. coli from the pGem 2 vector containing the gene coding for full-length c-Myb protein.

Construction of Dm-Myb R123 expression vector. Fig. 1 summarizes the strategy used for the construction of the expression vector for Dm-Myb R123. We have made use of the pET 11d vector system [23] to express the Dm-Myb R123. Two specific primers, P1 and P2, were designed to selectively amplify the region coding for the DNA binding domain: P1, CGG ATC CTC TAG AGT CTG CGT TCG ACC TC; P2, TGC AGA TCT TGG ~ GGCAAA CGC TGG TCC. The N-terminal primer (P2) contained a BglII site followed by a BarnHI site, while the C-terminal primer (P1) had a protein synthesis termination signal inside a XhoI site followed by a BamHI site. A linearized cDNA clone of Dm-Myb, assembled in pGem-2 vector, served as a template for PCR amplification (cycle 1: denaturation at 95°C, 2 min annealing and extension at 75°C for 45 s, 20 loops). The success of the PCR was checked by electrophoresis on agarose gel. The coding fragment was gel-purified, digested sequentially with BamHI and NcoI, purified and cloned into NcoI/BamHI-digested pET 11d by standard protocols [24]. The recombinants were first verified for the presence of the insert by suitable restriction analysis and were authenticated by sequence analysis. The construct, pET-1ld (R123), was transformed into the Escherichia coli strain BL21 (DE3) [24].

Expression of the recombinant protein. The transformants were grown at 37°C in Luria Bertani medium (10 g Bactotryptone, 10 g NaCl and 5 g yeast extract, pH 7.0) supplemented with 100 µg/ml ampicillin. An overnight culture was diluted and recombinant protein expression was induced by addition of isopropyl β-D-thiogalactopyranoside to 25 μM when A600 reached around 0.7. Cells were harvested after 3 h by centrifugation at 5000 rpm for 10 min at 4°C using a GSA rotor, resuspended in 20 ml buffer A (8 M urea, 20 mM Tris/HCl, 2 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 7 mM 2-mercaptoethanol, pH 8.0)/l culture, lysed by sonication (five 1-min periods at 40% maximum intensity) using a Branson sonifier and centrifuged at 10000 rpm for 10 min using a Sorvall SS34 rotor to remove the cell debris. The supernatant contained the solubilized Dm-Myb R123 protein along with other soluble cellular proteins and DNA.

Purification of the recombinant protein. 15% SDS/PAGE [25] analysis of the bacterial lysate of induced and uninduced cultures of the bacterium harboring the pET 11d (R123) expression plasmid showed that Dm-Myb R123 was present in bacteria only after induction. Gel analysis was subsequently followed at all other stages of purification of the protein. The general procedure involved four steps. First, DNA, RNA and phospholipids were removed by 0.03% poly(ethyleneimine) precipitation. Second, the Dm-Myb R123 was precipitated with 50% ammonium sulfate and pelleted at 15000 rpm for 10 min using a SS34 rotor. The pellet obtained from 11 culture was dissolved in 10 ml buffer A. In the next step, the solutions were passed through a DEAE-cellulose column (5X45 cm) previously equilibrated and developed with the lysis buffer. The elution of the proteins was monitored by measuring the absorbance at 280 nm and the fractions containing the Dm-Myb R123 were pooled. In the final step the pH of the pooled fractions was adjusted to 6.5 by adding Mes and loaded on SP-Sephadex (2.5X15 cm) equilibrated with buffer B (8 M urea, 20 mM Mes, 0.15 M NaCl). The column was washed extensively with the buffer B and the recombinant protein was eluted by a linear gradient of 0.15–1.0 M NaCl in buffer B. The recombinant protein eluted at 0.83 M NaCl and the fractions containing the Dm-Myb R123 on the basis of SDS/PAGE analysis were pooled. After SP-Sephadex column chromatography, the preparation was essentially homogeneous, >95% by SDS/PAGE.

Refolding of the protein. The purified protein from the SP-Sephadex chromatography step containing 8 M urea was gradually dialyzed to a final protein concentration of 0.2 mg/ml and further dialyzed to a concentration of 0.1 M with 20 mM Tris/HCl, 1 mM EDTA, 10 mM dithiothreitol, pH 8.5 and allowed to rock gently at 41°C for 24 h. The protein was then concentrated by ultrafiltration using an Amicon YM-10 membrane and was equilibrated to the storage buffer (20 mM Tris/HCl pH 7.0) by dialysis. The protein folding was monitored by characteristic changes in fluorescence and circular dichroic spectra.

Spectroscopic analysis. CD Spectra were acquired on a Jasco-J 600 spectropolarimeter. The instrument was calibrated with (+)-10-camphorsulfonic acid as standard. Ultraviolet absorption spectra were recorded on a Shimadzu UV-2100 spectrophotometer and fluorescence spectra were recorded on a Shimadzu RF-540 spectrofluorimeter. The spectra were corrected for dilution and contributions from buffer and DNA. Two-dimensional NMR spectra were recorded on a 500-MHz Bruker AMX system.

DNA binding experiments. The following duplex oligonucleotides have been used for DNA binding experiments; the specific Myb binding site has been underlined, and only one strand is indicated. Specific: (a) MRE-26 [3], TCGACACATTAAACGTTTTTACGC; (b) MRE-16 [20], CCTAATTGCACTACAT; (c) MRE-12, ACCGTTAAGCGT. Non-specific: (a) DSNS-12, GGTACGCGTGAC; (b) SSNS-21, GTTCAGCGGGGATCCCTGAA.
Initial electrophoretic mobility shift assays (EMSAs) were performed with MRE-26. DNA binding experiments in solution were monitored by the changes in the protein fluorescence at 342 nm ($
abla_{ex} = 295$ nm) at 22°C. Equilibrium reverse titrations [26–31] were performed by adding increasing aliquots of DNA to Dm-Myb R123 and incubating the mixture for 15 min at 22°C. The fluorescence intensities were noted ($F_{\text{obs}}$) and corrected ($F_{\text{corr}}$) for inner filter effect by using Eqn (1) [32]:

$$F_{\text{corr}} = F_{\text{obs}} \text{anti-log} \left( \frac{A_{ex} + A_{em}}{2} \right)$$

(1)

where $A_{ex}$ and $A_{em}$ refer to absorbances at the excitation and emission wavelengths, respectively. Fractional fluorescence quenching ($F_0$) on DNA binding was calculated for each data point as:

$$F_0 = \frac{F_0 - F_{\text{corr}}}{F_0}$$

(2)

where $F_0$ is the fluorescence of the free protein.

The stoichiometry of DNA–protein interaction was determined as follows. The observed fluorescence ($F$) at any point in the titration curve is the sum of the fluorescence of the bound and free protein

$$F = f_b F_b + f_f F_f$$

(3)

where $F_b$ and $F_f$ are the fluorescence intensities of the bound and the free proteins respectively and $f_b$ and $f_f$ are their fractions. The mole fraction ($M_k$) of the DNA for the initial points where all the added DNA is in the bound form is given by the equation

$$M_k = \frac{N P_b}{N P_b + P_f}$$

(4)

where $N$ is the number of binding sites, $P_b$ and $P_f$ are the total and bound concentrations, respectively. In the initial stages $NP_b \ll P_f$ and hence $NP_b$ can be neglected in the denominator. Thus,

$$M_k = \frac{P_b}{P_f} = f_b.$$  

(5)

Substituting for $f_b$ from Eqn (5) into Eqn (3):

$$F = \frac{M_k}{N} F_b + \left(1 - \frac{M_k}{N}\right) F_f.$$  

(6)

Rearranging Eqn (6):

$$F - F_0 = \frac{M_k}{N} F_b - \frac{F_0}{N}.$$  

(7)

Thus, from the plot of $M_k$ versus $\Delta F/\Delta F_{obs}$, the stoichiometry of the Dm-Myb R123–DNA interaction can be determined. In the present experiments, the plots were found to be linear for the DNA sequences investigated and stoichiometry was found to be 1:1.

For 1:1 stoichiometry, the following equilibrium holds

$$P + L \rightleftharpoons PL.$$  

The binding constant is then given by

$$K = \frac{L_o}{P_f L_p}$$

where $K$ is the binding constant, $L_o$ is the concentration of the bound DNA, $L_o$ and $P_f$ are the concentrations of free DNA and protein, respectively.

Rearranging,

$$\frac{1}{K} = \frac{P_f}{P_o - L_o}.$$  

And

$$L_o = \frac{P_o L_o}{P_f}.$$  

By using Eqn (3), the above equation can be written in terms of observable fluorescence as

$$\frac{\Delta F}{F_0} = \frac{K L_o \Delta F_{obs}}{(1 + K L_o) F_0}.$$  

(8)

Here $\Delta F$ is the quenching observed at any particular concentration of DNA and $\Delta F_{obs}$ is the quenching at infinite DNA concentration. For practical purposes Eqn (8) was rearranged to the equivalent double-reciprocal relation:

$$\frac{1}{\Delta F} = \frac{1}{K L_o} \frac{\Delta F_{obs}}{\Delta F_{obs}} + \frac{1}{\Delta F_0}.$$  

(9)

The free DNA ($L_o$) values were calculated by subtracting the bound DNA concentration from the total DNA concentration. The concentration of the bound DNA ($L_o$) was obtained from Eqn (10), which can be easily derived using Eqn (3) for 1:1 stoichiometry of interaction,

$$L_o = P_b = f_b P_0 = \frac{F_0}{F^o} P_0$$

(10)

where $F^o$ refers to the value of $F_0$ at infinite DNA concentration.

A double-reciprocal plot (1/$\Delta F$ vs 1/$L_o$) yields a straight line whose slope yields (1/$K L_o$) and the intercept yields 1/$\Delta F_0$, from which the value of $K$ can be calculated.

**Fluorescence quenching by acrylamide.** Fluorescence quenching experiments designed to probe the solvent accessibilities of the Trp residues that may be involved in DNA binding were performed with the neutral quencher acrylamide. The data were analyzed according to the Lehrer equation [33]:

$$\frac{F_0}{F_0 - F} = \frac{1}{f_k \tau_o (Q)} + \frac{1}{f_s}.$$  

(11)

where $F_0$ and $F$ are the fluorescence intensities of the protein in the absence and presence of the quencher, $f_s$ is the fraction of the quenchable fluorescence, $k_q$ is the rate constant for the quenching, and $\tau_o$ is the fluorescence life time of the fluorophores in the absence of the quencher.

**RESULTS**

**Bacterial expression of Dm-Myb R123 protein.** We placed the codons for a stretch of 160 amino acids encompassing only the R123 region of Myb with an engineered N-terminal start and C-terminal stop codons into the bacterial expression vector pET 11d. Upon characterization of the recombinant and induction with isopropyl $\beta$-D-thiogalactopyranoside, a novel protein of expected size ($\approx$19.5 kDa) could be seen in 15% SDS/PAGE analysis (Fig. 2, lane 2). The level of expression of the putative Dm-Myb R123 was generally high as early as 1 h post-induction but steadily increased to nearly 20% of the total cellular protein in 3 h. Preliminary fractional experiments showed that a part of the expressed protein was associated with inclusion bodies. By using buffers containing 8 M urea in a three-step purification protocol, we could obtain Dm-Myb R123 protein >95% in purity. The SDS/PAGE profile through various stages of purification and after refolding is shown in Fig. 2.

**Characterization of Dm-Myb R123.** Several criteria were applied to ensure that the bacterially expressed protein was indeed
Fig. 2. 15% SDS/PAGE profile of the Dm-Myb R123 through various stages of purification. Lane 1, BL21 (DE3)/pET 11d (R123) crude cell lysate without induction; lane 2, BL21 (DE3)/pET 11d (R123) crude cell lysate after induction; lane 3, after poly(ethyleneimine) precipitation of the crude cell lysate; lane 4, solution obtained on dissolving 50% ammonium sulfate precipitate; lane 5, DEAE-cellulose pool containing R123; lane 6, SP-Sephadex pool containing R123; lane 7, after refolding the protein. Molecular mass markers are indicated with arrows and expressed in kDa.

Dm-Myb R123: (a) the Dm-Myb protein was reactive to an antibody raised against N-terminal half of the Dm-Myb protein; (b) the ratio Trp/Tyr present in the protein determined by absorption [34, 35] was as expected; (c) two-dimensional (2D) NMR spectra (Fig. 3) of the protein showed unambiguous resonance signals emanating from the different residue types as expected for Dm-Myb R123. These results confirm that the bacterially expressed protein truly represents the Dm-Myb R123 DNA binding domain. Finally, the DNA binding activity of the purified protein was checked by gel shift assay. The protein was found to retard the 26-residue DNA containing the specific target sequence [3] in presence of 1000-fold excess of poly(dI)·poly(dC) (data not shown).

Predominantly helical structure of Dm-Myb R123. The indication that the helical structure was a component of Dm-Myb R123 secondary structure came from the CD experiments. Representative CD spectra of unfolded (8 M urea) and folded Dm-Myb R123 are shown in Fig. 4a. Upon quantitation [36] in aqueous solution, the helical content was found to be 58.6% The CD spectra indicated further that the β-structure content in Dm-Myb R123 protein is likely to be minimal. These results are similar to the data obtained with the R23 and R3 proteins [16, 37]. The helical content could be further enhanced by using helix-promoting solvents such as trifluoroethanol (Fig. 4b) indicating that Dm-Myb R123 still retained a potential to increase its helical content.

Binding of Dm-Myb R123 to oligonucleotides. Earlier studies on the R123 protein have noted that its binding affinity to DNA oligomers is length-dependent [20]. We have therefore studied binding of DNA oligomers of 26, 16 and 12 base pairs containing the Myb responsive element (MRE) to the Dm-Myb R123. When slightly more than the stoichiometric amounts of these oligomers were added to the R123 protein, a substantial quenching of the protein fluorescence in the ranges of 46–52% was observed (Fig. 5a). That the quenching was derived from bona fide DNA–protein interactions and was not due to the non-specific affects became clear from the control experiments. The quenched fluorescence of R123·DNA complexes showed a substantial fluorescence recovery upon treatment with 2 M NaCl. Similarly, when non-specific DNAs were added to the protein, the observed maximal quenching was in the range of 12–14% (Fig. 5b) and again 2 M NaCl could restore a substantial part of the quenched fluorescence. On the other hand, 2 M NaCl had no significant effect on the R123 fluorescence itself. We reason that the specific and non-specific DNAs interact with the Trp of the R123 in its bound state resulting in the observed quenching and that the differential quenching observed in the two cases is due to different numbers or types of Trp residues involved. This is in accordance with our earlier observation that DNA carrying MRE could displace a hydrophobic-site-specific probe, 6-(p-toludino)-2-naphthalene sulfonate from its binding site while a non-specific oligomer failed to do so [39].

We then exploited the fluorescence quenching property of DNA to assess quantitative parameters of the Dm-Myb R123.
Fig. 4. CD spectra of Dm-Myb R123. (a) CD spectra of the native (2) and denatured (1) Dm-Myb R123 in 10 mM Tris/HCl pH 7.0 at 25°C. The protein concentration was 90 μM and 8 M urea was used for denaturation. (b) CD spectra of Dm-Myb R123 (70 μM) in the presence of (1) 0%, (2) 10%, (3) 25% or (4) 50% trifluoroethanol. For each spectrum 10 scans were taken and averaged. Each spectrum was recorded with 0.2-nm wavelength increments and the signal was acquired for 1 s at each wavelength.

Fig. 5. DNA binding activity of Dm-Myb R123 monitored by change of fluorescence. The protein concentration was 7 μM. A 1:1 protein–DNA complex was prepared by incubating equimolar amounts of protein and DNA in 10 mM Tris/HCl pH 7.0 for 15 min at room temperature. The oligomers used for (A) specific (MRE-12) and (B) non-specific (DSNS-12) are as described in Experimental Procedures. In both A and B, curve 1 is the fluorescence spectrum of the protein alone, curve 3 is the fluorescence spectrum of the protein–DNA complex and curve 2 is the spectrum obtained after adjusting the solution containing the protein–DNA complex to 2 M NaCl. The samples were excited at 295 nm and the spectra were corrected by subtraction of the response of buffer and buffer + DNA as appropriate.

Table 1. Binding constants for different Dm-Myb R123 complexes.

<table>
<thead>
<tr>
<th>DNA</th>
<th>Apparent binding constant M⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRE-26</td>
<td>3.8 × 10⁶</td>
</tr>
<tr>
<td>MRE-16</td>
<td>2.3 × 10⁷</td>
</tr>
<tr>
<td>MRE-12</td>
<td>1.7 × 10⁷</td>
</tr>
<tr>
<td>DSNS-12</td>
<td>2.5 × 10⁶</td>
</tr>
<tr>
<td>SSNS-21</td>
<td>5.9 × 10⁸</td>
</tr>
</tbody>
</table>

in Fig. 6a for which the added DNA can be considered to be almost entirely bound to the protein were analyzed according to Eqn (7) (Fig. 6b) and the stoichiometry of interaction was found to be 1:1. It was then possible to use the ΔF and ΔFₐ values to calculate the free ligand concentration (L₀) at higher input values of the DNA (Eqn 10) and from a plot 1/ΔF versus 1/L₀ using 4 or 5 points below saturation in Fig. 6a the binding constant of the interaction were calculated (Fig. 6c). Binding constants determined in this manner for various specific and non-specific targets show a dependence on DNA length in our experiment (Table 1) which is in agreement with the earlier data [20]. It may be noted that, although the binding constants of the specific targets with respect to the R123 changed among the oligomers, the maximal quenching did not change appreciably, indicating that it is not determined by the relative affinity of the DNA to the protein.

Specific DNA binding and the helical content of R123. A previous study [21] hypothesized that there may be a conformational change in R123 when it is bound to the DNA. As a component of this conformational change the authors were able to detect a 10% increase in the helical content of R23. We have performed a number of experiments with Dm-Myb R123 and various DNA oligomers to address this issue. The effect of DNA on the conformation of Dm-Myb R123 was investigated by spectroscopy; the results are presented in Fig. 7A. It is seen clearly that there is really an insignificant change in the CD spectra of the protein on interaction with specific or non-specific DNAs. In a separate set of experiments, we held the total input of DNA as constant but varied DSNS-12 and MRE-12 such that the mole fraction of MRE-12 steadily increases from 0 to 1. CD
ellipticity parameters determined in three separate experiments for all the samples again failed to show any significant changes (Fig. 7B). Our results thus demonstrate that DNA binding to R123 does not lead to a convincing increase in the helical content at least, under the experimental conditions tried. This is similar to the observations made by Sarai et al. in mouse R123 [17].

**Role of Cys90 in DNA binding.** Cys90 in *Dm-Myb* R123 is a counterpart of a highly conserved cysteine seen with several other vertebrate Myb proteins that has recently been implicated in the redox regulation [21, 38]. In order to test whether Cys90 was important for the sequence-specific DNA binding in *Drosophila* protein, the following experiments were performed. We studied first whether Cys90 of *Dm-Myb* R123 was solvent-exposed by comparing its reactivity with 5,5′-dithiobis(2-nitrobenzoic acid) (Nbs₂) in its native and denatured (8 M urea) state.

As Fig. 8A shows, regardless of whether protein was in the native state or denatured, Cys90 reacted with Nbs₂ quantitatively and with similar kinetics, indicating that it was solvent-exposed. The resulting thio-nitrobenzoate (Nbs) derivative of native R123 upon Nbs₂ reaction was then checked for its specific DNA binding ability. The extent of Trp quenching observed were seen to be similar for both specific and non-specific DNA targets indicating that the protein has compromised its specific DNA binding property upon Nbs derivation. Addition of dithiothreitol to non-specific DNA—R123 complex did not lead to significant changes in the maximum quenching of R123 protein fluorescence by DNA, although the Nbs was released quantitatively as determined by absorbance at 412 nm. In contrast, in similar experiments with specific DNA, the maximum quenching was seen to be enhanced to nearly 50%, clearly indicating that the sequence-specific DNA binding property has been restored to the *Dm-Myb* R123 protein after reduction and removal of the
A

B

Fig. 8. Solvent accessibility and the role of Cys in DNA binding. The solvent accessibility of the Cys residue was determined by reacting the R123 protein with 5,5′-dithiobis(2-nitrobenzoic acid) (Nbs,). The reaction was performed with 12 μM R123 protein and 1 mM Nbs, in potassium acetate pH 4.5 at room temperature and monitored spectroscopically at 412 nm for both native (⊙) and denatured protein (○). 1:1 protein–DNA complexes with MRE-12 and DSNS-12 were formed as described in the legend of Fig. 5 and the solvent accessibility of Cys after interacting with the two targets was determined by treating with Nbs, (▲). Trace obtained with the MRE-12–protein complex, (▼) protein–DSNS-12 complex. (B) Role of Cys in DNA binding. A thionitrobenzoate (Nbs) derivative of the native protein was generated as described in (A) and was dialyzed against 10 mM Heps pH 7.0 to remove excess of Nbs, and checked for its ability to bind to specific and non-specific target DNA by monitoring the changes in the Trp fluorescence as described in the legend of Fig. 5. (1) Fluorescence of native protein; (2, 3) controls showing effect of derivatization; (4) effect of dithiothreitol on derivatized protein (the reaction was quantitative as expected amount of Nbs was liberated); (5, 6) derivatized protein–DSNS-12 complex before and after treating the complex with 10 mM dithiothreitol respectively; (7, 8) interaction with MRE-12.

Nbs moiety. A separate set of reciprocal experiments were performed by treating the complexes of R123 with specific or non-specific DNA with Nbs,. The nature of the Nbs, reactivity was found to be very different in these two cases. In the case of non-specific DNA, the protein–DNA complex showed stoichiometric reactivity and kinetics comparable to that of the free protein. On the other hand, with specific DNA–R123 complex, the kinetics was much slower and a very small fraction of the protein was modified on the same time scale. These results clearly show that a free Cys90 is important for specific DNA binding. Furthermore, a specifically bound DNA occludes the free thiol group from Nbs, accessibility whereas a non-specific DNA fails to do so. These results are similar to the N-ethylmaleimide reactivity pattern observed earlier [21] and again underscores the differences in the mode of specific and non-specific DNA binding to the protein.

Involvement of Trp in DNA binding. We have determined in a previous study [39] that three of the nine Trp in free R123 are quenched by acrylamide by virtue of their solvent accessibility. We have further probed the solvent-accessible Trp of various specific and non-specific DNAs, albeit indirectly, by using a collisional quencher acrylamide. The dynamic quenching experiments with the protein–DNA complexes were performed and illustrative results for specific and non-specific DNAs are displayed in Fig. 9. The results obtained for various oligomers are summarized in Table 2. It is evident that, with non-specific DNA–R123 complexes, about 2 mol Trp/mol remain solvent exposed. In contrast, the specific DNA–R123 complex contained apparently a single Trp residue exposed to the solvent.

Fig. 9. Lehrer plots of quenching of the fluorescence emission of 7 μM R123–MRE-12 complex (A) and R123–DSNS-12 complex (B). F₀ and F are the fluorescence values at 342 nm in the absence and presence of the quencher. The values of fₐ (fractional accessibility to the quencher) were calculated from the intercept of the plot.
Thus, the non-specific DNA occluded or quenched only one of the three solvent-exposed Trp whereas specific DNA could do the same to two of them. In this respect again, the binding mode of non-specific and specific DNA targets are clearly distinguishable from each other.

DISCUSSION

We have succeeded in the bacterial expression of R123 protein encompassing a 160-amino-acid DNA binding domain derived from the Drosophila homologue of c-Myb. Our experiments demonstrated that this protein was capable of binding with high affinity the same cognate binding sequence recognized by vertebrate proteins. This is not unexpected since Drosophila R123 shares nearly 66% identity with the vertebrate proteins.

Our determination of 58.6% of the helical content for R123 again underscores the similarity of Drosophila protein with its vertebrate counterparts in terms of its secondary structure content. It is thus expected that the DNA binding site of this protein as well as the mechanistic aspects of its binding to DNA are likely to be similar to those in all the Myb proteins. The observation that a known helix-promoting solvent, trifluoroethylamine, induced extra helical content signifies a degree of flexibility in the R123 structure.

We could not detect any increase in the helical content of Myb R123 protein upon its binding to DNA. Myrset et al. [21] reported a 10% increase in the helical content of chicken R23 on interacting with the specific DNA. However, Ebaneth et al. [40] failed to detect any increase in the helical content of chicken R23 but they observed such an increase only in the case of R123. This is in contrast to the earlier reports on R123 derived from mouse where no change in the helical content was observed on interacting with specific DNA [17]. The divergent results in all the cases are probably due to the intrinsic structural features of the protein or due to the different experimental conditions used: for example, the chicken R123 used by Ebaneth et al. [40] has an additional 39 N-terminal amino acids as compared to that of mouse R123 [17]. We have estimated the helical content of Dm-Myb R123 to be around 59% as compared to that of chicken R123 (49%) which also suggests differential structural features in these proteins.

Our studies of protein–DNA interaction in solutions have provided an insight into the manner in which specific and non-specific DNA are bound to the R123 protein. First, we found that the binding affinity of R123 protein to its small specific targets could be dependent on the length of the DNA for smaller oligomers. In agreement with earlier observations [20], we found that the binding affinity decreased in the order 26 bp > 16 bp > 12 bp. Notably, the 12-bp DNA contained a sequence with two palindromic binding sites but the binding stoichiometry still remained 1:1, thus indicating that binding of one protein precluded the binding of the other. The fact that all specific oligomers (26 bp, 16 bp and 12 bp) effected the protein quenching to nearly the same extent (46–52%) suggests that the extent of quenching is not related to binding affinity but rather to the mode of DNA binding. Thus, non-specific and specific 12-bp DNAs have similar binding affinities but the extent of quenching seen in these two cases were drastically different. This observation may have the following important implications: the specific 12-bp DNA is perfectly capable of interacting with R123 in a sequence-specific manner in solution although its binding affinity is low. But in standard electrophoresis mobility shift assays (EMSA), this interaction in all likelihood would have scored negative due to competition from non-specific carrier DNA and lead to possible erroneous conclusions. It is therefore important to exercise caution while interpreting EMSA data, particularly on site-directed mutants when loss of specific interactions is to be distinguished from the loss of binding affinity.

The increase in the binding affinity observable with longer DNA targets has been attributed to the stabilization contributed by the R1 repeat by its non-specific binding effects on the DNA. This can only account for part of the results because the binding affinities of R123 and R23 on the same high-affinity target were found to differ by only a factor 7 [22]. Clearly, the stability of the protein—specific-DNA complex is not solely determined by the sequence-specific interactions; other factors are important and need to be discovered to explain the differences in quantitative terms. Recent observations that Myb bends the target DNA upon binding [41] is one such factor.

In reversible Cys90 modification experiments, we obtained clear implications that a free thiol group at this position was important for sequence-specific DNA binding. In this respect, the Drosophila protein seems to be similar to its vertebrate counterparts. Our observation that the specific rather than non-specific DNA occludes Cys90 suggests that a redox regulation process is likely to work on free c-Myb protein but not when it is already bound to DNA. The redox regulation may therefore achieve the purpose of actively regulating the DNA binding step rather than its dissociation from the protein—DNA complex.

Our experiments designed to probe the status of the solvent-exposed Trp in R123–DNA complexes have led to potentially interesting results. With the approximation that the quantum yields of all the individual exposed Trp are similar, the data in Table 2 suggest the presence of one and two solvent-exposed Trp in specific and non-specific DNA–R123 complexes respectively. A minimum of four or five Trp seem to be involved directly and/or indirectly in the specific DNA binding process. In contrast, the non-specific DNA binding does not seem to involve the buried Trp at all. This is not unreasonable in that physical clustering of several Trp and a Tyr in R123 has already been noted by us in a previous study [39]. This hypothesis on the role of Trp in DNA binding, however, does not exclude other possibilities.

In summary, we have demonstrated that a 160-amino-acid segment derived from the N-terminal portion of Drosophila c-Myb is fully competent to function as a specific DNA binding protein and our results indicate that several of the Trp and the unique Cys90 constitute the essential elements of architecture of its DNA binding site.

The authors thank Dr William Studier for the gift of T7 expression vector. We thank P. N. Bhavsar for helpful discussions. LCP thanks the Lady Tata Memorial Trust for financial support.

REFERENCES

The cellular proto-oncogene product Myb acts as a transcriptional activator of the long terminal repeat (LTR) sequences and transactivates LTR-mediated transcription. The proto-oncogene c-myb binds to and activates the promyelocyte-specific mim-1 gene. Functional analysis of the c-myb protein in T-Lymphocytic cell lines shows that it trans-activates the c-myc promoter, which is consistent with the proposed structure of the Myb oncoprotein based on X-ray crystallography and NMR spectroscopy. Recognition of specific DNA sequences by the c-myb binding domain, which forms a hydrophobic core, is a critical aspect of Myb function. The role of three repeat units in the DNA binding domain of the Myb oncoprotein has been studied, and it is proposed that one of these repeat units functions as a DNA binding domain. The tryptophan cluster in the Myb protein is essential for its DNA binding activity, and the interaction of v-Myb and c-Myb with DNA has been characterized. The tryptophan cluster has been shown to be essential for Myb DNA binding, and the interaction of c-Myb with DNA has been studied using molecular cloning and NMR spectroscopy. The characterization of the c-Myb DNA binding domain using NMR spectroscopy and computational methods has provided insights into the mechanism of DNA binding by Myb. The Myb oncoprotein induces DNA bending, which is essential for its function as a transcriptional activator.