The polypyrimidine tract binding (PTB) protein interacts with single-stranded DNA in a sequence-specific manner

Franck Brunel, Mario M. Zakin, Henri Buc and Malcolm Buccke

ABSTRACT

Polypyrimidine tract binding (PTB) protein is a cellular factor whose function is unknown. Various RNA or single-stranded DNA sequences have been shown to interact with PTB. In this paper, using laser UV cross-linking and electrophoretic mobility shift assays to probe DNA–protein interactions, we demonstrate that PTB binding at a single-stranded DNA target is highly sequence-specific. We provide data showing that PTB interacts with the top strand of the adenovirus major late promoter transcriptional initiator, a sequence rich in pyrimidine residues. We also demonstrate that PTB is organised into at least two different binding domains.

INTRODUCTION

Polypyrimidine tract binding (PTB) protein (1–3), also named PYBP (4), is an ubiquitous cellular factor binding to RNA or single-stranded DNA, and is highly conserved during evolution. It has not been possible, to date, to assign an unambiguous biological function to PTB, although its implication in important biological mechanisms (2–6) has been previously suggested.

PTB was identified as a protein interacting with polypyrimidine-tracks present in pre-mRNAs, and proposed to be involved in splicing mechanisms (2,3,7). In addition, we and others have previously reported (4,5) that highly purified PTB was able to interact with pyrimidine-rich single-stranded DNAs present in at least three regulatory cis-acting elements of hepatic genes.

The adenosine major late promoter (AdMLP) contains a transcriptional initiator (8) that is a target for different proteins (9–12). A consensus sequence of transcriptional initiators has been derived showing that these regions are rich in pyrimidine residues (9,13). Therefore, we decided to investigate the specificity of PTB for this pyrimidine-rich sequence using laser UV photocrosslinking. Irradiation of a nucleoprotein complex induces zero length covalent bonds between amino acids and bases in direct contact (14–17). We use this technique to show that recombinant PTB interacts with a high affinity with the adenovirus major late promoter transcriptional initiator in a sequence-specific manner. We also demonstrate that PTB is organised into at least two different binding domains providing support for previous structure predictions (4).

MATERIALS AND METHODS

Preparation of recombinant PTB and PTBΔ337–531 proteins

PTB and its N-terminal moiety (PTBΔ337–531) were expressed in Escherichia coli as glutathione-S-transferase (GST) fusion proteins and purified using glutathione–Sepharose beads. The pGEX-2T plasmid (Pharmacia) was modified in order to create a unique XbaI restriction site where PTB cDNA could be cloned in frame with the GST sequence. The oligonucleotides 5′-GTCTAGAT-3′ and 5′-ATCTAGAC-3′ were phosphorylated, annealed and cloned in the SmaI site of pGEX-2T. PTB open reading frame was PCR amplified using a set of two primers designed to create an XbaI site both upstream of the AUG codon and downstream of the stop codon (4): forward primer: 5′-CTAGTCTAGAACCATGGACGCTCGGCGTCCCAGAC-3′; reverse primer: 5′-CTAGTCTAGAAGCTTCGCAGGCCCCACAGG-3′.

The PCR-amplified fragment was then cut with XbaI and introduced into the same site of the engineered pGEX-2T vector giving rise to the pGST-PTB plasmid. The in-frame insertion was checked by nucleotide sequencing.

The amino acid terminal moiety of PTB was constructed as follows: pGEX-GST-PTB was digested with EcoRI and religated. EcoRI sites in pGEX-GST-PTB are located at the codon encoding amino acid 336 and in the vector, upstream at three different frame stop codons. The resulting construct (PTBΔ337–531) produced a truncated PTB protein containing amino acids 1–336.

Escherichia coli strain DH5α containing pGST-PTB or pGST-PTBΔ337–531 were grown in LB broth supplemented with ampicillin (100 µg/ml). Cells were induced with IPTG (100 µg/ml) when the OD reached 0.4 AU. After a further 2 h, cells were harvested by centrifugation at 4000 g for 30 min. Fusion proteins were purified over a glutathione–Sepharose affinity column from sonicated bacterial lysates as described (18) with minor modifications. Products eluting from the glutathione–Sepharose beads were digested with 0.2% (w:w) thrombin for 1 h at 25°C and loaded onto a heparin–Sepharose column equilibrated in BC buffer containing 50 mM NaCl (BC buffer consists of 50 mM Tris–HCl, pH 7.9, 0.2 mM EDTA, 1 mM DTT, 0.05% NP-40, 10% glycerol). Under these conditions, the GST flows through the resin and bound proteins are eluted with steps of increasing salt concentration in BC buffer. Peak fractions were pooled, protein concentration measured

Received February 8, 1996; Revised and Accepted March 21, 1996

* To whom correspondence should be addressed
we asked the question about the specificity of the PTB binding to the transcription start site (9) that is rich in pyrimidine residues (13), since the adenovirus core promoter contains a region encompassing a DNA sequence (5) that is shown to be involved in the specificity of DNA binding (11).

Laser UV cross-linking of nucleoprotein complexes

Complexes were formed at 4°C in a volume of 10 μl in a buffer containing 25 mM Tris, pH 7.5, 50 mM NaCl, 0.2 mM EDTA, 1 mM DTT, 10% glycerol, 1 mg/ml BSA, 10 ng/μl poly(dI–poly(dC), 0.5 mM radiolabelled probe. Following a 20 min incubation, a single (5 ns duration) high intensity pulse of radiation at 266 nm (0.5 × 10¹¹ W/m² intensity providing a dose of ∼100 J/m²) was applied to the sample.

SDS–PAGE of cross-linked complexes

Irradiated samples were diluted 1:1 with SDS–PAGE loading buffer (final concentration = 3% SDS, 5% glycerol, 31 mM Tris–HCl, pH 6.8, 100 mM DTT) heated at 90°C for 3 min and applied to SDS–PAGE (10% acrylamide) at 200 V. Gels were dried and autoradiographed. For quantification purposes, gels were scanned using a PhosphorImager (Molecular Dynamics).

Gel filtration of proteins

Samples were applied to a Superdex 75 column (exclusion range 3–70 kDa) on the SMART (Pharmacia) system. Chromatograms were carried out at 15°C at 40 μl/min, Kav values were calculated for peaks measured at 215 nm and compared with values obtained for proteins having known molecular masses and Stokes radii separated on the same column under identical conditions.

RESULTS

Purification of the recombinant proteins PTB and PTBΔ337–531

cDNAs coding for PTB or for its N-terminal moiety (PTBΔ337–531) (Fig. 1A) were cloned in pGEX-2T plasmids, then introduced in bacteria and the corresponding proteins, fused to the glutathione S-transferase (GST), overexpressed. The two proteins were isolated from bacterial lysates using affinity column chromatography with a glutathione–Sepharose support (18). The GST and the PTB and PTBΔ337–531 recombinant proteins were released by thrombin cleavage and applied to a heparin–Sepharose column. As shown in Figure 1B, GST passed through the heparin column whereas PTB or PTBΔ337–531 were retained and could be eluted by a salt gradient. As judged by Coomassie Blue staining of the SDS–PAGE analysis gels the recombinant proteins are ∼90% pure, and have apparent Mr values of 58 and 32 kDa for PTB and PTBΔ337–531 respectively. These values agree with the calculated molecular masses (5690 and 35267 Da respectively) obtained from the corresponding nucleotide sequences (4). Gel filtration upon Superdex-75 using the SMART system (Pharmacia), also confirmed the degree of purity, giving apparent Mr values of ∼58.4 and 35.9 kDa for PTB and PTBΔ337–531 respectively (data not shown) by comparison with globular proteins of known molecular mass.

Specificity of DNA binding

Since the adenovirus core promoter contains a region encompassing the transcription start site (9) that is rich in pyrimidine residues (13), we asked the question about the specificity of the PTB binding to this element, as a model to study its interaction with single-stranded DNA. To address this question, we used laser UV cross-linking to induce direct covalent links between amino acids and nucleic acids involved in intimate contacts. The oligonucleotide probes used consisted of the pyrimidine-rich top strand (Y) and the purine-rich bottom strand (R) of the AdMLP initiator region (see Table 1). These probes were radioactively end-labelled at the 5′ position. Probes were either used as single-stranded oligonucleotides or as a hybrid (RY) with the 5′ position of the probes and irradiated with a single pulse at 266 nm as described in Materials and Methods. Samples were applied to SDS–PAGE and cross-linked material identified (Fig. 2) as labelled bands migrating slower
Figure 2. Specific interaction between PTB and the pyrimidine-rich strand of the adenovirus major late promoter (AdMLP) initiator. PTB was incubated with the indicated labelled oligonucleotide probes. After a single (5 ns) pulse of 266 nm radiation from a NdYAG laser as described in Materials and Methods, samples were separated on SDS–PAGE. Cross-linked and free oligonucleotide were visualised by autoradiography. B) UV-induced cross-linked probe (bound); F: free probe; Y strand: pyrimidine-rich single-stranded probe (top strand of the AdMLP); R strand: purine-rich single-stranded probe (bottom strand of the AdMLP); RY*: double-stranded oligonucleotide probe labelled on the Y strand. The concentration of PTB was 0 nM (lanes 1, 7 and 13); 27.5 nM (lanes 2, 8 and 14); 55 nM (lanes 3, 9 and 15); 110 nM (lanes 4, 10 and 16); 220 nM (lanes 5, 11 and 17) and 440 nM (lanes 6, 12 and 18). The concentrations of radiolabelled probe were estimated as being of the order of 80 pM.

Table 1. Substrates used for binding studies

<table>
<thead>
<tr>
<th>WT</th>
<th>GCGTTCGGCTCCCTACTCCCTCCCGATCGTG</th>
<th>undetected</th>
<th>+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>GCGTTCGGCTCCCTACTCCCTCCCGATCGTG</td>
<td>Mut +3, +5</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>GCGTTCGGCTCCCTACTCCCTCCCGATCGTG</td>
<td>Mut +6 → +8</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>GCGTTCGGCTCCCTACTCCCTCCCGATCGTG</td>
<td>Mut +3 → +8</td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>GCGTTCGGCTCCCTACTCCCTCCCGATCGTG</td>
<td>Mut +5 → +2</td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td>GCGTTCGGCTCCCTACTCCCTCCCGATCGTG</td>
<td>Mut +5 → +2, +3 → +5</td>
<td></td>
</tr>
</tbody>
</table>

than the free probe and with apparent $M_r$ greater than free PTB. No labelled products were observed in the absence of irradiation (data not shown).

PTB cross-linked exclusively to the Y strand but not to the purine-rich (R) strand (Fig. 2). Although the purified protein was made up of only one discrete subunit, we invariably noticed up to three separate bands retarded on the gel, independent of the preparation of PTB used. In contrast, when we tested the PTB deletion PTBΔ337–531, only one complex could be seen. A possible explanation for the appearance of several cross-linked bands following irradiation of PTB complexes may be that there is substantial photodegradation of the protein. To check this, we carried out the cross-linking experiments in the absence of carrier protein (BSA) so as to be able to follow the fate of the protein on SDS–PAGE. Whilst cross-linking was as efficient as in the presence of carrier, no detectable degradation of PTB was observed following irradiation, as demonstrated by the silver staining of the SDS–PAGE analysis (data not shown). These experiments also confirmed that the radioactive retarded bands contained PTB and the probe.

Recombinant PTB protein therefore exhibits a specificity for the pyrimidine-rich sequence as previously demonstrated for the native protein. Of particular interest is the observation that PTB formed no cross-links with the double-stranded DNA probe (Fig. 2). It should be noted that upon overexposure of the gel no retarded bands could be seen, even after extended exposure to PhosphorImager intensifying screen was carried out. This result is in contrast with previous data based upon electrophoretic mobility shift assays. A technical problem common to all of these studies resides in the possibility of there being residual non-hybrid single-stranded oligomers that could interact with PTB to give retarded complexes. To avoid this possibility, we generated the equivalent double-stranded probes from plasmids and again found that there were no cross-linked products with PTB (Fig. 3). Denaturation of the double-stranded DNA labelled at the Y strand, and incubation with PTB followed by irradiation did, however, lead to a labelled retarded band on SDS–PAGE (Fig. 3). Denaturation of the double-stranded DNA labelled at the Y strand, and incubation with PTB followed by irradiation did, however, lead to a labelled retarded band on SDS–PAGE (Fig. 3). As with the assays using single-stranded oligonucleotides as probes, more than one complex can be seen (see legend of Fig. 3). It should be noted that in the experiments shown in Figure 3 the bands migrated slower than those corresponding to the oligomer probes (Fig. 2). This was due to the presence of the plasmid polylinker, increasing the size of the probe to 80 nucleotides (a corresponding increase in mass of 26 kDa). We also failed to observe retarded bands in native gels of
non cross-linked double-stranded plasmid derived DNA in the presence of PTB (data not shown). The inclusion of competitive quantities of single-stranded R strand, single-stranded Y strand and the plasmid derived double-stranded DNA in cross-linking experiments showed that whereas increasing concentrations of unlabelled Y strand oligomer effectively decreased the amount of retarded cross-linked complex (Fig. 4) no effect of the double-stranded plasmid derived fragment was observed. These competition experiments conclusively demonstrate that PTB binds to pyrimidine-rich single-stranded sequences.

**Sequence specificity of binding**

Having ascertained that PTB cross-links exclusively to single-stranded DNA, we then asked the question as to whether there is a requirement for a specific DNA sequence. We synthesized single-stranded oligonucleotides corresponding to the Y strand of the AdMLP initiator region with point mutations within the pyrimidine tract (Table 1). We then carried out cross-linking experiments with PTB and PTBΔ337–531 (Fig. 5) on the single-stranded oligomers. As shown above (Fig. 2) both proteins cross-linked to the wild-type sequence. Binding of both PTB and PTBΔ337–531 to oligonucleotides containing the point mutations at positions +6 to +8 (Fig. 5, mutant M2) is severely reduced. Mutations at other positions selected did not qualitatively appear to interfere with binding although mutations at positions −5 to −2 (M4) gave essentially only two instead of the three retarded bands observed on SDS–PAGE with PTB and the wild-type sequence (Fig. 5). PhosphorImager densitometric analysis of the types of gels presented in Figure 5 produced titration curves as shown in Figure 6. Classical Scatchard type analysis of these curves showed single site non-cooperative binding (Fig. 7). Under the conditions used the apparent $K_d$ was estimated as being that concentration of ligand required to give 50% saturation and correlated well with that calculated from the Scatchard analysis (Fig. 7). A comparison of the relative binding constants obtained from these curves is given in Table 2. PTB and PTBΔ337–531 both bind wild-type pyrimidine tracts with an estimated $K_d$ of ~1–3 nM demonstrating a relatively high degree of specificity. The degree of cross-linking of the truncated protein (PTBΔ337–531) to the wild-type tract was significantly reduced (Fig. 5) although the calculated $K_d$ was only slightly larger (Table 2). This differential in binding was respected and in certain cases accentuated for all the different oligomers used in this study. In all cases, PTBΔ337–531 bound weaker than the wild-type protein.

**Table 2. Apparent binding constants ($K_d$) determined from curves of the type shown in Figure 7**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_d$ (nM)</th>
<th>PTB</th>
<th>PTBΔ337–531</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>1.4</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>28.7</td>
<td>95.1</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>95.6</td>
<td>not measurable</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>34.3</td>
<td>73.4</td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>9.5</td>
<td>21.8</td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td>20.8</td>
<td>63.1</td>
<td></td>
</tr>
</tbody>
</table>

Substrates refer to the oligomers listed in Table 1.

Cross-linking efficiency is known to be a function of the local concentration of pyrimidine and notably thymine bases (16,20). Consequently, the substitution of a reactive base such as a T by a less reactive base such as G should decrease the overall amount of complex cross-linked. However, if, the substitution has no effect on the affinity of the protein for the DNA then it should not alter the apparent $K_d$. Mutations at positions +6 to +8 lead to increases in $K_d$ by a factor of $\leq 100$ (M2 mutation) for PTB and practically abolished binding by PTBΔ337–531. Mutations at other positions produced smaller increases in $K_d$ in the case of PTB but were somewhat larger for PTBΔ337–531. Our results show that for both proteins, binding was most efficient in the order: Wt > M4 > M5 > M1 = M3 > M2. Note that this order does not correlate with the concentration of Ts in the region from −6 to +8. At least with respect to the PTB, it would appear that three regions are important in binding; first, the sequence from +6 to +8, secondly, positions +3 to +5, and finally, the region −5 to −2. This order was respected for the PTBΔ337–531 protein, except that PTBΔ337–531 binding was more severely reduced by the +3 to +5 and +6 to +8 mutations. Paradoxically, whereas mutations in the region +3 to +5 and +6 to +8 independently decreased or abolished binding of both PTB and PTBΔ337–531 respectively, the effects of the presence of both mutations were less dramatic than either alone. Furthermore, the decrease in binding due to mutations in the +3 to +8 region were partially offset by contemporary changes in the −5 to −2 region, even though changes in the −5 to −2 region alone decreased binding affinity. This latter effect was more marked for PTBΔ337–531 than for PTB.

Finally, we constructed labelled single-stranded RNA molecules containing equivalent sequences to the AdMLP wild-type pyrimidine and purine tracts and attempted to laser UV cross-link over the range of PTB concentrations used for the DNA binding
Figure 5. SDS–PAGE of PTB or PTBΔ337–531 cross-linked to different oligomers. PTBNt refers to PTBΔ337–531. The recombinant proteins were cross-linked over a range of concentrations (0–70 nM) to oligonucleotide substrates (5 pM). Wt, M1, M2 and M4 are the sequences listed in Table 1.

studies. Even at elevated concentrations of PTB no retarded RNA bands were observed (data not shown) and we conclude that PTB cannot be cross-linked to RNA using laser UV irradiation.

DISCUSSION

PTB (2,3), previously named PYBP (4), is a 58 kDa polypeptide recently purified, whose cDNA has been cloned and sequenced. In this paper, recombinants PTB and PTBΔ337–531 (Fig. 1A), a deleted mutant containing the N-terminal moiety of the protein, were obtained. The two proteins were overexpressed in bacteria as polypeptides fused with the C-terminal moiety of the GST. After purification of the fusion proteins by affinity chromatography and further thrombin treatment, purification of the recombinant proteins from the carrier GST was performed by a heparin–Sepharose column (Fig. 1B). Since almost all of the DNA-binding proteins bind to this resin, this chromatographic step seems to be a convenient way to separate the carrier protein, which flows through the resin, from the recombinant polypeptides.

SDS–PAGE analysis shows that PTB and PTBΔ337–531 have apparent Mr values that correspond respectively to those calculated from the nucleotide sequences (Fig. 2). Gel filtration under non denaturing conditions suggests that both proteins are monomeric and migrate as expected for globular proteins.

We have previously determined by electrophoretic mobility shift assays that native PTB is able to interact with pyrimidine-rich single-stranded DNA sequences (4). In order to study the specificity of recombinant PTB binding to DNA, we used a single short impulse of laser UV radiation to directly cross-link proteins to DNA (17). As shown in Figure 2, the recombinant PTB cross-links only with pyrimidine-rich DNA single-stranded sequences as already demonstrated for wild-type PTB. In these experiments we used the AdMLP transcription initiator region as a model (Table 1). Figure 5 shows that the PTBΔ337–531 protein is also able to interact with the same AdMLP single-stranded Y sequence. It has previously been demonstrated that the 195 amino acid C-terminal of the recombinant mouse PTB was sufficient for specific binding to pre-miRNAs (1). We performed preliminary experiments with the PTBΔ1–331 protein and noticed that this protein was also able to cross-link to the AdMLP Y sequence (data not shown). Thus, PTB displays more than one potential DNA- and RNA-binding site (21), one situated in the N-terminal moiety and another one in the C-terminal part.

Primary structure information coupled to a predicted secondary structure of the protein, indicate that PTB contains four internal repeats (4). Each of these repeats appears to contain an RNA recognition motif (RRM) present in proteins interacting with RNA or single-stranded DNA (20). Since PTBΔ337–531 as well as PTBΔ1–331 contains two RRM motifs, we could imagine that two RRM sequences may constitute potential nucleic acid binding sites implicated in the interaction of PTB derivatives with the Y stranded DNA sequences. We also produced a truncated version of PTBΔ337–531 retaining only one RRM and observed that this truncated version does not bind the AdMLP initiator (data not shown). The same characteristic is shared by other proteins containing RRMs. For example, U2AF, a genuine splicing factor does not target the intron pyrimidine tract in a functional assay when one RRM is deleted (22), indicating possible cooperation between RRMs.

When using PTB, but not PTBΔ337–531, irradiation of the protein–DNA complex gave three separate bands on analytical SDS gels (Figs 2 and 3); although strikingly, only two bands were seen when PTB was cross-linked with the mutant M4 (positions –5 to –2, Table 1) (Fig. 5). Since we observe, on average, ~10% cross-linking, then the expected amount of retarded material resulting from a second cross-link on the same molecule would be ~0.1%. This is far less than the material present in the different cross-linked bands and we thus conclude that these species result from single cross-linking events. Similar results have been previously reported in other systems (17). One explanation for the appearance of several cross-linked bands following irradiation is that there is photodegradation of the protein. However, no detectable degradation of the
Figure 6. Binding curves derived from SDS–PAGE of PTB cross-linked to different oligomers. The fraction of cross-linked complex produced after laser UV irradiation at various concentrations of PTB protein was estimated from PhosphorImager densitometry of retarded labelled bands after SDS–PAGE. Curves were fit to simple hyperbolas. Wt, M1–M5 are the sequences listed in Table 1.

protein as deduced from silver staining of irradiated complexes was observed, nor was the capacity of irradiated PTB to form retarded complexes on native gels impaired. A second, more plausible explanation for the appearance of several bands is that each band corresponds to the binding of the Y DNA strand to different potential DNA-binding sites of PTB (see above). The PTB-DNA populations thus obtained may adopt different configurations, not disturbed by the presence of SDS in the gel, giving rise to complexes possessing different electrophoretic mobilities.

In contrast with previous data based on electrophoretic mobility shift assays (4), it was not possible to detect any cross-link between PTB and the double-stranded AdMLP DNA probe (Fig. 3). In a former paper (4), and in the experiments presented in Figure 2, the double-stranded probes were generated by annealing single-stranded oligonucleotides. In the experiments shown in Figures 3 and 4, the double-stranded probes used were obtained directly from plasmids. Under these experimental conditions, the existence of residual non-hybrid single-stranded oligomers that would interact with the protein to give retarded complexes is unlikely. Furthermore, the irradiation of PTB in the presence of single-stranded DNA derived from denatured double-stranded DNA labeled at the Y strand, followed by irradiation, again gave characteristic retarded bands on SDS–PAGE (Fig. 3). Finally, significant decreases in the amount of cross-linked material was seen only with Y strand competitor and not with either the R strand single-stranded DNA or with the double-stranded DNA (Fig. 4).
Figure 7. Scatchard analysis of the binding of PTB protein to the wild-type oligomer. Curves of the type shown in Figure 6 were transformed for Scatchard analysis. The percentage cross-linking (y) is related to the amount of complex present (ES) according to the simple function y = λ (ES). The ratio of V to the calculated amount of free protein (L) is represented graphically as a function of L (expressed in nmol on the ordinate axis). The figure shows a representative plot for PTB binding to the wild-type oligomer sequence.

Until now, the DNA or RNA sequences requirements for PTB binding, other than the presence of Y tracts, were still difficult to precise. It was previously shown that the protein interacts with Y-rich sequences within the 3′-splice site of pre-mRNAs, near the branch point of the intron; that it had a higher affinity for uridine residues than for cytidine residues, and that it did not bind to RNAs with Y tracts <10 nucleotides (3,7). Studies using β-tropomyosin pre-mRNA as a model indicated that the interaction of PTB with an RNA requires more than a sequence of consecutive pyrimidines (23). Recently, it was demonstrated that PTB has a more distinct RNA-binding specificity than both the human splicing factor U2AF65 and the drosophila splicing regulator Sex-lethal (Sxl) (6). We and others (4,5) have reported that PTB binds specifically to the Y-rich strand of at least three regulatory cis-elements of hepatic genes. Using one of these elements, we have demonstrated that PTB interacts weakly with a derivative where only one Y tract has been mutated, and does not interact with another mutant, in which all Y tracts have been removed (4). In order to better characterise the DNA sequence specificity of PTB binding, we carried out cross-linking assays using several single-stranded oligomers containing mutations within the Y region of the AdMLP initiator region (Table 1). We were able to quantitatively assess apparent binding affinities to each of the mutated sequences for both PTB and PTBΔ531 as shown in Table 2. Mutations at positions +6 to +8 (M2) effectively abolished binding whereas mutations from positions –5 to –2 (M4, M5) also severely decreased binding but to a lesser extent, as was also observed for mutations in positions +3 to +5 (M1). These results indicate that even if the nucleotides 6, 7 and 8 of the +3 to +8 region are critical for binding, a second region comprising nucleotides –5 to –2 may be involved in the interaction of the AdMLP Y strand with PTB. This may explain the observation that PTB produced at least three bands after cross-linking to the wild-type DNA site, due to the presence of a mixed population containing molecules of PTB cross-linked at different positions on the protein (as discussed above), but also at different positions on the DNA. Furthermore, this is consistent with the fact that the N- and C-terminal moieties of PTB were shown to cross-link to pyrimidine-rich sequences (this paper; 1; and our unpublished results). It is thus extremely interesting to note that only two main retarded bands were observed when PTB was cross-linked to the –5 to –2 mutated sequence. Coupled with the observations that the truncated protein (PTBΔ337–531) only gave one cross-linked species on SDS–PAGE and that mutations of the oligomers in the region +3 to +8 severely restricted or totally abolished binding of PTBΔ337–531, this would suggest that the N-terminal part of PTB is involved in contacting this part of the pyrimidine-rich sequence and that the C-terminal part contacts the –5 to –2 region. Mutations in a nucleic acid sequence that binds on a protein surface clearly change not only the availability of a given base to make specific Van der Waals contact with an amino acid but also perhaps the path of the nucleic acid on that surface. Although the data presented here demonstrate for the first time that the sequences in question confer a high degree of specificity for an intimate contact with PTB, they do not allow determination of those bases in contact in a defined context.

In conclusion, since the AdMLP displays a transcriptional initiator sequence whose structural feature is a high content in pyrimidine residues, we were prompted to study PTB interaction on this element. Using DNA-binding assays, we demonstrated in fact that PTB interacts with a high affinity and in a sequence-specific manner with the pyrimidine-rich strand of the transcriptional initiator (Table 2 and Figs 5 and 6). We were also able to dissect this interaction to suggest that the N-terminal part of PTB (PTBΔ337–531) interacts with bases +3 to +8 and the C-terminal with bases –5 to –2 on the pyrimidine-rich strand of the transcriptional initiator.

Laser UV cross-linking, by definition, reports only residues in intimate contact, thus our inability to cross-link PTB to RNA sequences (see above), suggests that the nature of the interaction of PTB with RNA is different to that with DNA. We have demonstrated a strong local sequence requirement for binding to DNA, and this seems to be different for the interactions with RNA. This does not necessarily preclude the same domain on PTB being involved in both interactions but may suggest that the more intimate DNA binding is due to PTB having to bind in vivo to locally melted regions of a DNA molecule.

ACKNOWLEDGEMENTS

F.B. and M.M.Z. were supported by the CNRS (URA 1129) and by the Commission of European Communities (CEC contract No BIO2-CT93-0319). H.B. and M.B. were supported by a grant from the INSERM (No 91-0902). We are grateful to Mme O. Delpech and E. Croullebois for their secretarial assistance and would like to thank members of the Unité d’Expression des Gènes Eucaryotes for fruitful discussions and support during completion of this work.

REFERENCES