Protein Engineering vol.10 no.8 pp.935–941, 1997

Use of protein A gene fusions for the analysis of structure–function relationship of the transactivator protein C of bacteriophage Mu

Amitabha De[1], B.D.Paul, V.Ramesh[2] and V.Nagaraja[3]

Centre for Genetic Engineering, Indian Institute of Science, Bangalore, 560 012, India

[1]Present address: National Diary Development Board, Anand 388 001, India
[2]Present address: Department of Biology, Massachussets Institute of Technology, Cambridge, MA, USA

[3]To whom correspondence should be addressed

Authors A.De and B.D.Paul have equal status

A sensitive dimerization assay for DNA binding proteins has been developed using gene fusion technology. For this purpose, we have engineered a gene fusion using protein A gene of *Staphylococcus aureus* and *C* gene, the late gene transactivator of bacteriophage Mu. The *C* gene was fused to the 3′ end of the gene for protein A to generate an A–C fusion. The overexpressed fusion protein was purified in a single step using immunoglobulin affinity chromatography. Purified fusion protein exhibits DNA binding activity as demonstrated by electrophoretic mobility shift assays. When the fusion protein A–C was mixed with C and analyzed for DNA binding, in addition to C and A–C specific complexes, a single intermediate complex comprising of a heterodimer of C and A–C fusion proteins was observed. Further, the protein A moiety in the fusion protein A–C does not contribute to DNA binding as demonstrated by proteolytic cleavage and circular dichroism (CD) analysis. The assay has also been applied to analyze the DNA binding domain of C protein by generating fusions between protein A and N- and C-terminal deletion mutants of C. The results indicate a role for the region towards the carboxy terminal of the protein in DNA binding. The general applicability of this method is discussed.

Keywords: Dimerization/DNA binding protein/Mu C/protein A gene fusion

Introduction

Gene fusions have now become convenient tools to address several important questions in molecular biology. Gene fusion technology has been employed when other means of overexpression fail to produce the desired protein in sufficient quantities. Further, facile purification of the protein in soluble and active form can be achieved in a single step involving specific affinity chromatography. In some cases, the properties of the protein in question can easily be assayed by choosing the appropriate fusion partner. To meet these requirements, several fusion vectors (Uhlen and Moks, 1990) have been developed for the expression of various proteins in *Escherichia coli*. Among them, glutathione *S*-transferase (GST; Smith and Johnson, 1988), maltose binding protein (MBP; di Guan *et al*., 1988) and protein A (Nilsson *et al*., 1985) gene fusion systems are widely employed.

In this paper, we describe yet another application for gene fusion technology. A dimerization assay has been developed and has been applied to studies on the structure–function relationship of DNA binding proteins. The protein under study is the C protein, a site-specific DNA binding protein, which activates transcription of late genes of bacteriophage Mu (Heisig and Kahmann, 1986). Earlier, we had engineered the hyperexpression of the protein by the removal of a secondary structure at the translational initiation region in the *C* mRNA (Ramesh *et al*., 1994a). The overexpressed protein has been purified both by DNA affinity as well as anti-C IgG affinity chromatography. The purified recombinant protein is composed of 138 amino acids and binds to its cognate site with a very high equilibrium constant, of the order of 2.0×10^{12} M^{-1} (Ramesh *et al*., 1994a,b). Obtaining large quantities of protein has enabled us to initiate structure–function studies such as the determination of the oligomerization and DNA binding domain of the C protein. To assess the oligomerization status of C protein an assay involving the association of C protein with a truncated derivative of C was considered, as in the case of GCN4 (Hope and Struhl, 1987). However, in the absence of appreciable differences in size between C protein and deletion mutants of C, unlike GCN4 and its truncated products, gel electrophoresis technique is inadequate for clear resolution of protein–DNA complexes to assess dimerization. To overcome these problems, we have engineered the construction of fusion proteins to evaluate the quaternary structure of the C protein in solution and when bound to DNA. Protein A (Surolia *et al*, 1982) of *Staphylococcus aureus* was chosen as the fusion partner. The results presented here show that the C protein binds to DNA as a dimer and the region towards the C-terminal is involved in DNA binding. The method employed has the potential to be utilized extensively in structure–function analysis of proteins, especially those of small sizes.

Materials and methods

Restriction enzymes and DNA modifying enzymes were purchased from Boehringer Mannheim, New England Biolabs and Pharmacia and used according to the manufacturers' recommendations. Labeled nucleotides were from Bhabha Atomic Research Centre, India. Plasmid pRIT2T, *E.coli* strain N4830-1 and IgG Sepharose 6 FF and Heparin Sepharose CL6B were from Pharmacia. Other reagents were from Sigma.

Construction of plasmids

A 0.8 kb *Bam*HI–*Pst*I fragment containing the *C* gene derived from plasmid pVN8 (Balke *et al*., 1992) was cloned at the corresponding sites in the multiple cloning site located 249 codons after protein A gene in the vector pRIT2T (Nilsson *et al*., 1985) to obtain the A–C fusion plasmid pVNIII. The *C* gene deletion mutants, ΔN18ΔC13 and ΔN18ΔC27 were generated using the polymerase chain reaction. The former lacks the first eighteen and the last thirteen amino acids. The latter lacks the first 18 and the last 27 amino acids. The template used for the generation of the mutants were pVR125 and pVR111, respectively, which are pKK223-3 derivatives.

pVR125 harbors the *C* gene which is deleted for the last 13 amino acids and pVR111 has the *C* gene deleted for the last 27 amino acids. The forward primer used for deletion of the first 18 amino acids had the sequence 5′-ATCGGACATATGG-ATAACATCCC-3′. The reverse primer used anneals downstream of a *Hin*dIII site in vectors pVR125 and pVR111 and has the sequence 5′-CAGGCTAAAATCTTCTC-3′. To generate fusions to protein A gene, the PCR products were first cloned into pUC19 as follows. The PCR products were filled in using the Klenow fragment of DNA polymerase I of *E.coli* and digested with *Hin*dIII. The vector pUC19 was digested with *Pst*I, blunted with T4 DNA polymerase and digested with *Hin*dIII. The processed PCR products were then ligated to the digested vector to generate truncated versions of C. Next an *Eco*RI–*Pst*I fragment (the PCR product has a *Pst*I site at the 3′ end) was released from the recombinants and cloned into the corresponding sites of pRIT2T in frame with the gene for protein A resulting in the formation of two biterminal deletion mutants as protein A fusions, namely, A-CΔN18ΔC13 and A-CΔN18ΔC27. All the DNA manipulations were according to standard procedures (Sambrook *et al*., 1989).

Induction and expression analysis

Escherichia coli N4830-1 and *E.coli* BL26(DE3) were used as hosts for λP_R (pRIT2T,pVNIII, A-CΔN18ΔC13 and A-C ΔN18ΔC27) and phage T7 promoter (pVR7) based expression vectors, respectively. In the case of λP_R based vectors, transformants were grown at 30°C in Luria Bertani (LB) medium containing 100 μg/ml ampicillin to a cell density of 0.6 A_{600} and induced for 15 min at 42°C followed by incubation for 45 min at 40°C. pVR7 transformants were grown at 37°C in LB containing 100 μg/ml ampicillin, to a cell density of 0.6 A_{600}, induced with 0.3 mM isopropyl thiogalacto-β-D-pyranoside (IPTG) for 2 h. Cells were harvested, lysed and run on a 12% (w/v) SDS–PAGE (Laemmli, 1970) and analyzed by Coomassie blue staining.

Purification of the proteins

Cells from a 1 l culture were harvested and washed with buffer 1 [20 mM Tris–HCl, pH 7.5, 5 mM $MgCl_2$, 1 mM ethylenediamine tetraacetic acid (EDTA) and 5% v/v glycerol]. The cells were suspended in 10 ml buffer 1 and disrupted by sonication and centrifuged for 30 min at 20 000 *g* (S20 fraction). The S20 supernatant was further centrifuged at 100 000 *g* for 3 h at 4°C. Nucleic acids were removed by polyethylenimine precipitation (1% v/v) followed by centrifugation for 10 min at 10 000 *g*. The resulting supernatant was subjected to ammonium sulfate fractionation (0–45% w/v) and centrifuged for 10 min at 10 000 *g*. Precipitated protein was dissolved in 100 mM Tris–HCl (pH 8.0). The solution was loaded on a pre-equilibrated IgG Sepharose column (bed volume 3 ml) and washed with 30 ml of 100 mM Tris–HCl (pH 8.0). Protein A or fusion protein A–C was then eluted with 0.1 M Glycine–HCl (pH 3.0). One ml fraction volumes were collected and neutralized using 1 M Tris–HCl, pH 8.0 and dialysed against buffer 1. C protein purification was as described earlier using anti-C IgG affinity chromatography (Ramesh *et al*., 1994b). In the case of the deletion mutants, the S100 supernatants were passed through an IgG Sepharose column, followed by anti-C IgG Sepharose column. The eluted proteins were further loaded onto a Heparin Sepharose column and eluted with high salt. The purity of the proteins were checked on a 10% (w/v) SDS polyacrylamide gel.

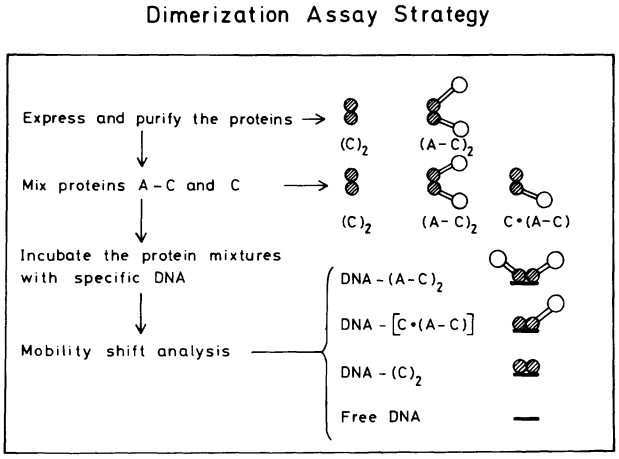

Fig. 1. Schematic representation of the dimerization assay strategy. (**Top**) The presence of both C-protein and fusion protein A–C in dimeric form. (**Middle**) Three types of protein molecules. (**Bottom**) The different protein–DNA complexes are depicted.

Gel filtration chromatography

FPLC Hiload^{TM} 26/60 Superdex ^{TM}75 gel filtration column (bed volume 24 ml) was equilibrated with buffer 2 (20 mM Tris–HCl, pH 7.5, 150 mM KCl, 5 mM $MgCl_2$ and 7 mM 2-mercaptoethanol). Purified protein solutions (100–300 μg in 200 μl) were loaded onto the column. In cases where mixture of proteins were loaded, preincubation was carried out for 1 h at room temperature and elution was performed with the same buffer at a flow rate of 0.8 ml/min.

Electrophoretic mobility shift assay (EMSA)

EMSA was carried out with a labeled 220 bp *Eco*RI–*Bam*HI fragment of pUW4 bearing the C-binding site (Ramesh *et al*., 1994a). The proteins were incubated separately with labeled DNA for 10 min at 4°C in buffer 3 (20 mM Tris–HCl, pH 8.0, 5.0 mM $MgCl_2$, 70 mM KCl, 20 mM NaCl, 0.5 mM EDTA, 1.0 mM Dithiothreitol (DTT), 3.0% (v/v) glycerol). In cases where trypsin was added to preformed complex, incubation was carried out for an additional 30 min at 37°C in presence of 1 mM $CaCl_2$. The samples were electrophoresed either on a 4% (w/v) polyacrylamide gel for 6 h or a 1% (w/v) agarose gel for 2 h in 0.5× TBE (45 mM Tris-borate, 1 mM EDTA) buffer. In the case of the deletion mutants, electrophoresis was carried out on a 2% agarose–2% acrylamide (w/v) composite gel (Topol *et al*., 1985) at 4°C at 100 V for 2 h in 0.5× TBE.

Circular dichroism (CD) analysis

A JASCO model J500A spectropolarimeter was used for recording the CD spectrum. Samples were prepared at room temperature in buffer 3. Ten μM of 220 bp *Eco*RI–*Bam*HI fragment bearing the C-binding site was incubated with 10 μM C protein, Protein A or protein A–C and analyzed. Sonicated calf thymus DNA served as non-specific DNA control and did not exhibit any change in spectrum in the presence of protein. Concentration of DNA was calculated by A_{260}. Protein concentration was estimated using Bradford reagent (Bradford, 1976).

Results

Dimerization assay strategy

The schematic representation of the dimerization assay strategy for the DNA binding proteins is depicted in Figure 1. The assay relies on the heterodimerization of two proteins having significantly different molecular masses, resulting in the forma-

tion of complexes of different sizes, which can be resolved on native gels. The larger protein in the heterodimerization reaction is a fusion protein between the smaller protein (the protein being characterized) and a protein which serves to increase the molecular mass of the protein. Dimerization can generate homodimers of both the protein of interest and the fusion protein. It can also result in the formation of a heterodimer when a monomer of the protein under study associates with a fusion protein monomer. The size of the fusion protein being distinctly larger than that of the protein of interest, the detection of a heterodimer is facilitated. The following criteria were considered for the selection of an appropriate fusion partner. (i) The fusion partner should not bind DNA as it would interfere with the assay system. (ii) It should exist in a monomeric form since the assay has been designed to evaluate the oligomerization status of the protein under study. (iii) It should serve as an 'affinity handle' (Nilsson and Abrahamsen, 1985) for facile purification. Amongst the variety of fusion systems available, protein A system was chosen for constructing a fusion as it satisfies the above mentioned criteria and has several advantages. Firstly, protein A fusions can be purified to >95% yields using immunoglobulin G (IgG) affinity chromatography (Nilsson *et al*., 1985), as protein A binds to the F_c portion of IgG with a very high affinity (Lindmark *et al*.,1983). Secondly protein A has higher solubility over GST and MBP. Thirdly, the junction of the two proteins is flexible (Nilsson and Abrahmsen, 1990) and as a result the two proteins in the fusion are likely to reflect their individual properties and chance of interfering with each other properties is minimal.

Expression and purification of proteins

In order to test the above scheme, we constructed a gene fusion between transactivator C protein and protein A. Plasmid pRIT2T has the *S.aureus* protein A gene under transcriptional control of the inducible λ P_R promoter (Figure 2A). The construct was transformed into *E.coli* N4830-1, a strain which harbors the temperature sensitive λcI857 repressor. Expression of protein A and fusion protein A–C could be achieved by heat induction which results in inactivation of the thermolabile lambda repressor. The overexpressed protein A and protein A–C were purified using an IgG Sepharose 6FF column (Figure 2B, lanes 4 and 6). The overproduced C protein was purified by anti-C IgG affinity chromatography (Figure 2B, lane 2). The affinity purified C protein, protein A and fusion protein A–C have the expected molecular masses. The 46.5 kDa size observed for fusion protein A–C corresponds to the fusion of a 30 and 16.5 kDa protein (Figure 2B, lane 6).

The fusion protein forms heterodimers in solution

The oligomerization status of the purified proteins was determined by gel filtration chromatography. Protein A was found to be present only in monomeric form (Figure 3B). On the other hand, both C protein and fusion protein A–C are present in solution in a dynamic equilibrium of monomeric (peak II for C protein and peak V for protein A–C) and dimeric (peak I for C protein and peak IV for protein A–C, respectively) forms (Figure 3a and c). A new peak (VI) with a retention time of 7 min corresponding to a molecular mass of 60 kDa was observed when a mixture of protein A–C and C protein were chromatographed, indicating the specific interaction between the proteins (Figure 3e). Absorbance profile of protein mixtures containing protein A and C protein does not reveal the presence of any (Figure 3d) new peaks with a different

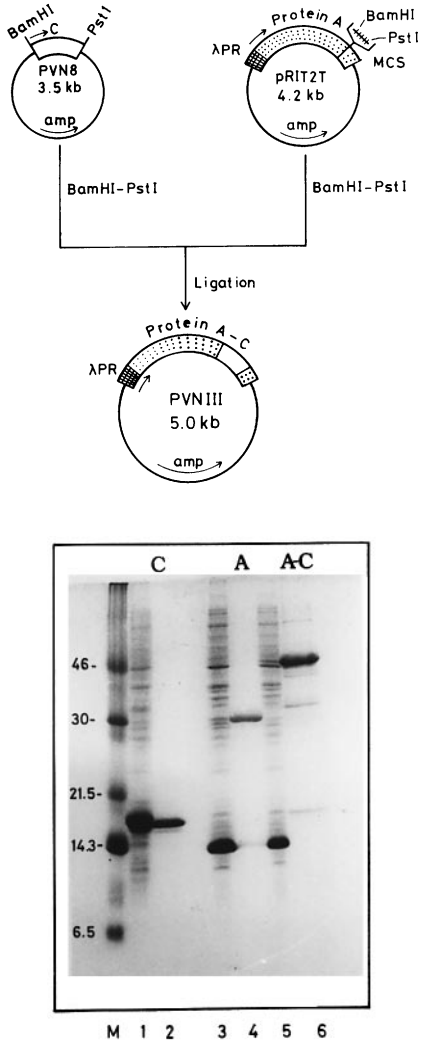

Fig. 2. (**A**) Construction of pVNIII plasmid. Boxes show the positions of different genes; C protein gene (open), protein A gene (stippled), λP_R promoter (cross hatched). *Bam*HI and *Pst*I restriction sites are present in multiple cloning site (MCS). (**B**) SDS–PAGE analysis of the expressed proteins. Molecular weight marker with sizes in kDa is shown on the left margin; lane 1, crude extracts of induced *E.coli* BL26(DE3) harboring pVR7; lane 2, purified C protein eluted from anti-C IgG Sepharose column; lane 3, crude extracts of induced *E.coli* N4830-1 harboring pRIT2T; lane 4, purified protein A eluted from IgG Sepharose 6FF; lane 5, crude extracts of induced *E.coli* N4830-1 harboring pVNIII; lane 6, purified A–C protein eluted from anti-C IgG Sepharose 6FF column.

retention time. The molecular weight of protein present in different peaks were determined by using standard protein markers. The peak fractions corresponding to those shown in Figure 3a–e were analyzed on a 12% (w/v) SDS–PAGE to identify the component proteins (data not shown). Peak VI in Figure 3e contains two proteins with molecular mass of 46 and 16.5 kDa corresponding to A–C and C, respectively, indicating that C protein can form a heterodimer with protein A–C. The other peak fractions contained the individual proteins. Under all experimental conditions, no other heterodimeric species were observed.

DNA binding studies with the A–C fusion protein

The purified proteins eluted from the affinity column were dialyzed and checked for their DNA binding activity in a gel retardation assay using a fragment bearing the C-binding site. The C protein forms a specific complex with the DNA (Figure

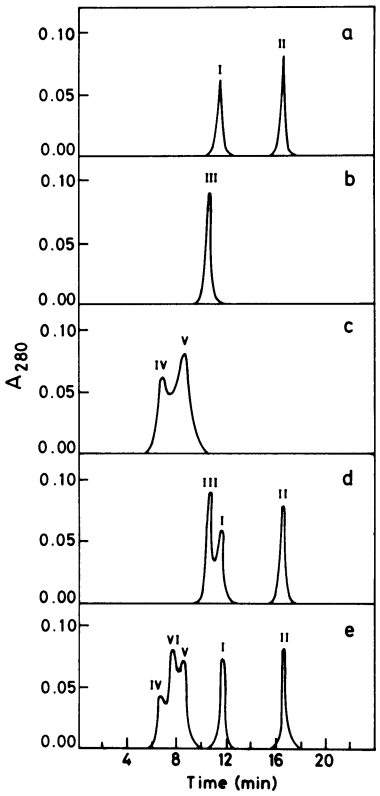

Fig. 3. FPLC analysis of C protein, protein A and fusion protein A–C. (**a**) 100 µg C protein; (**b**) 200 µg protein A; (**c**) 300 µg fusion protein A–C; (**d**) Mixture of 200 µg protein A and 100 µg C protein; (**e**) mixture of 300 µg A–C protein with 100 µg C protein.

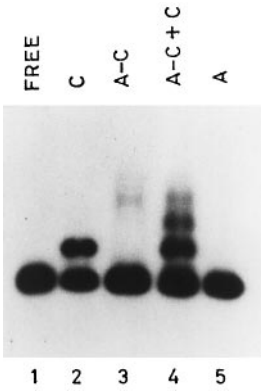

Fig. 4. Analysis of DNA binding with C, A and A–C proteins on a 1% (w/v) agarose gel. The end-labeled 220 bp DNA fragment (140 pmol, 10 000 c.p.m.) was incubated with different proteins. Lane 1, free DNA; lane 2, 10 ng C protein; lane 3, 30 ng protein A–C; lane 4, 30 ng protein A–C mixed with 10 ng C protein; lane 5, 200 ng protein A.

4, lane 2; Ramesh *et al*, 1994a), whereas protein A fails to bind DNA (Figure 4, lane 5). The fusion protein A–C also forms a specific complex with DNA (Figure 4, lane 3) albeit with a lower affinity. The equilibrium binding constant of protein A–C binding to DNA was determined as described previously (Ramesh *et al*, 1994a) and is at least 3 orders of magnitude lower ($1x10^9$ M^{-1}) than that of the C protein. This could be as a result of attachment of a protein of large size. Next, we tested the ability of heterodimer to bind to DNA. When a mixture of C and A–C was used in the assay, a single intermediate complex was observed. With increasing amounts of protein A–C added, increased complex formation was

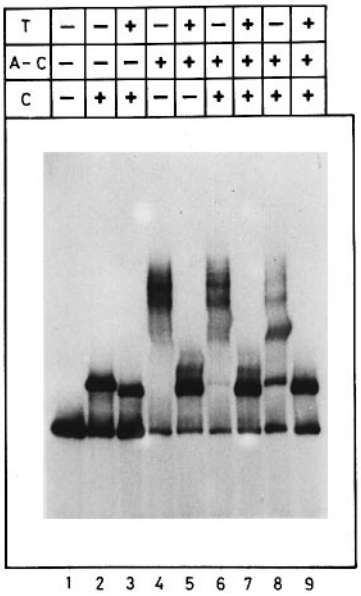

Fig. 5. Trypsin digestion of protein when complexed to DNA. The 220 bp end-labeled DNA (140pmol, 10 000 c.p.m.) fragment was incubated with proteins as specified and electrophoresed on a 4% (w/v) native polyacrylamide gel. Lane 1, free DNA; lanes 2 and 3, 10 ng C protein; lanes 4 and 5, 90 ng protein A–C; lanes 6 and 7, 90 ng protein A–C + 2 ng protein C; lanes 8 and 9, 90 ng protein A–C + 10 ng protein C; lanes 3, 5, 7 and 9, preformed complexes treated with 20 ng trypsin.

observed in addition to C and A–C specific complexes (Figure 4, lane 4).

The appearance of the intermediate complex reflects the binding of (A–C).C heterodimer to DNA in accordance with the strategy depicted in Figure 1.

C protein alone is responsible for binding to DNA in A–C fusion

The above results suggest that the DNA binding property of C is responsible for the binding of the A–C fusion to DNA. This point was further substantiated by proteolytic digestion experiments. Trypsin treatment on preformed DNA–C protein complex results in the generation of a faster moving complex (Figure 5, lane 3). Trypsin digestion of free protein, followed by addition of DNA does not produce any complex. Trypsin treated protein (A–C)–DNA complex (lane 5) moves with the same mobility as that of trypsin treated DNA–C protein complex. Further, upon trypsin treatment, the intermediate complex (lanes 6 and 8) shows an identical pattern (compare lanes 7 and 9 with lane 3), indicating that only C protein is responsible for binding to DNA. The molecular weight of trypsin resistant products in all the cases was 11 kDa as determined by SDS–PAGE (data not shown).

Study of interaction of C protein with its cognate site using circular dichroism

Conformational changes occurring in DNA due to the binding of C protein to its cognate site was detected by CD studies. The spectra were recorded at room temperature from 240–320 nm. Each spectrum is an average of four recordings stored and normalized in a data processor. The fragment containing C protein binding site exhibits a positive ellipticity around 260 nm (Figure 6). In presence of C protein, ellipticity of DNA at peak position is altered (Figure 6a). The pattern obtained is characteristic of DNA binding proteins and agrees with those reported earlier (Torigoe *et al.*, 1991). A similar

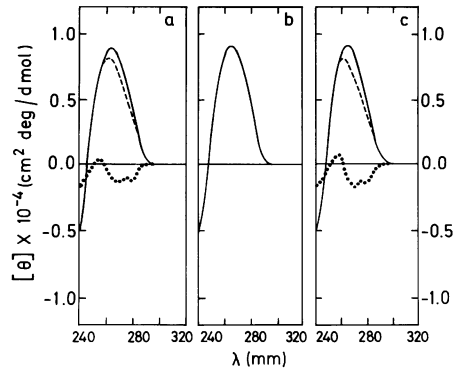

Fig. 6. CD spectrum of protein–DNA interaction. The DNA fragment bearing the C-binding site (10 μM) was incubated with 10 μM of either C protein, protein A or protein A–C in buffer 3 and analyzed (panel a, b and c, respectively). The uninterrupted curve represents the spectrum in the absence of protein. The dashed lines represent the spectrum in the presence of protein. The dotted line represents the difference spectrum.

alteration in spectrum was observed when the fusion protein A–C was added instead of C (Figure 6c), while the characteristic spectrum of DNA was unaltered with protein A (Figure 6b). CD studies were also done with calf thymus DNA which lacks a C-binding site both in the presence and in the absence of C protein. The spectra of DNA was similar in both the cases (data not shown). The decrease in the positive ellipticity, θ, (Figure 6a and c) indicates that DNA undergoes a conformational change upon complex formation both with C protein and fusion protein A–C. The difference spectra indicates that it is a specific DNA–protein interaction (Torigoe *et al.*, 1991). These data further establish that C protein moiety in the A–C fusion is involved in DNA binding and protein A does not participate in this process.

Determination of the DNA binding region

The assay described above has been applied to determine the DNA binding region of C protein. Analysis of the derived amino acid sequence of the C protein had revealed the presence of a helix–turn–helix (HTH) motif located towards the C-terminal of the protein (Figure 7a, Bolker *et al.*, 1989),

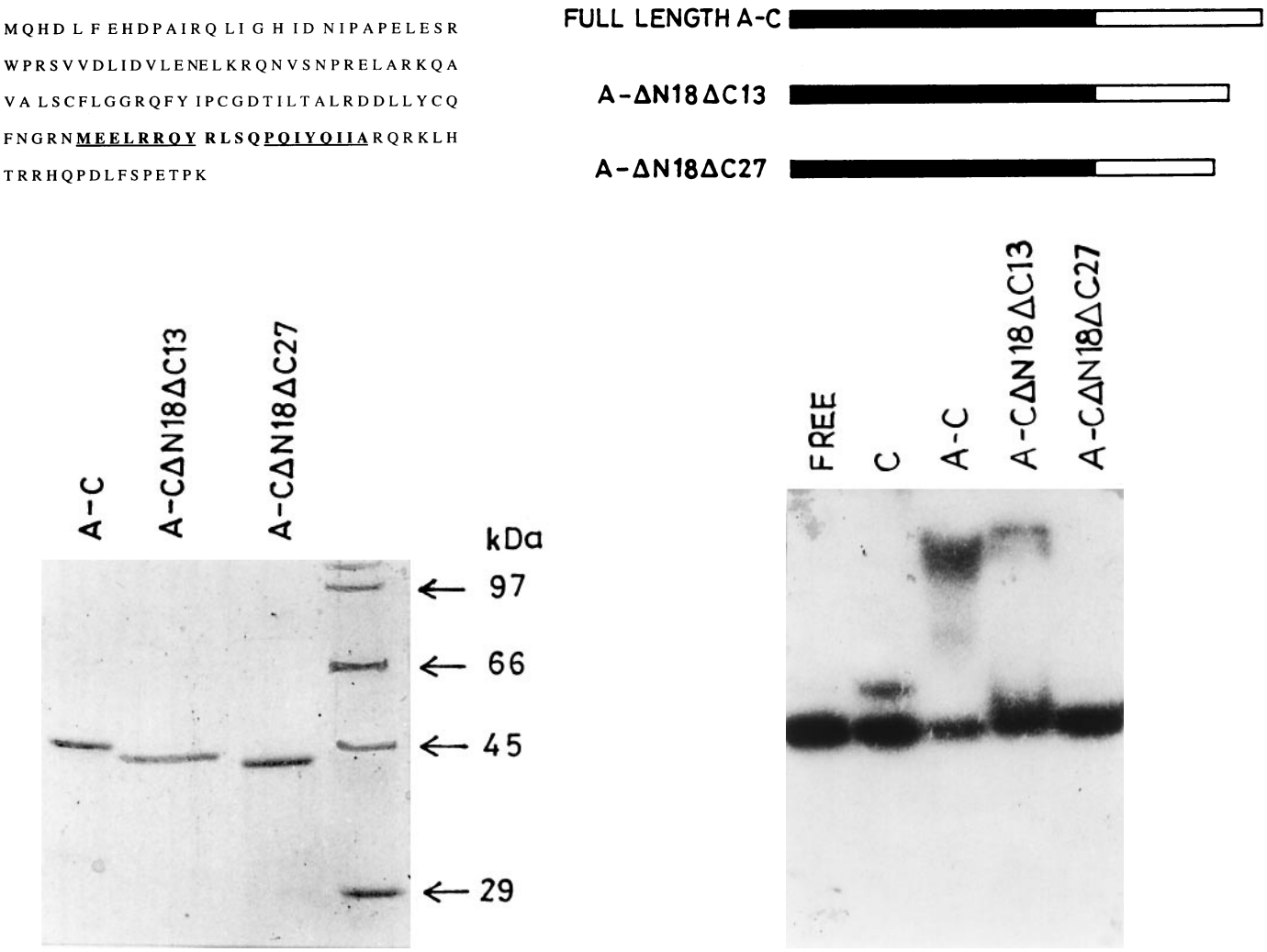

Fig. 7. (**A**) The derived amino acid sequence of C protein. The helix–turn–helix motif is indicated in bold. The component helices are underlined. (**B**) Schematic representation of the protein A fusions of biterminal deletion mutants of C. The solid rectangular region represents the protein A moiety. The open rectangular region represents C or truncated versions of C. (**C**) Purified mutant A–C fusion proteins. 100 ng of each of the A–C fusion proteins were incubated with a labeled 220 bp *Eco*RI–*Bam*HI fragment (10000 c.p.m.) bearing the C-binding site on ice for 10 min. The reaction mixes were electrophoresed on a 2% acrylamide, 2% agarose (w/v) composite gel at 100 V for 2 h.

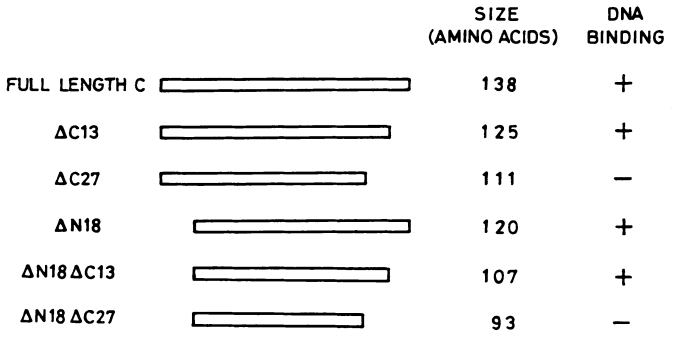

Fig. 8. Schematic representation of the deletion mutants of C and their DNA binding activity.

which is a candidate motif for DNA binding. In order to assess its role in DNA binding, N- and C-terminal deletion mutants of C were generated, namely, ΔN18, ΔC13 and ΔC27 using the polymerase chain reaction. ΔN18 is an N-terminal deletion mutant lacking the first eighteen amino acids. ΔC13 and ΔC27 are C-terminal deletion mutants lacking the last 13 and the last 27 amino acids, respectively. However, purification of these proteins posed problems, therefore DNA binding activity was assayed in crude extracts. ΔN18 and ΔC13 bind DNA, whereas ΔC27 does not bind DNA (data not shown). Next, we generated biterminal deletion mutants to determine the minimal essential region of C required for DNA binding. The biterminal mutants constructed lacked the first 18 amino acids as this region was not involved in DNA binding. The deletion mutants generated either formed inclusion bodies or were not stable. Therefore protein A fusions of these deletions mutants were considered. Two of these biterminal deletion mutants were constructed as fusions to protein A (Figure 7b). The biterminal deletion mutant A–CΔN18ΔC13, has an intact HTH motif whereas A–CΔN18ΔC27 has a part of the motif deleted. These fusion proteins were expressed and purified by IgG chromatography (Figure 7c). The purified proteins were assayed for their DNA binding activity. The mutant A–CΔN18ΔC13 binds DNA (Figure 7d, lane 4), whereas the mutant A–CΔN18ΔC27 does not bind DNA (Figure 7d, lane 5). Although A–CΔN18ΔC13 is smaller in size than A–C (Figure 7c), the complex shows anamolous mobility. This could be a property of the mutant protein, although the protein migrates according to its size on a SDS–polyacrylamide gel (Figure 7c). Figure 8 summarizes the DNA binding activities of the deletion mutants of the C protein. Our results show that (i) the first 18 amino acids are not involved in DNA binding. (ii) The last 13 amino acids are also dispensable for DNA binding. (iii) Deletion of the last 27 amino acids abolish DNA binding, implicating a role for the HTH motif at the C-terminal of the C protein. (iv) Biterminal deletion mutants generated as protein A fusions confirm these results and also facilitate localization of the minimal region required in protein–DNA interaction. (v) Protein A fusions can be effectively employed to investigate the DNA binding activity of deletion mutants of the protein under study.

Discussion

In this paper, we have presented a strategy to assay the dimerization status of DNA binding proteins. A gene fusion between protein A and the DNA binding protein C was constructed and used for this purpose. Gel filtration experiments with the purified proteins and subsequent EMSA show that C protein exists as a dimer in solution and when bound to DNA. Further, the C protein is solely responsible for dimerization and DNA binding activity of the fusion protein. This conclusion is based on the size of the protease resistant complex in gel mobility shift assays and CD studies. Both the shift in DNA peak and the resulting difference spectra (Fig. 6a and c) indicate specific protein–DNA interaction resulting in conformational changes in DNA. The characteristic shift is identical to both C and A–C fusion protein. It should be noted here that protein A does not bind to the specific DNA sequence employed in these experiments. Binding of protein A to other DNA, such as calf thymus DNA, lambda DNA, pBR322 DNA, poly[d(G–C)] and poly[d(A–T)] is also not observed (data not shown). Further, protein A is a monomeric protein (Surolia *et al.*, 1982) in solution and remains in that form under all the experimental conditions described in this paper. We have also used fusions of protein A with deletion mutants of C protein to determine the region of C involved in DNA binding. The resulting fusion proteins were obtained in soluble form in contrast to the deletion mutants generated earlier which formed inclusion bodies. Moreover, the protein A facilitates purification using IgG chromatography. As a result, protein A gene fusion system is suitable to analyze oligomerization status and DNA binding potential of any regulatory protein. The gene fusion technique has been widely used for the expression of genes and purification of the expressed proteins. Here, we show that gene fusion technique can be applied effectively for dimerization assay and to assess DNA binding activity of a regulatory protein. The method may have its own limitations when dealing with proteins having the DNA binding domain at the N-terminal end. However, in the case of protein A fusions, the junction comprises a flexible linker (Nilsson and Abrahmsen, 1990). It is clear from the results that this approach could have wider applications for studying oligomerization potential of different classes of proteins. The method becomes particularly useful for small proteins where full length and truncated proteins may not have appreciable difference in sizes (as in the case of C protein) and as a result the DNA–protein complex will have very insignificant difference in mobility compared with the authentic complex in EMSA. The technique can easily be adapted to test the dimerization potential of minimal regions of polypeptides. In spite of tremendous developments in expression technology, expression of small polypeptides is still a major problem, due to instability caused by protease degradation (Uhlen and Moks, 1990). Further, purification of small polypeptides (<40 residues) from crude cell extracts poses additional problems. Once expressed and purified as fusion proteins, the peptides could be assayed directly or after protease cleavage at the junction. Protein A fusion may have added advantage in that fusion proteins are frequently in soluble form (Uhlen and Moks, 1990). Our findings agree with the earlier observations as protein A–C fusions are soluble. Further, the elution of the protein at low pH does not seem to affect the renaturation of fusion protein. Apart from the applications described here, the method could be conveniently applied to assay other protein–protein, enzyme–substrate/inhibitor, receptor–ligand interactions.

Acknowledgements

We thank Molecular Biology Unit, Banaras Hindu University for use of their CD facility. A.D. is recipient of Department of Biotechnology Postdoctoral fellowship. B.D.P is a recipient of the Council of Scientific and Industrial Research (CSIR) fellowship. V.R. was supported by Jawaharlal Nehru Centre

for Advanced Scientific Research. S.Basak and D.R.Radha are gratefully acknowledged for help in some of the experiments. This work is funded by a grant from Department of Science and Technology, Government of India.

References

Balke,V., Nagaraja,V., Gindlesperger,T. and Hattman,S. (1992) *Nucleic Acids Res.*, **20**, 2777–2784.
Bolker,M., Wulczyn,F.G. and Kahmann,R. (1989) *J. Bacteriol*., **171**, 2019–2027.
Bradford.M. (1976) *Anal. Biochem.*, **72**, 248–254.
DiGuan,C., Li,P., Riggs,P.D. and Inouye,H. (1988*) Gene*, **67**, 21–30.
Heisig,P. and Kahmann,R. (1986) *Gene*, **43**, 59–67.
Hope,I.A. and Struhl,K. (1987) *EMBO J*., **6**, 2781–2784.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Lindmark,R., Thoren-Tolling,K. and Sjoquist,J. (1983) *J. Immunol. Methods*, **62**, 1–13.
Nilsson,B. and Abrahmsen,L. (1990) *Methods Enzymol*., **185**, 144–161.
Nilsson,B., Abrahmsen,L. and Uhlen,M. (1985) *EMBO J*., **4**, 1075–1080.
Ramesh,V., De,A. and Nagaraja,V. (1994a) *Protein Engng.*, **7**, 1053–1057.
Ramesh,V., De,A. and Nagaraja,V. (1994b) *Protein Exp. Pur.*, **5**, 379–384.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press.
Smith,B.D. and Johnson,K.S. (1988) *Gene*, **67**, 31–40.
Surolia,A., Pain,D. and Khan,M.I. (1982) *Trends Biol. Sci*., **7**, 74–76.
Topol,J., Ruden,D.M. and Parker,C.S. (1985) *Cell*, **42**, 527–537.
Torigoe,C., Kidokoro,S., Takimoto.M., Kyogoku,Y. and Wada,A. (1991) *J. Mol. Biol*., **219**, 733–746.
Uhlen,M. and Moks,T. (1990) *Methods Enzymol*., **185**, 129–143.

Received June 6, 1996; revised April 4, 1997; accepted April 9, 1997