Nucleic Acids Research, Vol. 19, No. 19 5191–5198

A DNA binding domain is contained in the C-terminus of *wild type* p53 protein

Orit Shohat Foord, Protima Bhattacharya, Ziv Reich[1] and Varda Rotter*
Department of Cell Biology and [1]Department of Organic Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel

Received July 23, 1991; Revised and Accepted September 13, 1991

ABSTRACT

In the present study we evaluated the DNA binding activity of *wild type* and *mutant* p53 proteins that were isolated from bacterial expression vectors. A comparison of the binding activities of the various purified p53 proteins, assessed by their ability to bind DNA cellulose columns, indicated that *wild type* p53 has a higher affinity to DNA than have *mutant* p53 forms. Furthermore, only *wild type* p53 was able to bind genomic DNA upon electrophoretic protein blotting. As specific deletion of the C-terminal region of *wild type* p53 totally abolished binding to genomic DNA, it was concluded that the 47 C-terminal amino acids contain the DNA binding region. The fact that the N-terminus contains a transcription activation region whereas the C-terminus contains a DNA binding domain places p53 in the family of typical transcription factors. Our experiments show that the topographical positioning of these domains plays an important role in the activity of *wild type* p53.

INTRODUCTION

Wild-type p53 protein is a growth regulator involved in the cell cycle, and it most likely functions at the transition of cells from the G_0/G_1 into the S-phase. The fact that transfection of *wild type* p53 induced growth arrest of proliferating cells was the basis for the assumption that p53 is a suppressor gene (1–8). Inactivation of *wild type* p53 by point mutations or deletions in mouse and human tumors was suggested to be a key event in the induction of the malignant process (9–15). The majority of tumor cells were found to express elevated levels of *mutant* p53 proteins, suggesting that, in addition to inactivation of *wild type* p53 (which is probably an early key event in the induction of malignant transformation), accumulation of *mutant* p53 protein may play, at least in certain types of cells, an active role in the establishment of the malignant phenotype.

The biochemical activity and function of the *wild type* p53 protein are still unknown. However, accumulation of p53 protein in the cell nucleus (16–19) and spatial regulation of the protein during the cell cycle (5) strongly suggest that the *wild type* p53 protein functions within this cellular compartment. In the cell nucleus, p53 may be associated directly with DNA replication or other regulatory cell cycle functions. *Wild type* p53 was found to compete with DNA α polymerase in the replication of SV40 DNA regulated by the large T antigen (20–22). Recently, it was shown that p53 is a target substrate for a cell cycle protein. Indeed it was found that p53 can be phosphorylated at a specific serine by $p34^{cdc-2}$, a cell cycle dependent protein (23,24). The fact that *wild type* p53 has been considered to be a transcriptional factor suggests that it can facilitate the expression of other target genes by direct transactivation (25,26). In the event that p53 is associated *in vivo* with either DNA replication or transcriptional transactivation, its activity requires either direct DNA binding as shown by several investigators (27,28) or interactions with DNA binding proteins. In any case, the cell nucleus is expected to be the natural milieu for its function.

The approach that we took to investigate the possible interactions between p53 protein and the target DNA consisted in purifying the proteins from a bacterial expression vector and studying their interactions with DNA. Using the *E.coli* expression vector (29), we purified and isolated authentic *wild type* and *mutant* p53 proteins. In our present studies we show, in agreement with previous findings, that *wild type* p53 has a higher DNA binding affinity than *mutant* p53 forms (27). Using a direct DNA binding assay we found that while *wild type* p53 bound genomic DNA, no binding was evident when *mutant* p53 was tested instead. Furthermore, a truncated *wild type* p53, which had no C-terminus, failed to bind total genomic DNA. This suggests that while the N-terminus of *wild type* p53 contains the transcription activation region (25,26), the C-terminus contains the DNA binding domain of this protein.

MATERIALS AND METHODS

Bacterial Strains

E.coli B strain BL21 (DE3) (F′-ompTr_B-m_B-) (29) was obtained from W. Studier (Brookhaven National Laboratory Upton). It contains a single copy of the gene for T7 RNA polymerase in the chromosome under the control of the inducible *lac*UV5 promoter. Expression of an eukaryotic gene in this system requires that the gene be located downstream to appropriate bacterial regulatory sequences that are necessary for efficient transcription and translation. In this system, expression of the

* To whom correspondence should be addressed

T7 polymerase is controlled by an inducible *lac*UV5 promoter integrated in the bacterial genome. Expression of the target gene contained in the multicopy plasmid is regulated, in turn, by the induced T7 polymerase (29,30). Expression of the heterologous gene, in our case p53, was induced during cell growth by adding β-D-thiogalactopyranoside (IPTG) to the growth culture.

Plasmids

Construction of the p53 expression plasmids was done by insertion of p53 cDNAs into the pET-8c plasmid described by Rosenberg et al. (30); expression of the inserted p53 cDNA was controlled by the T7 polymerase promoter. All cDNAs were cloned as NcoI-BamHI fragments into the same sites in the pET-8C vector. The first ATG, the start codon of the native protein, lies within the NcoI cloning site. pET-p53M8, pET-p53M11 coding for full length mutant p53 protein and pET-p53cD contain a mouse p53 full length cDNA, coding for *wild type* p53 protein (31, 32). pET-p53cDΔBs-B and pET-p53cDΔS-B contain a mouse cDNA cleaved at BspMI site and StuI site, respectively; these sites were filled in and ligated to the BamHI, coding for p53 proteins with deleted C-terminus. pET-p53cDΔRV has a deletion of an RV fragment. pET-p53H1 and pET-p53H19 contain a human p53 full length cDNA (32). pLR589, kindly donated by G. Lozano, codes for a *wild type* p53-GAL4 fusion protein. Expression of the protein is regulated by an LTR, p53 coding sequence containing genomic and cDNA fragments; GAL4 DNA binding sequences were fused at a StuI into GAL4 sequences, retaining the correct reading frame. pJ4ΩcDGAL-5 was derived from pJ4ΩcD (34), constructed by subcloning the GAL4 DNA binding domain into the 5′ end of p53 full length *wild type* coding cDNA, retaining the correct reading frame. Expression in this plasmid is regulated by LTR. pJ4ΩcDGAL-3Δ was generated by subcloning a SacII-BamHI 1 kb fragment containing the GAL4 DNA binding domain (obtained from pLR589) into SacII-BglII digested pJ4ΩcD. pJ4ΩcDGAL-5Δ was generated by deletion of the SacII-BglII followed by blunt ligation. The target reporter gene G4C, containing the GAL4-DNA binding domain upstream to the E1b transcription activation region (35), was obtained from M. Green, Harvard University.

Synthesis of p53 Protein

E.coli BL21 (DE3) cells containing the different plasmids were grown in 37° LB medium containing ampicillin (50 mg/ml). IPTG at a final concentration of 0.4 mM was added when the cells reached 1 O.D. Addition of IPTG to a growing culture of BL21(DE3) lysogen induces T7 RNA polymerase, which in turn transcribes the target DNA in the plasmid. Three hours after the induction cells were harvested by centrifugation and the pellet was washed and resuspended in TE containing 50 mM NaCl and 10 mg/ml lysozyme. The lysate was sonicated for 3×15 sec and subjected to low speed centrifugation. The pellet contained the total membrane fraction, and the supernatant contained the soluble fraction. Insoluble material containing the p53 protein was washed 3× with 4 M urea 0.1 M Tris-HCl pH 8.5 and solubilized by dissolving in 7 M guanidine-HCl, 50 mM Tris-HCl pH 9.0, 2 mM EDTA. Finally, the cell lysates were clarified and dialyzed against 50 mM NaCl, 10 mM Tris-HCl pH 7.8 and 1 mM EDTA.

DNA Cellulose Chromatography

Bacterial cell extracts were applied onto a 1 ml column of native calf thymus double-stranded DNA-cellulose (Sigma), equilibrated with 10 mM Tris-HCl pH 7.8 1 mM EDTA 50 mM NaCl. The column was washed with 20 bed volumes of equilibrated buffer to remove unbound protein. Bound protein was eluted by a salt gradient at 0.5 M up to 2.0 M NaCl in equilibration buffer. Fractions were diluted to 150 mM NaCl and immunoprecipitated with anti-p53 monoclonal antibodies.

Antibodies

The following anti-p53 antibodies were used: monoclonal anti-p53 PAb-240, PAb-242, PAb-246 (36,37); PAb-421 (38); 200.47 (39), and RA3−2C2 (40). The antibodies were either supernatants of growing cell lines or ascitic fluids obtained from the peritoneal cavity of hybridoma-bearing syngeneic mice which were spun to remove tissue debris, diluted in PBS, and used without further purification. Anti-immunoglobulin antibodies used were ^{125}I−sheep anti-mouse IgG, 0.5×10^6 CPM/ml (Amersham).

Immunoprecipitation and Western Blot Analysis

Radiolabeled protein samples were immunoprecipitated by a 2 h incubation at 4°C. Antigen-antibody complexes were collected by binding to Staph A on ice for 1 h (41). Complexes were centrifuged and washed three times with phosphate lysis buffer. For Western blotting (42) proteins were transferred from 10% (w/v) acrylamide gels to nitrocellulose paper (Schleicher and Schuell) for 4 h at 200 mA. Detection of specific p53 protein was done using monoclonal anti p53 antibodies that were detected by a radiolabeled second anti-mouse specific antibody.

Electrophoretic Protein Blotting and Subsequent DNA Hybridization

The blots were prepared as described above for Western blotting and washed 2 h in Standard Binding Buffer (SBB) (10 mM Tris-HCl pH 7.0, 50 mM NaCl, 1 mM EDTA, 0.02% BSA, 0.02% polyvinyl pyrrolidone, 0.02% Ficoll). Hybridization was preformed in a heat-sealable plastic bag for 3 h at room temperature. To the specific radiolabeled MboI digested DNA (2×10^5 cpm/ml), a nonspecific competitor DNA (10 μg/ml) was added. The blots were washed 3× with SBB for 20 min and exposed.

DNA Transfections and CAT Assays

HeLa cells were grown in DMEM and 10% FCS. Ten micrograms of the various DNA plasmids were transfected into cells at 40% confluence by calcium phosphate co-precipitation (43). After 5 h in the presence of the precipitate, the cells were treated with 20% glycerol in complete medium for 1 min. The cells were washed and maintained in complete medium. After 48−72 h, extracts were prepared and assayed for CAT activity (44), using equivalent amounts of protein.

RESULTS

Construction of Bacterial p53 Expression Vector

To study possible interactions between p53 protein and DNA target sequences, it was important to establish an efficient expression system for the synthesis of the p53 protein. To that end, we adapted the *E.coli* T7 RNA polymerase expression vector, pET-8c (29,30).

The protein expressed in this system is expected to be produced at high protein concentrations and to maintain its authentic structure. To facilitate the comparison of DNA binding affinities between *wild type* p53 and several *mutant* p53 forms, we subcloned a variety of p53 cDNA clones in the pET-8c vector.

In all cases the protein expressed was expected to be translated off the natural ATG site of the p53 sequence, giving rise to a native p53 protein without any fused bacterial residues. In our study we expressed the cDNA clones in the bacterial cells:

(i) **p53cD.** This cDNA clone was isolated from a normal T cell library (32). It codes for the *wild type* p53 product. In our previous studies we found using various assays that the protein that it encodes does not induce malignant transformation but functions instead as a suppressor protein (45).

(ii) **p53M11 and p53M8.** These two *mutant* coding p53 cDNAs were both isolated from a λgt10 cDNA library prepared from chemically transformed Meth A fibroblasts (31,32). Sequence comparisons of the various cDNA clones indicated that the *mutant* p53 protein encoded by p53M8 differs from the *wild type* p53 by a single point mutation, changing cysteine 132 into a phenylalanine. Furthermore p53M8 represents an mRNA molecule that was generated by alternative splicing of the 3′ terminus, coding for a smaller-size protein that lacks a C-terminal specific antigenic epitope (PAb-421). It contains a 96 base-pair insert that shares complete homology with an intron-specific sequence within the extreme 3′ end of intron 10, bordering the acceptor splicing site of exon 11. The 96 bp insert, which contains a premature protein termination signal, accounts for the transcription of a 96 bp longer transcript which nevertheless is translated into a shorter protein. In addition, 3 base substitution at nucleotide position 395, 503 and 729 were found between p53-M8 and p53-M11. Each of these substitutions predicted an amino acid difference between them. A comparison of p53M11 and *wild type* p53 indicated the existence of 2 point mutations, changing methionine 243 into an isoleucine and glutamine 168 into a glycine (31). Both p53M8 and p53M11 code for *mutant* p53 proteins that efficiently enhance malignant transformation of rat primary embryonic fibroblast (45).

(iii) **p53H1 and p53H19.** These two human cDNA clones were isolated from an SV40 transformed human fibroblast SV80 library, were also analyzed. The code for p53 polymorphic *wild type* human p53 proteins that vary in one nucleotide which modifies amino acid 72. While an arginine is expressed in the faster migrating species (p53-H1), a proline is found in the slower migrating protein (p53-H19) (33).

Characterization of the Bacterial p53 Proteins

Figure 1 illustrates the induction of pET-p53H19 that represents a typical pattern of bacterial p53 expression upon induction with IPTG. It is clear that a specific band of 53 kD is induced as early as 10 min after induction. Within 2 h, this 53 kD band comprises more than 50% of the total bacterial protein. In our first attempt to purify the p53 from the bacteria, we noticed that the proteins fractionated with the membranes upon cell solubilization, suggesting that the p53 protein synthesized in the bacteria is most likely concentrated in inclusion bodies. Our approach to purifying the p53 proteins therefore involved sequential washings of the bacterial membrane fraction with 4 M urea and 7 M guanidine-HCl. To confirm the identity of the induced 53 kD protein, we metabolically labeled IPTG-induced bacteria and immunoprecipitated cell lysates with specific anti-p53 PAb-421 monoclonal antibodies (38). In addition, total bacterial cell extracts were exposed to a Western blot analysis, using a second radiolabeled antibody. Figure 2 shows a typical experiment where the human encoded p53 protein (pET-p53H19) was studied and compared with the *in vivo* expressed protein. In both experiments we found that the bacterial p53 contained in the insoluble fraction retained the expected size and expressed the specific antigenic determinants. Table 1 summarizes the patterns of antigenic determinants expressed in the various bacterial p53 proteins. As expected murine p53 encoded by p53cD cDNA gave rise to a p53 protein that expressed the *wild type* p53 specific antigenic determinant recognized by PAb-246 (36) but lacked the PAb-240 *mutant* specific epitope (37). *Mutant* proteins encoded by p53M8 or p53M11 did not express the *wild type* specific PAb-246 determinant but expressed the PAb-240 *mutant* specific epitope. All murine p53 proteins expressed the RA3−2C2 (40), PAb-242,

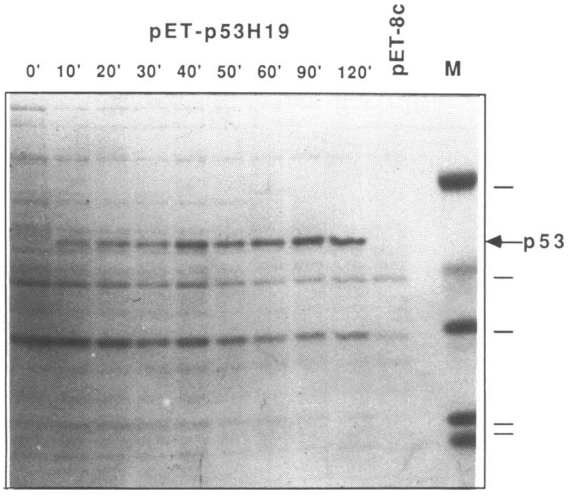

Figure 1. Induction of p53 expression in *E.coli*. *E.coli* cells harboring the vector pET-p53H19 were induced with IPTG and at the indicated time points after induction, 200 μl were removed and separated on a 10% acrylamide gel, followed by staining with Coomassie blue. Lane M represents the following size markers (from top to bottom): Bovine serum albumin (69 kD), ovalbumin (46 kD), Carbonic anhydrase (30 kD), Trypsin inhibitor (21.5 kD), and Lysozyme (14.3 kD). Arrow indicates the expected p53 proteins.

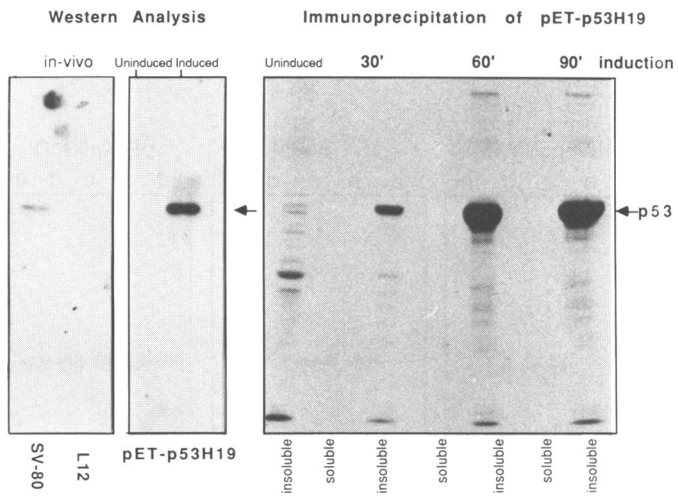

Figure 2. Partial purification and characterization of bacterial p53 protein encoded by pET-p53H19. Western blot analysis of human p53 protein expressed in bacteria, compared with the human p53 protein expressed in the SV80 transformed cell line *in vivo* (33). L12 is a p53 non-producer cell line (17) Immunoprecipitation of human p53 protein at 30, 60, and 90 min after IPTG induction. The PAb-421 anti p53 monoclonal antibody was added to soluble and insoluble fraction of cell lysate.

and PAb-248 mouse-specific determinants. In agreement with our previous experiments we found that *mutant* p53 encoded by p53M8, which has an alternatively spliced C-terminus, lacked the PAb-421 specific determinant (38). Both *wild type* p53, encoded by p53cD and *mutant* p53, encoded by p53M11, expressed the PAb-421 determinants. Human p53, encoded by either p53H1 or p53H19, expressed the PAb-421 determinant. The fact that *in vivo* size differences between the various p53 forms were retained in the bacterially expressed p53 protein forms (data not shown), coupled with the observation that these proteins expressed the expected antigenic determinants, strongly suggests that these bacterial products represent the authentic *in vivo* produced p53 proteins.

DNA Binding Affinities of Bacterially Expressed p53 Proteins

Recently it was shown by several investigators that p53 is indeed able to bind to both ds-DNA and ss-DNA (27,28), suggesting that DNA binding is an intrinsic property of p53. To determine whether the bacterially expressed p53 protein binds to DNA, we applied the various bacterial proteins onto DNA-cellulose columns and evaluated the binding affinities of the various products. Insoluble, partially purified p53 was first solubilized by dissolving it in 6 M urea and eventually dialyzed against TE with 50 mM NaCl. The soluble fractions of the bacterial p53 protein were chromatographed on native calf thymus ds-DNA cellulose in the presence of 50 mM NaCl. Elution was performed by increasing NaCl concentrations. The various fractions obtained were adjusted to 150 mM NaCl and immunoprecipitated with the anti-p53-specific monoclonal antibodies PAb-421 and RA3−2C2. Figure 3 shows a comparison in the elutions patterns of *wild type* and *mutant* p53 bacterial proteins. In general, as can be seen in this figure, the bacterially expressed p53 protein encoded by pET-p53M8, pET-p53M11 coding for *mutant* p53 proteins and pET-p53cD coding for *wild type* p53 protein bound to ds-DNA cellulose. However, there were variations in the DNA binding affinities. *Wild type* p53 exhibited a higher DNA binding activity than did *mutant* p53 forms. While *mutant* p53 was eluted around 1.25 M NaCl, *wild type* p53 required up to 2 M NaCl for the same elution patterns. These results agree with previously reported data suggesting that *wild type* p53 and *mutant* p53 vary in their DNA binding affinities (27). These variations are probably due to subtle changes in the tertiary structures of the various p53 protein.

To further asses the differences in the DNA binding affinities of *wild type* p53 versus *mutant* p53 forms, we modified the DNA binding assay. The various partially purified bacterial p53 proteins were separated by electrophoresis and transferred to nitrocellulose that was hybridized with radio labeled total MboI-digested genomic DNA. Figure 4 shows that while *wild type* p53 specifically bound genomic DNA, no binding was evident when *mutant* p53 proteins encoded by either pET-p53M8 or pET-p53M11 were used instead (A). Figure 4 shows that under the same experimental conditions *wild type* and *mutant* p53 bound specific anti-p53 monoclonal antibodies to an equal extent (B). This Southern-Western analysis represents therefore, a precise and simple assay for determining the specific DNA binding affinity of *wild type* p53 protein. Development of this assay permitted further definition of the specific DNA binding domain of the p53 protein.

Mapping of the DNA Binding Region of Wild Type p53 Protein

To localize the protein region responsible for site specific DNA binding, we performed a deletion mutagenesis of the *wild type* p53 protein. The various truncated versions of *wild type* p53 were expressed in *E.coli*. Bacterial cells expressing *wild type* p53 derivatives were lysed, the proteins were partially purified and DNA binding affinity properties of the modified proteins were examined. As shown in Fig. 5, after removal of 83 amino acids from the C-terminus of the molecule encoded by clone pET-p53cDΔBs-B no DNA binding activity could be measured by the Southern-Western assay (Figure 5, A). Under the same experimental conditions, the two p53 derived proteins efficiently bound the anti-p53 PAb-246 monoclonal antibody (Fig. 5, B).

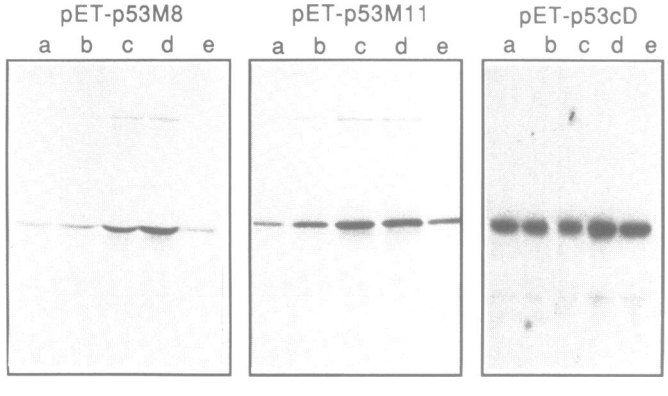

Figure 3. Evaluation of DNA binding affinities of the various bacterial p53 proteins. pET-p53M8 and pET-p53M11 coding for *mutant* p53 protein and pET-p53cD coding for *wild type* p53 were bound to calf thymus DNA-cellulose. Proteins bound to the column were eluted at 0.5 M, 0.75 M, 1 M, 1.5 M and 2 M NaCl (lanes a-e respectively) were immunoprecipitated with anti-p53 monoclonal antibodies and then subjected to SDS-PAGE and autoradiography.

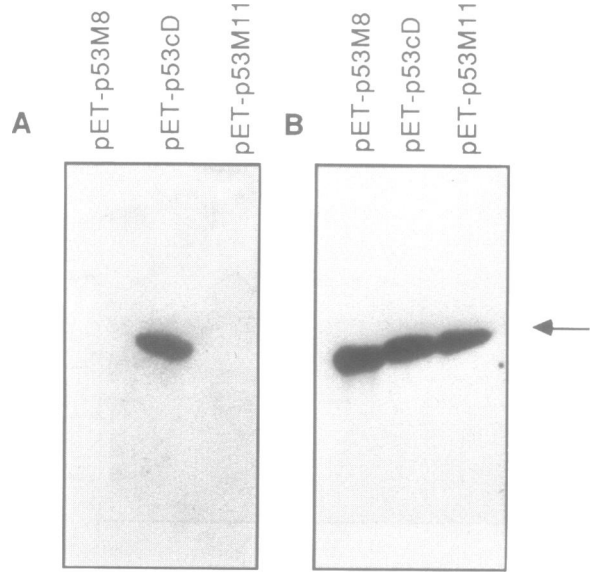

Figure 4. Specific DNA binding mediated by *wild type* p53 protein. The various bacterial p53 proteins were transferred onto nitrocellulose blots and hybridized with either radiolabeled MboI digested genomic DNA (A) or anti p53 monoclonal antibodies that was detected by a ^{125}I-sheep anti-mouse second antibody (B). Arrow indicates p53 protein.

This suggests that the lack of DNA binding is an intrinsic property of the protein encoded by the pET-p53cDΔBs-B plasmid.

To further define the DNA binding domain, smaller deletions at the C-terminus were tested. Figure 6 shows that when the protein was deleted at the StuI site generating a protein 47 amino acids short of the C-terminus almost all the binding was lost. However some residual activity was left. In repeated experiments some low affinity binding was detected. The faint band in lane pET-p53cDΔS-B was consistently detected. No residual binding was evident when the pET-p53cDΔBs-B protein was used under the same conditions. To exclude the possibility that the above deletions reflect a non-specific abolishment of DNA binding typical for *mutant* p53, we deleted the *wild type* p53 at the N-terminal part of the molecule. pET-p53cDΔEcoRV was generated by removing 70 bp from the N-terminus of the p53cD sequence; it codes for a shorter p53 protein. Although changed in structure, it clearly exhibited a DNA binding affinity equal to that found with the intact *wild type* p53 protein. Western blot analysis showed that, although this shorter protein lacks the PAb-242 determinant, because the deletion maps within that region, it retained its *wild type* structure, reflected in its capacity to bind the PAb-246 *wild type* specific determinant (see Table 1). These results define a 47 amino acid region contained at the C-terminus of the protein as the specific DNA binding region.

Transcriptional Activity of Wild Type p53

Transcription factors have been shown to contain two main functional domains: a DNA binding domain and a *transcriptional activation* region (46–50). It was shown previously that *wild type* p53 functions as a transcriptional factor. GAL4-p53 fusion proteins were found to induce CAT activity driven by a target plasmid containing a DNA binding site and an E1b transcription element (25,26). This activity was specific for *wild type* p53 and the functional region was mapped to the N-terminus of the molecule (25,26). Data presented here show that the C-terminus of *wild type* p53 contains the DNA binding domain. In our next experiment we evaluated the significance of the topographical

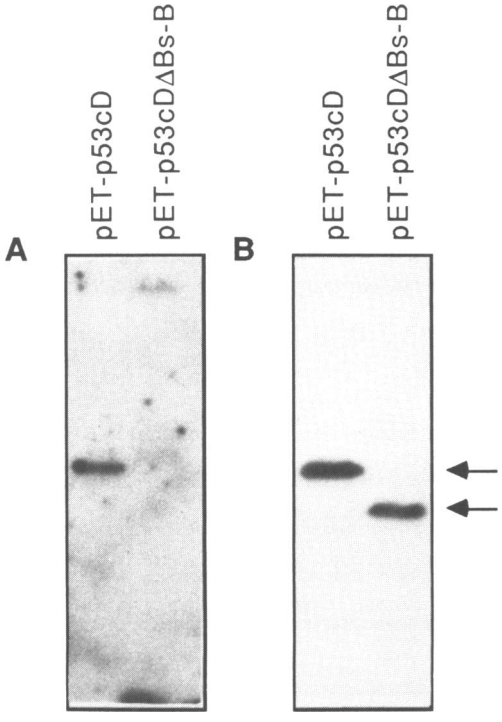

Figure 5. Mapping of the DNA binding domain of *wild type* p53 protein. Intact *wild type* p53 encoded by pET-p53cD or C-terminus deleted p53 encoded by p53cDΔBs-B were hybridized with either radiolabeled MboI digested genomic DNA (A) or anti-p53 monoclonal antibodies that was detected by a ^{125}I-sheep anti-mouse second antibody. Arrow indicates p53 protein. Upper arrow indicates intact p53 and lower arrow deleted p53 protein.

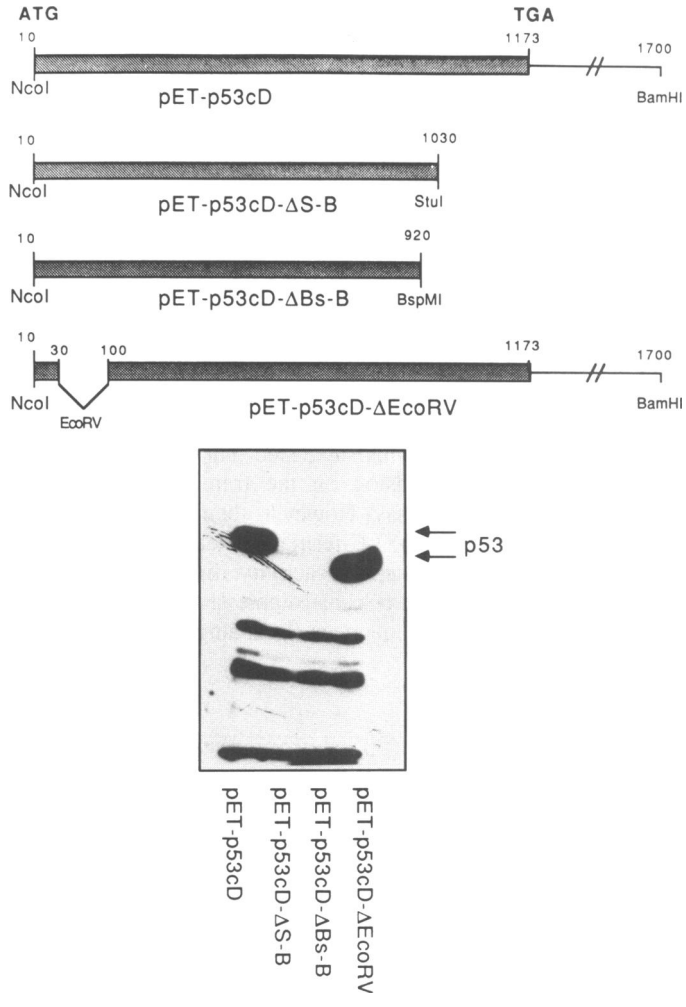

Figure 6. Comparison between various p53 derived proteins. Various p53 protein encoded by the different plasmids (top of the Figure) were hybridized to radiolabeled genomic DNA. p53 proteins are indicated by arrows.

Table 1. Characterization of antigenic determinants expressed in the bacterial p53 proteins

p53 cDNAs in *E.coli*	Species	p53 phenotype	p53 antigenic determinants PAb-240	PAb-242	PAb-246	PAb-421	RA3-2C2
p53cD	Murine	*Wild type*	–	++	+	++	+
p53cDΔBs-B	Murine	*Wild type*	–	++	+	–	+
p53cDΔS-B	Murine	*Wild type*	–	++	+	–	+
p53cDΔEcoRV	Murine	*Wild type*	–	–	+	++	+
p53M8	Murine	*Mutant*	+	++	–	–	+
p53M11	Murine	*Mutant*	+	+	±	+	+
p53H1	Human	*Wild type*	–	–	–	++	–
p53H19	Human	*Wild type*	–	–	–	++	–

relationship between these two domains. We compared the induced CAT activity mediated by p53 GAL4 proteins varying in their positioning of the GAL-DNA binding domain relative to the *transcriptional activation* region of the p53 protein. For that, we compared the CAT activity induced by a fusion GAL4-p53 protein in which the GAL-DNA binding sequences were fused at the N-terminus of p53 with the CAT activity induced by a p53-GAL4 protein in which the DNA binding GAL sequences were fused to the C-terminus. The transcriptional region contained in the N-terminus of the p53 protein is identical in all constructs used. HeLa cells were co-transfected with the various effector plasmids coding for p53-GAL4 fusion protein and the target plasmid G4C, which contains the specific GAL4-DNA binding sequences and the E1b transcription regulation region upstream to the CAT reporter gene. The results in Fig. 7 agree with previous reports showing a high CAT activity of the pLR589 plasmid coding for a p53 protein fused at the C-terminus with the GAL4-DNA binding domain. When pJ4ΩcDGAL-5 was used, almost no activity above that of the G4C clone alone was detected under the same experimental conditions. In addition to the difference in the positioning of GAL-DNA binding sequences, pLR589 also lacks the C-terminal region that contains the DNA binding domain. We found that deletion of a SacII-BamHI of the pJ4ΩcDGAL-5 gave rise to a protein with some what higher CAT activity (pJΩcDGAL-5Δ) than that of the parental plasmid. This suggested that the C-terminus of p53 exerts some suppression on the transcriptional activity measured in the present assay. However, the fact that the activity induced by this 5′ GAL4-p53 C-terminal deleted pJΩcDGAL-5Δ protein was still lower than the activity directed by pLR589 suggests that, in addition to possible suppressive activity mediated by the basic C-terminal sequences (in analogy with the human *wild type* p53 (26)), the major reason for the lower CAT activity was the topographical correlation between the GAL4 DNA binding sequences and the transcriptional activity region of *wild type* p53. We performed a control experiment in which, in the same vector, a SacII-BamHI fragment containing the GAL4 DNA binding sequences was constructed at the C-terminus (plasmid pJΩcDGAL-3Δ), instead of at the N-terminus plasmid (clone pJΩcDGAL-5Δ). In this experiment we obtained CAT activity levels similar to those with the pLR589 positive control. It should be noted that p53 sequences in the pJΩcDGAL-3Δ does not contain any p53 intron specific sequences which are found in the pLTR589 vector. The possibility that the variation in activity of the various effector plasmids resulted from variations in protein expression was excluded, since all plasmids were found to induce the expected fusion products in transient expression assays.

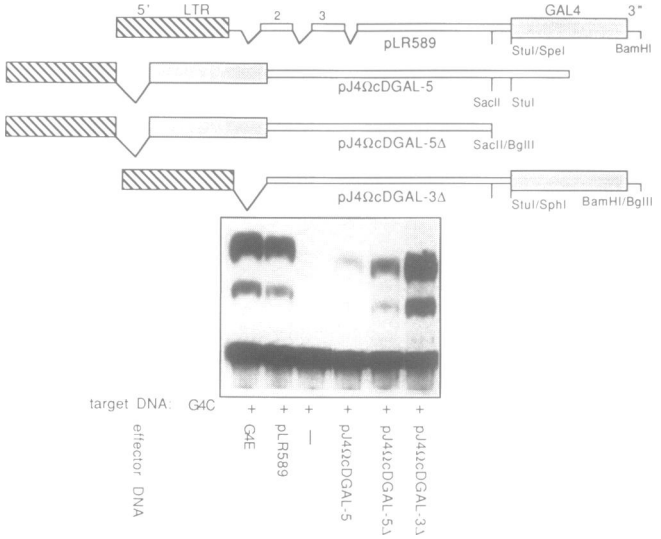

Figure 7. Transcription activity of various p53-GAL4 fusion proteins. The various p53-GAL4 fusion proteins encoded by plasmids are illustrated in the top part of the figure. In all constructs the N-terminus containing the transcriptional region was identical. Ten micrograms of the various plasmids were co-transfected with the effector plasmid G4C into HeLa cells, and CAT activity was determined after 48 h. G4E is the positive control containing the E1b-derived *transcriptional activation* domain. Striped box: the LTR promoter; dotted box: GAL4 DNA binding domain. Empty box: p53 coding sequences. pLR589 is a fusion p53 construct consisting of exons 2 and 3 fused into p53 cDNA sequences.

DISCUSSION

The p53 protein structure deduced from the DNA sequence shows that while the N-terminus is acidic in nature, the C-terminus is basic. In the course of our experiments we found that *wild type* p53 contains a specific DNA binding domain at the C-terminus of the molecule, a region which is highly conserved and which contains several potential active sites. In addition to the conserved serine (312 in the mouse), which is specifically phosphorylated by the *CDC2* cell cycle specific protein (23,24), this region contains three nuclear localization signals that account for the nuclear localization of the protein (34, 51, 52). Furthermore, it can be folded into an amphipathic α helix that facilitates DNA binding activity.

Previous reports showed that the N-terminus contains a specific *transcriptional activation* region. Indeed, the furthest 73 N-terminal amino acids were sufficient to induce transcriptional activity measured in a GAL4-p53 fusion system (26). In that respect, p53 represents a typical transcription factor, the activity of which is dominated by two separate domains (46−50). For the protein to be active requires the integrity of these domains. Like p53, c-myc contains an active transcription domain in the N-terminus, whereas the C-terminus contains the DNA binding domain. In other nuclear transcription factors, activation domains

```
1.....10.....20.....30.....40.....50.....60.....
MTAMEESQSDISLELPLSQETFSGLWKLLPPEDILPSPHCMDDLLLPQDVEEFFEGPSEALRVSG

...70.....80.....90.....100.....110.....120.....130
APAAQDPVTETPGPVAPAPATPWPLSSFVPSQKTYQGNYGFHLGFLQSGTAKSVMCTYSPPLNKL
(p53M8)                                (p53M11)
.F.....140.....150.....160.....G.....180.....190.....
FCQLAKTCPVQLWVSATPPAGSRVRAMAIYKKSQHMTEVVRRCPHHERCSDGDGLAPPQHLIRVE
                                                (p53M11)
..200.....210.....220.....230.....240..I....250.....260
GNLYPEYLEDRQTFRHSVVVPYEPPEAGSEYTTIHYKYMCNSSCMGGMNRRPILTIITLEDSSGN
                                                BspMI
.....270.....280.....290.....300.....|310.*.....320.....
LLGRDSFEVRVCACPGRDRRTEEENFRKKEVLCPELPPGSAKRALPTCTSASPPQKKKPLDGEYF
                                                NLSI
                StuI                                LQPRAFQALIKEESPNC(p53M8)
..330.....340..|....450.....360....*..370.....380.....390
TLKIRGRKRFEMFRELNEALELKDAHATEESGDSRAHSSYLKTKKGQSTSRHKKTMVKKVGPDSD
          α amphipathic helix          NLSII     NLSIII
```

Figure 8. Amino acid sequence of the intact *wild type* murine p53 protein. Protein structure was deduced from the sequence of a p53 cDNA clone isolated from a normal T cell cDNA library (32). NLSI, NLSII and NLSIII: the nuclear localization containing sequences (34). The α amphipathic helix underlines the corresponding amino acids. Sequence variations are marked by p53M8 and p53M11 accordingly.

are found in the C-terminus and the DNA binding domain maps to the N-terminus. This is the case with the c-*rel* (49) and the c-*ets* (53).

The location of the DNA binding domain at the extreme C-terminus and that of the *transcriptional activation* region at the N-terminus probably creates a suitable tertiary structure that permits a correct bending of the protein to facilitate DNA binding while inducing transcriptional activity. That this specific molecular organization is required for the activity of the protein is supported by the results presented here, showing that transcription activity of p53-GAL4 proteins was detectable only when the GAL-DNA binding domain was fused at the C-terminal part of the p53 protein, substituting for the endogenous p53 DNA binding domain. Almost no transcription activity was found when the GAL-DNA binding domain was fused at the N-terminus of the p53 protein, suggesting that topographical proximity between the transcription activation region and the DNA binding domain is an important factor in regulating transcriptional activity of p53. When this configuration is impaired, the activity of *wild type* p53 protein is abolished.

Both DNA binding activity and induction of transcriptional activity are mediated solely by the *wild type* p53. Inactivation mutations impair these activities, suggesting that subtle variations in the protein structure induce conformational modifications that are sufficient to interfere with both activities.

In our present experiments we estimated p53 DNA binding activity by two different assays. One measured the capacity of p53 protein to bind DNA cellulose and the other measured DNA binding by South-Western analysis. Whereas *wild type* p53 efficiently bound DNA in both assays, *mutant* p53 showed a lower DNA affinity in the DNA cellulose assay and no binding by the South-Western assay. Therefore we concluded that the two p53 molecule types vary in their DNA binding affinities and that the South-Western analysis most likely measures DNA binding of more specific affinities.

It is worth mentioning that although *wild type* p53 and *mutant* p53 encoded by p53M11 share complete homology at the C-terminus (which contains the DNA binding domain), they exert different DNA binding activities. By the direct Southern-Western method we found that, while *wild type* p53 efficiently bound genomic DNA, *mutant* p53 encoded by p53M11 lacked this activity. We therefore concluded that, in addition to the specific DNA binding domain contained at the C-terminus, perfect binding requires additional structural motifs mapping outside the 83 C-terminal amino acids.

A structural analysis of the C-terminal amino acids did not reveal the existence of known DNA binding motifs. A similarity search, conducted with the aid of the Pearson and Lipman scoring matrices (54), indicated no homology with other known DNA binding motifs (or review see 50). An analysis of the predicted protein structure showed no appropriately spaced cysteine or histidine residues known to stabilize a zinc finger structure. Furthermore, no alien structure representing a known leucine zipper cluster arrangement (55) and no significant linear homology with sequences speculated to be crucial for the formation of α helix-Loop-helix configuration could be detected (56–58). This suggesting that p53 most likely contain a novel DNA binding motif.

Interestingly, however, a secondary structure prediction (59) enabled the identification of an amphipathic α helix within the region shown to contain the DNA binding domain. Immediately adjacent to the amphipathic α helix, there is a conserved FREL motif common to DNA binding proteins that is preceded by a basic region (see Fig. 8). It should be noted that the region spanning amino acids 333 to 373 is flanked by two glycine residues, which consequently breaks any structural continuity of the presumed helical structure. The fact that removal of the entire C-terminus totally abolished DNA binding, while the StuI deletion, which cuts within the amphipathic α helix (Figure 8) leaving the upstream region intact, shows some residual DNA binding strongly suggests that the integrity of the entire region which may contain more than one potential motif is essential for high affinity DNA binding.

Therefore, it seems that the p53 DNA binding region represents a unique unknown DNA binding motif. The possibility that additional potential DNA binding sites located outside the 83 C-terminal amino acids are required for efficient binding suggests that a tertiary interaction, rather than local secondary interactions, is essential for stabilization of the p53 DNA binding region.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Leo and Julia Forchheimer Center for Molecular Genetics, the Rockefeller-Weizmann Fund and the German-Israeli Foundation (GIF). V.R. is the incumbent of the Norman and Helen Asher Professorial Chair in Cancer Research at the Weizmann Institute and the recipient of a Career Development Award from the Israel Cancer Research Fund.

REFERENCES

1. Eliyahu, D., Michalovitz, D., Eliyahu, S., Pinhasi-Kimhi, O. and Oren, M. (1989) Proc. Natl. Acad. Sci USA 86, 8763–8767.
2. Finlay, C.A., Hinds, P.W. and Levine, A.J. (1989) Cell 57, 1083–1093.
3. Mercer, W.E., Shields, M.T., Amin, M., Sauve, G.J., Appella, E., Romano, J.W. and Ullrich, S.J. (1990) Proc. Natl. Acad. Sci. USA 87, 6166–6170.
4. Michalovitz, D., Halevy, O. and Oren, M. (1990) Cell 62, 671–680.
5. Shaulsky, G., Ben-Ze'ev, A. and Rotter, V. (1990) Oncogene 11, 1707–1711.
6. Baker S.J., Markowitz, S., Fearon, E.R., Wilson, J.K.V. and Vogelstein, B. (1990) Science 249, 912–915.
7. Diller, L., Kassel, J., Nelson, C.E., Gryka, M.A., Litwak, G., Gebhardt, M., Bressac, B., Ozturk, M., Baker, S.J., Vogelstein, B. and Friend, S.H. (1990) Mol. Cell. Biol. 10, 5772–5781.
8. Chen, P.-L, Chen, Y., Bookstein, R. and Lee, W.-L. (1991) Science, 250, 1576–1580.
9. Baker S.J., Markowitz, S., Fearon, E.R., Willson, J.K.V. and Vogelstein, B. (1989) Science 244, 217–221.
10. Takahashi, T., Nau, M.M., Chiba, I., Birrer, M.J., Rosenberg R.K., Vinocour, M., Levitt, M., Pass, H., Gazdar, A.F., and Mina, J.D. (1989) Science 246, 491–494.
11. Kelman, Z., Prokocimer, M., Peller, S., Kahn, Y., Rechavi, G., Manor, Y., Cohen A., and Rotter, V. (1989) Blood. 74, 2318–2324.
12. Rodrigues, N. R., Rowan, A., Smith, M. E. F., Kerr, I. B., Bodmer, W. F., Gannon, J. V. and Lane, D. P. (1990) Proc. Natl. Acad. Sci. USA 87, 7555–7559.
13. Mulligan, L. M., Matlashewski, G. J., Scrable, H. J., and Cavenee, W. K. (1990) Proc. Natl. Acad. Sci. USA 87, 5863–5867.
14. Ahuja, H., Bar-Eli, M., Advani, S.H., Benchimol, S. and Cline, M.J. (1989) Proc. Natl. Acad. Sci. USA 86, 6783–6787.
15. Nigro, J.M., Baker, J.S., Preisinger A.C., Jessup J.M., Hostetter, K., Cleary, K., Bigner S.H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F.S., Weston, A., Modali, R., Harris, C.C. and Vogelstein, B. (1989) Nature 342, 705–708.
16. Rotter, V., Abutbul, H. and Ben-Ze'ev, A. (1983) EMBO J. 2, 1041–1047.
17. Shaulsky, G., Goldfinger, N., Ben-Ze'ev, A. and Rotter, V. (1990) Mol. Cell. Biol. 10, 6565–6577.
18. Martinez J., Georgoff, I., Martinez J., and Levine, A. J. (1991) Genes and Development 5, 151–159.

19. Ginsberg D., Michalovitz D., Ginsberg D. and Oren M. (1991) Mol. Cell. Biol. 11, 582–585.
20. Gannon, J.V. and Lane, D.P. (1987) Nature 329, 456–458.
21. Braithwaite, A.W., Sturzbecher, H.W., Addison, C., Palmer, C., Rudge, K. and Jenkins, J.R. (1987) Nature, 329, 458–460.
22. Wang, E.H., Friedman, P.N. and Prives, C. (1989) Cell 57, 379–392.
23. Bischoff, J.R., Friedman, P. N., Marshak, D. R., Prives C., and Beach, D. (1990) Proc. Nat. Acad USA 87, 4766–4770.
24. Stürzbecher, H.W., Maimets, T., Chumakov, P., Brain, R., Addison, C., Simanis, V., Rudge, K., Philp, R., Grimaldi, M., Court, W. and Jenkins, J.R. (1990) Oncogene 5, 795–801.
25. Raycroft, L. Wu, H., and Lozano, G. (1990) Science 249, 1049–1051.
26. Fields, S. and Jang, S.K. (1990) Science 249, 1046–1049.
27. Kern, S.E., Kinzler, K.W., Baker, S.J., Nigro, J.M., Rotter, V., Levine, A.J., Friedman, P., Prives, C. and Vogelstein, B. (1991) Oncogene 6, 131–136.
28. Steinmeyer, K. and Deppert, W. (1988) Oncogene, 3, 501–507.
29. Studier, F.W. and Moffet, B.A. (1986) J. Mol. Biol. 189, 113–130.
30. Rosenberg, A.H., Lade, B.N., Chui, D.-S., Lin, S.-W., Dunn, J.J. and Studier, F.W. (1987) Gene 56,125–135.
31. Wolf, D., Harris, N., Goldfinger, N. and Rotter, V. (1985) Mol. Cell. Biol. 5, 127–132.
32. Arai, N., Nomura, D., Yokota, K., Wolf, D., Brill, E., Shohat, O. and Rotter, V. (1986) Mol. Cell. Biol. 6, 3232–3239.
33. Harris, N., Brill, E., Shohat, O., Prokocimer, M., Wolf, D., Arai, N., and Rotter, V. (1986) Mol. Cell. Biol. 6, 4650–4656.
34. Shaulsky, G., Goldfinger, N., Tosky, M.S., Levine, A.J. and Rotter, V. (1991) Oncogene, in press.
35. Lillie, J.W. and Green, M.R. (1989) Nature 338, 39–44.
36. Yewdell, J.W., Gannon, J.V. and Lane, D.P. (1989) J. Virol. 59, 444–452.
37. Gannon, J.V., Greaves, R., Iggo, R. and Lane, D.P. (1990) EMBO J. 9, 1595–1602.
38. Harlow, E., Crawford, L.V., Pim, D.C. and Williamson, N.M. (1981) J. Virol. 39, 861–869.
39. Dippold, W. G., Jay, G., DeLeo, A. B., Khoury, G. and Old, L. J. (1981) Proc. Natl. Acad. Sci. USA 78, 1695–1699.
40. Rotter, V., Witte, O.N., Coffman, R. and Baltimore, D. (1980) J. Virol. 36, 547.
41. Kessler, S.W. (1975) J. Immunol. 115, 1617–1624.
42. Towbin, H.H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350–4354.
43. Graham, F.L. and van der Eb, A.J. (1973) Virology 52, 456–467.
44. Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982) Mol. Cell. Biol. 2, 1044–1051.
45. Eliyahu, D., Goldfinger, N., Pinhasi-Kimhi, O., Shaulsky, G., Skurnik, Y., Arai, N., Rotter, V. and Oren, M. (1988) Oncogene 3, 313–321.
46. Ptashne, M. (1988) Nature 335, 683–689.
47. Ptashne, M. and Gann, A.A.F. (1990) Nature 346, 329–331.
48. Mitchell, P.J. and R. Tjian. (1989) Science 245, 371–378.
49. Bull, P., Morley, K.L., Hoekstra, M.F., Hunter, T. and Verma, I.M. (1990) Mol. Cell. Biol. 10, 5473–5485.
50. Johnson, P.F. and McKnight, S.L. (1989) Annu. Rev. Biochem. 58, 799–839.
51. Addison, C., Jenkins, J.R. and Sturzbecher, H.W. (1990) Oncogene 5, 423–426.
52. Dang, C.V. and Lee, W.M.F. (1989) J. Biol. Chem. 264, 18019–18023.
53. Wasylyk, B., Wasylyk, C., Flores, P.,l Begue, A., Leprince, D. and Stehelin, D. (1990) Nature 346, 191–193.
54. Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
55. Schuermann, M., Neuberg, M., Hunter, J.B., Jenuwein, T., Ryseck, R.-P., Bravo, R. and Müller, R. (1989) Cell 56, 507–516.
56. Murre, C., McCaw, P.S. and Baltimore, D. (1989) Cell 56,777–783.
57. Prendergast, G.C. and E.B. Ziff. (1989) Nature 341, 392.
58. Blackwood, E.M. and Eisenman, R.N. (1991) Science 25, 1211–1217.
59. Chou, PY and Fassman G.D. (1974) Biochemistry 13, 222–245.