Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots

Nikola P. Pavletich,[1] Kristen A. Chambers, and Carl O. Pabo

Howard Hughes Medical Institute and the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

Mutations in the p53 tumor suppressor gene are the most commonly observed genetic alterations in human cancer. The majority of these mutations occur in the conserved central portion of the gene, but there has been little information about the function of this region. Using proteolytic digestion of the 393-amino-acid human p53 protein, we have identified a 191-amino-acid protease-resistant fragment (residues 102–292) that corresponds to the central portion of p53, and we show that this core fragment is the sequence-specific DNA-binding domain of the protein. DNA binding is inhibited by metal chelating agents, and we find that the core domain contains zinc. Proteolytic digests also reveal a 53-amino-acid carboxy-terminal domain which we show to be the tetramerization domain of p53.

[*Key Words*: p53; DNA-binding domain; tumor suppressor; tetramerization domain]

Received August 16, 1993; revised version accepted October 7, 1993.

The p53 tumor suppressor gene is the most commonly mutated gene identified in human cancers, and mounting evidence points to the inactivation of p53 as a critical step leading to neoplastic transformation (for review, see Hollstein et al. 1991; Levine et al. 1991; Vogelstein and Kinzler 1992). Loss of p53 function results in an enhanced frequency of genomic rearrangements (Livingstone et al. 1992; Yin et al. 1992) and eliminates the growth arrest response induced by DNA damage (Kastan et al. 1992; Kuerbitz et al. 1992). These studies suggest that p53 controls a cell cycle checkpoint that is important for maintaining the integrity of the genome (Hartwell 1992; Lane 1992).

The exact mechanisms through which p53 exerts its biological function are not known, but its ability to bind to specific DNA sequences (Bargonetti et al. 1991; Kern et al. 1991; El-Deiry et al. 1992; Funk et al. 1992) and activate transcription (Fields and Jang 1990; Raycroft et al. 1990; Farmer et al. 1992; Kern et al. 1992; Unger et al. 1992), and its interactions with several cellular proteins, such as the MDM-2 oncogene product (Momand et al. 1992; Oliner et al. 1992), seem to play important roles. For example, tumor-derived p53 mutants that have lost their ability to cause cell cycle arrest also are inactive in DNA binding and *trans*-activation (Kern et al. 1992; Unger et al. 1992). Furthermore, many of these mutants apparently promote neoplastic processes by forming hetero-oligomers with wild-type p53 and abrogating its activity in a dominant negative manner (Milner and Medcalf 1991; Shaulian et al. 1992). The transcriptional activation function of p53 has been mapped to residues 1–42 at the amino terminus (Fields and Jang 1990; Raycroft et al. 1990; Farmer et al. 1992; Kern et al. 1992; Unger et al. 1992), and the oligomerization activity has been roughly mapped to the carboxy-terminal portion (Shaulian et al. 1992; Stürzbecher et al. 1992). It has been proposed that the DNA-binding domain may also reside in the carboxy-terminal region of p53 (Foord et al. 1991), but no data have been presented to show that this region is involved in sequence-specific DNA binding.

In an attempt to locate the sequence-specific DNA-binding domain, we used proteolytic digestion to identify the major structural domains in the human p53 protein, and we then expressed these domains in *Escherichia coli* and characterized their function in vitro. The proteolytic digestion experiments show that the amino-terminal portion of p53 is highly susceptible to proteolytic digestion and suggest that it is solvent exposed and loosely folded. In contrast, the highly conserved middle portion of p53 is strikingly resistant to proteolytic digestion and thus seems to be an independently folded, compact structural domain. Finally, the pattern of cleavage sites in the carboxy-terminal portion of p53 reveals a second, smaller structural domain. In vitro experiments using the recombinant domains show that the central, core domain contains the sequence-specific DNA-binding activity of p53, and this activity appears to be dependent on zinc binding. The carboxy-terminal do-

[1]Corresponding author. Present address: Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021 USA

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

main contains the tetramerization activity of p53. Contrary to a previous report (Stürzbecher et al. 1992), a 20-residue basic region does not seem to be necessary for tetramer formation.

Results

Proteolytic digestion of human p53 reveals two structural domains and a loosely folded amino-terminal region

The proteolytic digestion experiments were done with immunopurified human p53 (Bargonetti et al. 1992) and the protease subtilisin (subtilisin has relatively low sequence specificity, making it a useful probe for structural domains). The subtilisin concentration was varied over a wide range to determine which regions of the molecule are the most and the least susceptible to proteolysis. After subtilisin digestion, the proteolytic fragments were purified and analyzed by amino-terminal sequencing, mass spectroscopy, and amino acid analysis. The cutting points inferred from these analyses are presented in Figure 1A.

Figure 1B shows the pattern of fragments produced when human p53 is digested with increasing concentrations of subtilisin. Cutting occurs most readily at the extreme amino terminus (at residues 8, 18, 23) and at three clusters of sites in the carboxy-terminal portion of p53 (at residues 305 and 311; at residues 364, 366 and 368; and at residues 387 and 388). The digestion pattern in the carboxy-terminal region reveals the existence of a 53-amino-acid fragment (residues 311–363), which is relatively resistant to proteolytic digestion, and suggests that this fragment is an independently folded structural domain (Fig. 1B, lanes 2,3). This carboxy-terminal domain is very readily cleaved from p53, and we presume that it is connected to the rest of the protein by a flexible linker. Immediately after this domain, and separated from it by a protease-sensitive linker, there is a ~20-amino-acid region rich in basic amino acids (residues 368–387). It appears that this basic region has considerable secondary structure, as subtilisin does not readily cut internal to it, even though it cuts the flanking regions with ease.

Digestion with higher subtilisin concentrations gives additional cuts in the amino-terminal region, which is cleaved progressively at residues 39, 40, 44, 54, and 68 and eventually at residues 94, 96, and 102. The observed digestion pattern indicates that the region containing residues 1–67 is readily accessible to subtilisin, suggesting that it is generally solvent exposed and loosely folded. The transcriptional activation function of p53 has been attributed to residues 1–42 in this region (Unger et al. 1992), and the accessibility or flexibility of this region may be important for the protein–protein interactions involved in *trans*-activation. Interestingly, the amino-terminal portion of p53 also contains the binding site for the MDM2 oncoprotein (residues 1–118, Oliner et al. 1993).

At high concentrations, subtilisin also cuts the region between residues 293 and 306, giving a 191-amino-acid core fragment (residues 102–292) that is strikingly resistant to digestion and persists even at a fivefold excess (wt/wt) of subtilisin (Fig. 1B, lane 6). This clearly shows that the core fragment is an independently folded, compact structural domain. Consistent with this evidence for a tightly folded structure, the core domain contains the evolutionarily highly conserved regions of p53. These include residues 117–142, 171–181, 234–258, and 270–286. (The one conserved region of p53 that is not in the core domain is found near the amino terminus, at residues 13–19). The core domain also contains the majority of the mutation sites identified in tumors that are concentrated in the conserved regions in the central portion of p53 (Hollstein et al. 1991).

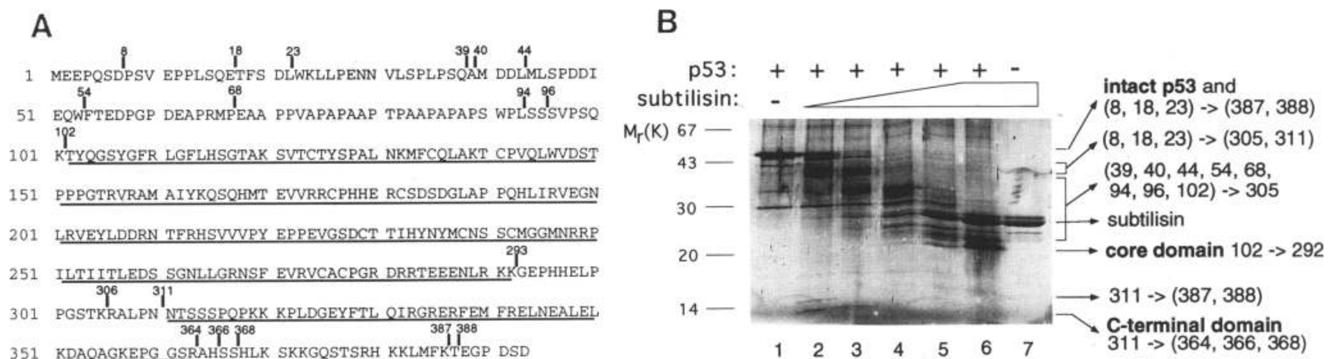


Figure 1. Digestion of human p53 by subtilisin. (*A*)Amino acid sequence of the human p53 protein showing the subtilisin cutting points (residues carboxy-terminal to the cleaved peptide bonds) identified in this study. The sequences of the protease-resistant core domain and the carboxy-terminal domain are underlined. (*B*)SDS-gel (12–20% polyacrylamide gradient) stained with silver showing a titration of the reaction of subtilisin with p53. (*Right*) The results of extensive microchemical and mass spectroscopic analyses are summarized; for lack of space, the data are not presented in this paper. Reactions contain the following ratios of subtilisin to p53 (wt:wt): (Lane *1*) No subtilisin; (lane *2*) 10%; (lane *3*) 25%; (lane *4*) 60%; (lane *5*) 200%; (lane *6*) 500%. Lane 7 contains a reaction similar to that of lane 6 but has no p53. Some of the fragments give broader bands on the gel, and this may reflect heterogeneity in the cleaving points. The carboxy-terminal fragments are not readily apparent in this gel because of their low mass relative to the rest of the fragments.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

The core domain contains the sequence-specific DNA-binding activity of p53

Using the gel mobility shift assay, we tested the products of the subtilisin digestion reaction for DNA binding and found that the fragments retained significant activity (data not shown). Binding to a p53 site could be observed even when the digestion mixture contained no detectable intact p53 and consisted predominantly of the core domain. The shifted DNA band had a mobility faster than that of the intact p53–DNA complex, suggesting that a proteolytic fragment of p53, presumably the core domain, was responsible for the DNA-binding activity. To test this more carefully (the subtilisin digests contained small amounts of other fragments as well), we expressed the core domain (residues 102–292) in *E. coli*, purified it to near homogeneity, and then performed gel mobility-shift experiments with a p53-binding site from the ribosomal gene cluster promoter (RGC site; Kern et al. 1991).

Figure 2A shows that the recombinant core domain can form a complex with the RGC site and gives a shifted band with a mobility faster than that of the intact p53–RGC site complex. (In the presence of 175 mM NaCl and 28 μg/ml of nonspecific competitor DNA, 50% binding is obtained with 0.2 μM intact p53 or with 1.1 μM core domain.) To determine whether the core domain binds specifically to the RGC site, we challenged the complex with either the wild-type RGC site or a mutant site. Figure 2A demonstrates that, like intact p53 (lanes 8–11), the core domain strongly prefers the wild-type RGC site over the mutant site (lanes 3–6).

We also compared the DNA sequence specificities of the core domain and of intact p53 by means of a methylation interference assay. Kern et al. (1991) reported that the RGC site contains several guanines which, when methylated by dimethylsulfate (DMS), interfere with the binding of p53. The methylation interference pattern that we observe with the core domain–RGC complex (Fig. 2B) reveals that this complex also is sensitive to guanine methylation. When this pattern is compared with the pattern observed with intact p53 (as reported by Kern et al. 1991) it appears that the methylation-sensitive guanines of the core domain–RGC complex form a subset of those found with the intact p53–RGC complex. The methylated guanines that interfere most strongly with the binding of intact p53 also interfere with the binding of the core domain (these are guanine 25 on the upper strand and guanines 20′, 21′, and 25′ on the lower strand; Fig. 2B). On the other hand, the methylated guanines that only partially interfere with the binding of intact p53 (Kern et al. 1991) show little or no interference with the binding of the core domain (these are guanines 15, 19, and 20 on the upper strand and guanines 30′ and 31′ on the lower strand; Fig. 2B). These results demonstrate that the core domain can make many of the key DNA contacts made by the intact p53 protein.

The methylation interference patterns of the core domain and intact p53 can also be interpreted by referring to the pentamer consensus sequence Pu-Pu-Pu-C-(A/T). The majority of the known p53-binding sites contain four repeats of this pentamer motif (although some of the repeats only partially match the consensus). The RGC site contains three repeats with a perfect match to the consensus and one repeat with a partial match (Fig. 2B). In the case of intact p53, the methylated guanines that interfere most strongly with binding occur in only two of these repeats; the third repeat shows only weak interference, and the fourth (imperfect) repeat shows no interference at all. As mentioned above, the core domain contacts mimic the strongest contacts seen with the intact protein. The methylated guanines that interfere strongly also occur in the first two repeats, whereas the third and fourth repeats do not show any significant interference. In summary, methylation interference shows that the core domain interacts preferentially with two of the pentamer motifs of the RGC site, and these are the same regions where the intact p53 makes its strongest contacts.

Two pentamer sequences are required for the binding of the core domain

p53 has been shown to form tetramers (Stenger et al. 1992; Friedman et al. 1993), and since binding sites typically contain four pentamer sequence motifs it seems possible that each p53 molecule interacts with a single pentamer motif. This model would predict that the isolated core domain, which is missing the oligomerization domain and appears to be a monomer in solution (gel filtration data; not shown), might bind to a single pentamer sequence. However, methylation interference experiments suggest that the situation is more complicated, and we did additional experiments to determine how many pentamers are required for tight binding by the core domain. We used competition experiments similar to those of Figure 2A, and we challenged the core domain–RGC complex with sites that contained one-, two-, or four-pentamer motifs. Figure 3A shows that the most effective competitors are the RGC site and a four-pentamer site. A site with two pentamers is also a good competitor, whereas a site with a single pentamer is a poor competitor (comparable to the mutant RGC site or to a site with no pentamer sequences). When we compare the amount of competitors required to give 50% reduction in binding (Fig. 3B), it becomes apparent that the core domain requires at least two pentamer sequences for specific binding. This finding suggests that two core domain molecules may interact with two pentamers in a cooperative fashion. Interestingly, when the amount of competitor is normalized for the pentamer content (Fig. 3B), the two-pentamer site seems to have the same affinity for the core domain as the four-pentamer site. Although cooperative binding (to form a dimer on the DNA) seems to be the most plausible explanation for our methylation interference and competition data, we cannot rule out the alternative possibility that one core domain molecule may interact simultaneously with two adjacent pentamers.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

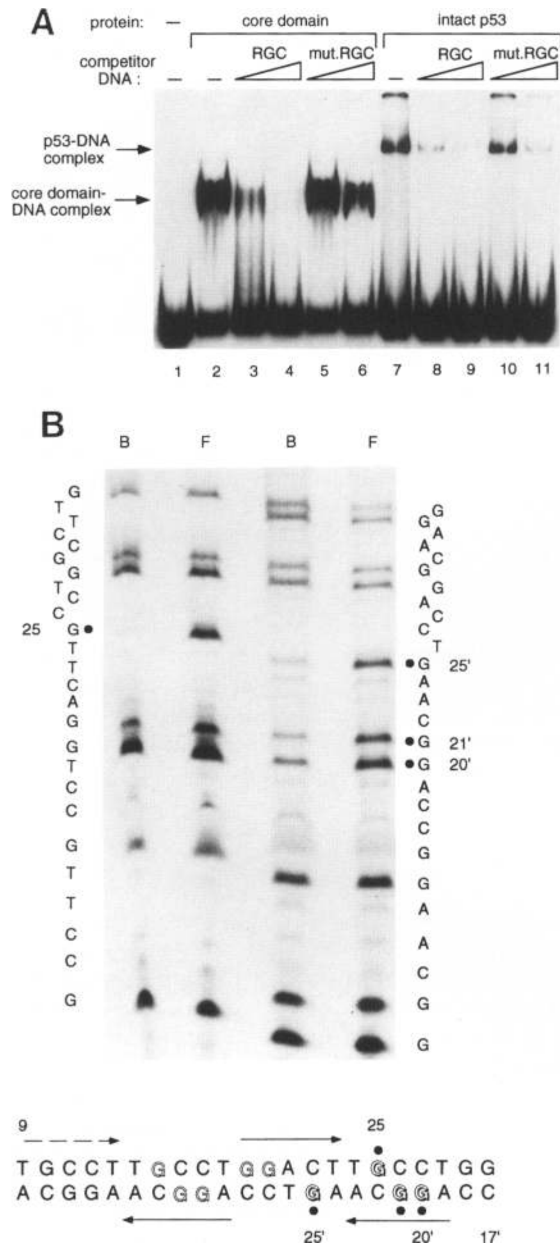


Figure 2. Comparison of the DNA-binding activities of the core domain and of intact p53. (*A*) Gel mobility shift experiments using the RGC site. Reactions in lanes *2–6* contain 1.1 μM of purified core domain peptide; those in lanes *7–11* contain 0.2 μM of immunopurified human p53. Lane *1* contains no protein. As sequence-specific competitors, we used either the RGC-binding site or a mutated RGC site: (Lanes *2,7*) No competitor; (lanes *3,8*) 100 ng of RGC site; (lanes *4,9* 1000 ng of RGC site; (lanes *5,10*) 100 ng of mutant RGC site; (lanes *6,11*) 1000 ng of mutant RGC site. Binding reactions contained 25 mM Tris-HCl (pH 7.2) 175 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 5% glycerol, 50 μg/ml of bovine serum albumin, 28 μg/ml of nonspecific DNA, and 1–2 nM radiolabeled RGC site DNA. (*B*) Methylation of critical guanine residues in the RGC site interferes with core domain binding. The two lanes at *left* show the piperidine cleavage pattern of the methylated upper strand. Lanes labeled B contain core domain-bound DNA recovered from the gel; Lanes labeled F contain free DNA recovered from the gel. The two lanes at *right* show the pattern obtained with the lower strand. Dots represent the methylation-sensitive guanine residues of the core domain–RGC complex. For comparison, the methylation-sensitive guanine residues of the intact p53–RGC complex (Kern et al. 1991) are indicated on the RGC sequence: Shadowed guanine residues show strong interference; outlined residues show weak interference. The locations of the pentamer sequence motifs are indicated by arrows. (*Solid arrows*) Pentamers that match the consensus; (dashed arrow) pentamer with two mismatches.

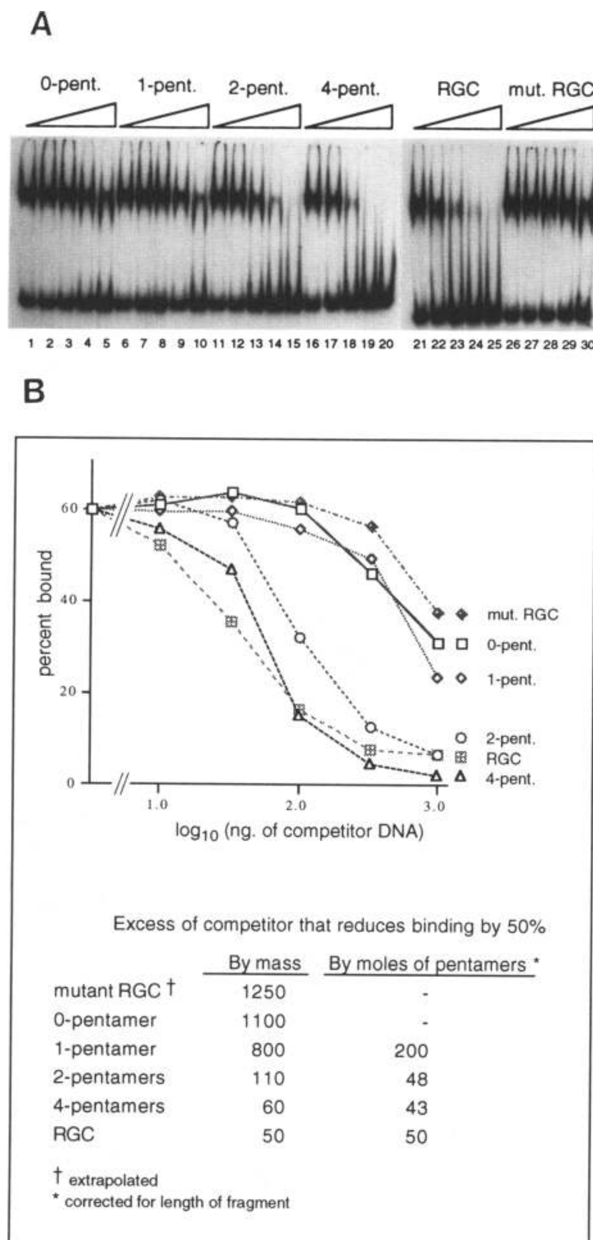
Zinc is required for the DNA-binding activity

The p53 protein contains 10 cysteines, all of them in the core domain, and this raises the possibility that p53 binds a metal ion. To address this possibility, we tested the effect of metal-chelating agents on DNA binding. Figure 4 shows that the metal-chelating agent 1,10-phenanthroline abolishes DNA binding by both the core domain and intact p53. Similar results were obtained with EDTA (data not shown). The metal seems to be tightly bound to p53, because relatively high concentrations of metal-chelating reagents are needed to abolish DNA binding (~5–10mM of 1,10-phenanthroline or EDTA).

To analyze the metal content of the core domain, we performed plasma emission spectroscopy (this analysis can detect 31 different elements, including Ca, Co, Cu, Fe, Mg, Mn, and Zn), and found that the only metal that the core domain contains is zinc. This result, taken together with our finding that metal-chelating agents abolish DNA binding, suggests that p53 is a zinc metalloprotein. Our conclusion is consistent with a recent report (Hainaut et al. 1993) showing that metal-chelating agents induce p53 to adopt a conformation that is recognized by monoclonal antibodies specific for the mutant form of the protein.

The plasma emission analysis was repeated several times, and the zinc content consistently corresponded to a ratio of ~ 0.5 mole of zinc per mole of core domain. Although this stoichiometry could in principle result

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press



Excess of competitor that reduces binding by 50%

	By mass	By moles of pentamers *
mutant RGC †	1250	-
0-pentamer	1100	-
1-pentamer	800	200
2-pentamers	110	48
4-pentamers	60	43
RGC	50	50

† extrapolated
* corrected for length of fragment

Figure 3. Competition experiments showing that the core domain requires two pentamer motifs for sequence-specific binding. (*A*)Binding reactions, similar to those of Fig. 2A, contained the core domain peptide (1.2 μM) and labeled RGC site (1 ng). The core domain–RGC complex was challenged with increasing amounts (10, 33, 100, 333, and 1000 ng) of the unlabeled competitor fragments (*top*). (Lanes *1–5*) A site with no pentamer motifs (but containing the same flanking sequences as the sites with pentamers); (lanes *6–10*) a site with a single pentamer; (lanes *11–15*) a site with two pentamers; (lanes *16–20*) a site with four pentamers; (lanes *21–25*) the RGC site; (lanes *26–30*) the mutant RGC site. (*B*) The results of the competition experiments were quantitated, and the amount of binding observed was plotted against the logarithm (base 10) of the amount of competitor used in each reaction. To facilitate comparisons, we list the amount of each competitor required to give a 50% reduction in binding. In calculating the molar excess of the pentamer competitors, we included the fourth, imperfect pentamer of the RGC site.

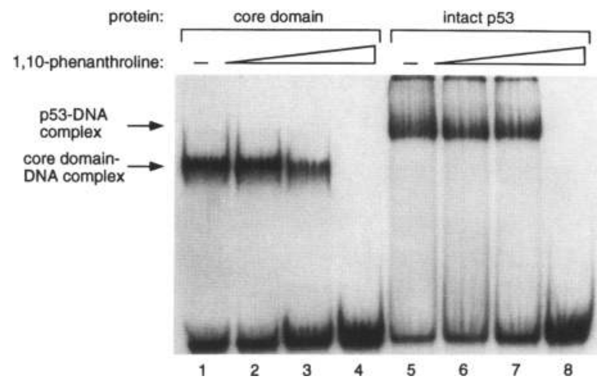


Figure 4. Gel mobility-shift assay showing that DNA-binding by the core domain and by intact p53 requires metal binding. Binding reactions, similar to those of Fig. 2A, with either the core domain (lanes *1–4*) or with intact p53 (lanes *5–8*) contained the following amounts of 1,10-phenanthroline: (lanes *1,5*) no 1,10-phenanthroline; (lanes *2,6*) 0.1 mM; (lanes *3,7*) 1 mM; (lanes *4,8*) 10 mM.

from a metal-bridged dimer, we think it is unlikely, since the core domain elutes from a gel-filtration column as a monomer. It seems more likely that some of the zinc dissociates from the protein during purification, since we find that the core domain is sensitive to oxidation.

The carboxy-terminal structural domain contains the tetramerization activity of p53

Previous studies (Shaulian et al. 1992; Stürzbecher et al. 1992) had mapped the oligomerization activity of p53 to the carboxy-terminal portion, and Stürzbecher et al. (1992) have reported that a stretch of basic residues in this region is required for the conversion of dimers to tetramers. The proteolytic digestion pattern that we observe indicates the presence of a 53-residue structural domain (residues 311–363) in this carboxy-terminal region of the protein. The short basic region implicated in oligomerization (residues 368–387) occurs immediately after this domain and is connected to it by a linker highly sensitive to subtilisin.

To determine whether the 53-residue structural domain has any oligomerization activity and to investigate the role of the basic region in oligomerization, we expressed the structural domain (residues 311–367) and a peptide consisting of this domain plus the basic region (residues 311–393) in *E. coli*, and assayed the oligomerization activity of the purified peptides. Each peptide elutes from a Superdex 75 gel-filtration column as a single, high-molecular mass complex, and the two peptides seem to have similar oligomerization states (Fig. 5A): The shorter peptide (residues 311–367) elutes at an estimated molecular mass of 37 kD (~6.4 molecules/oligomer) and the longer peptide (residues 311–393) at 53 kD (~5.9 molecules/oligomer). Since the elution profile on a gel-filtration column may be influenced by the molecular shape, we also performed glutaraldehyde cross-linking experiments to determine whether these oligomeric

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

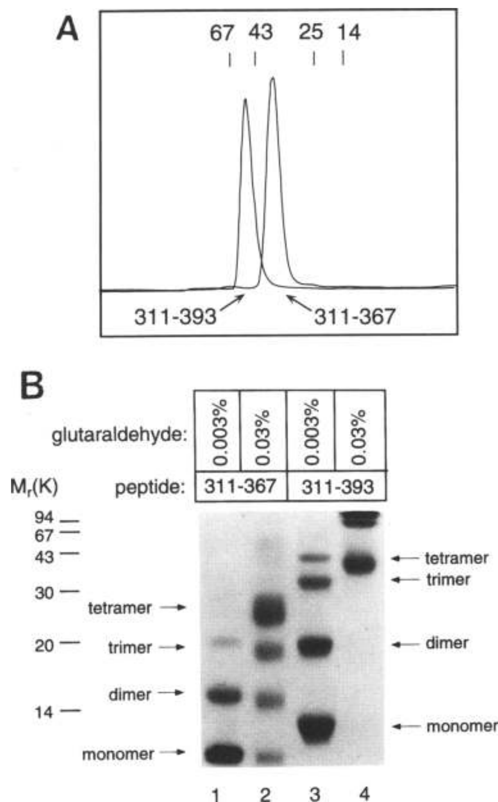


Figure 5. Comparison of the oligomerization activities of the carboxy-terminal domain and the carboxy-terminal domain plus basic region peptides. (*A*) Elution profiles of the two peptides on a Superdex 75 gel-filtration column. (*Top*) The position of the standards used to estimate the molecular mass. No other peaks are observed over a concentration range of 1 μM – 1 mM for either peptide. (*B*) Products of the reaction of glutaraldehyde with p53(311–367) and with p53(311–393) were separated on an SDS-gel (15% polyacrylamide) and visualized with Coomasie blue staining. The oligomerization states of the various products were estimated from their mobilities relative to those of the molecular mass markers.

complexes might actually be tetramers (Fig. 5B). Cross-linking of either peptide produces a ladder of four bands with sizes roughly corresponding to monomers, dimers, trimers, and tetramers. (Cross-linking of the longer peptide is more efficient, presumably because it contains 11 lysine residues, whereas the shorter peptide only contains 5 lysines.) Since the ladder ends with the tetramer band, and no pentamers, hexamers, and so forth, are observed, this suggests that the tetramer is the predominant form of these peptides. These results, taken together with the results of our gel-filtration experiments demonstrate that the carboxy-terminal structural domain is the tetramerization domain of p53 and show that the basic region is not required for tetramerization. It should be noted that cross-linking of the longer peptide also produces high-molecular-mass (>85 kD) products (Fig. 5B, lane 4). Although these products may represent nonspecific cross-linking because of the high lysine content of the basic region, it also is conceivable that the basic region is involved in the association of tetramers to form higher-order oligomers observed with intact p53 (Stenger et al. 1992). However, such high-molecular-mass forms were not observed in our gel-filtration experiments (Fig. 5A), and the significance of these bands is not clear.

Since it has been proposed that the DNA-binding domain of p53 may reside in the carboxyl terminus (Foord et al. 1991), we also tested our carboxy-terminal peptides for DNA binding. We found that the longer peptide (residues 311–393) has considerable affinity for DNA, but the shorter peptide (residues 311–367) which is missing the basic region, has no detectable affinity for DNA (data not shown). However, the interaction of the longer peptide with DNA is not specific for the p53-binding site, because we find that the mutant RGC site or nonspecific calf thymus DNA can effectively compete with the wild-type RGC site for binding. Although the affinity of this carboxy-terminal peptide for DNA may be an in vitro artifact, it is also possible that the basic region makes auxiliary contacts as p53 binds to DNA, or binds nonspecifically and helps regulate the activity of p53 (Hupp et al. 1992).

Discussion

Our proteolytic digestion experiments and our studies of fragments produced in *E. coli* provide new insights into the structural and functional organization of the p53 protein. The key finding of our study is that the central portion of p53 (residues 102–292) constitutes the sequence-specific DNA-binding domain of the protein. This finding reveals that the DNA-binding domain coincides with the major mutation hot spots (Fig. 6), and this helps us to understand why the tumor-derived mutant p53 proteins are defective in DNA-binding.

Several lines of evidence demonstrate that the core domain is the sequence-specific DNA-binding domain of p53. First, the core domain binds tightly to the RGC site, with an affinity comparable to that of intact p53. Second, competition experiments show that the core domain, like intact p53, strongly prefers to bind to the wild-type RGC site over the mutant RGC site. Third, methylation interference experiments show that the core domain makes many of the key DNA contacts made by intact p53. The methylated guanines that interfere most strongly with the binding of intact p53 are precisely the ones that interfere with the binding of the core domain. Fourth, the pentamer sequence motif that is required for the binding of intact p53 also is necessary for the binding of the core domain.

Although the core domain appears to be a monomer in solution, several observations suggest that it may bind as a dimer. A monomer might be expected to bind weakly, but we find that the core domain binds to the RGC site with an affinity only five- to sixfold lower than that of intact p53. We also find that the core domain requires at least two adjacent pentamers for specific binding. Again these results suggest that the core domain may form

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

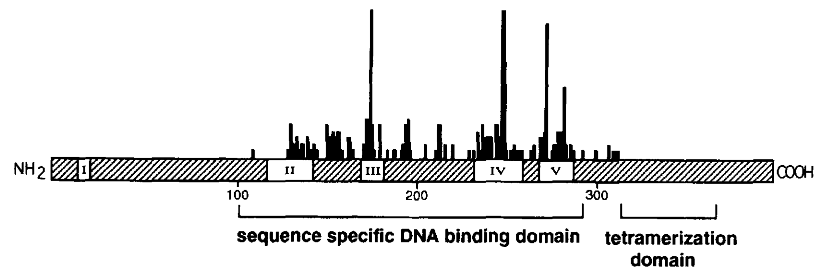


Figure 6. Structural domains of p53 identified in this study. Boxes with roman numerals indicate the five regions of p53 that are conserved across species; the bar graph above shows the approximate position and frequency of tumor-derived mutations (Hollstein et al. 1991). The positions of the DNA-binding domain (residues 102–292) and the tetramerization domain (residues 311–363) are indicated below.

dimers upon DNA-binding and thus recognize the two pentamers in a cooperative fashion.

In addition to the effects of oligomerization, there are likely to be other effects that modulate the DNA-binding activity of the intact protein. Hupp et al. (1992) have shown that the DNA-binding activity of p53 may be regulated by phosphorylation in the carboxyl terminus, and it is conceivable that there may be additional regulatory mechanisms. Possible regulatory effects may help to explain differences between our results and those of Shaulian et al. (1993), who showed that a p53 protein with a deletion in the oligomerization domain failed to bind to DNA in vitro but was functional in vivo. Their protein could *trans*-activate a target promoter containing a p53-binding site and also could suppress oncogene-mediated transformation. It is conceivable that their in vitro DNA-binding assay was not sensitive enough to observe the somewhat weaker binding of the monomer, but it is also likely that flanking regions (not present in an isolated core domain) may inhibit binding. Such inhibition of binding may be biologically relevant, but it is also possible that it represents artifacts resulting from the use of fragments that do not correspond to precise structural domains (i.e., misfolded regions at the ends may affect the activity of neighboring domains).

Our proteolytic digestion experiments reveal that p53 also contains a smaller structural domain located in the carboxy-terminal region (residues 311–363). Following this 53-residue domain, there is a flexible linker, and then we find a 20-residue highly basic region (residues 368–387). Our studies of peptides from this region show that the shorter carboxy-terminal domain is the tetramerization domain of p53, and the basic region seems to be involved in nonspecific DNA binding. Our findings are in general agreement with previous data that mapped the location of the oligomerization activity to the carboxy-terminal portion of p53 (Shaulian et al. 1992; Stürzbecher et al. 1992). However, some aspects of the oligomerization model of Stürzbecher et al. (1992) are inconsistent with our findings. They proposed that the region corresponding to the shorter carboxy-terminal domain is involved in dimerization and that the basic region is required to convert the dimers to tetramers. Several lines of evidence suggest that the basic region is not required for tetramer formation. First, a peptide containing the carboxy-terminal domain alone elutes from a gel-filtration column in a high molecular mass form, having essentially the same oligomerization state as the carboxy-terminal domain plus basic region peptide. Second, glutaraldehyde cross-linking experiments demonstrate that the preferred oligomerization state of both peptides is the tetramer. Third, Shaulian et al. (1992) have shown that a peptide corresponding to our carboxy-terminal domain (residues 302–360) has transforming properties very similar to those of a peptide containing the basic region (residues 302–390), suggesting that the two peptides have very similar oligomerization activities. (The transforming ability of these peptides seems to involve the formation of mixed oligomers with wild type p53.)

Our results, in conjunction with previous work from other groups, allow a clear delineation of the major domains of p53. The amino-terminal region contains the *trans*-activation domain, the central region contains the sequence-specific DNA-binding domain, and the carboxy-terminal region contains sites responsible for oligomerization and nonspecific DNA-binding. However, certain puzzles remain, and these may reflect (1) possible regulatory mechanisms, (where the presence of one domain may help to modulate the activity of another domain), or (2) additional minor roles of the domains, (thus the core domain could have some weak dimer contacts; the carboxy-terminal domain might be able to make some auxiliary contacts with the DNA). Further experiments will be needed to fully understand these effects. Nevertheless, the results presented here help us to understand the role of the conserved central region and the effects of mutations in this region. The identification of the structural domains of p53 also should aid in its structural characterization by crystallographic or nuclear magnetic resonance (NMR) spectroscopic methods.

Materials and methods

Proteolytic digestion and identification of the fragments

Proteolysis reactions contained 60–200 μg/ml of human p53 protein immunopurified from Sf21 cells expressing the recombinant baculovirus pEV55hwt (Bargonetti et al. 1992), in a buffer of 10 mM HEPES-Na at pH 7.5, 5 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol, and the protease subtilisin (the p53 preparations that we used in this study typically contain a number of minor contaminant proteins that have not been characterized). The reactions were typically carried out for 25 min and were stopped by the addition of 1 mM PMSF. For analysis, the fragments were either purified by reverse-phase HPLC on a C4 column, or separated by SDS-PAGE and transferred to Immobilon-P membrane (Millipore). Amino acid analyses were per-

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

formed on an ABI 420H amino acid analyzer with PITC derivatization, and amino-terminal sequencing was performed on an ABI 477A pulsed liquid microsequencer with on-line phenylthiohydantoin analysis. The mass of the purified fragments was determined either on an electrospray ionization triple-stage quadropole mass spectrometer (Finnigan) or on a matrix-assisted laser desorbtion time of flight mass spectrometer (Vestek and Finnigan). The analyses were performed at the Harvard Microchemistry Facility (Cambridge, MA) and at the Sloan-Kettering Microchemistry Facility (New York, NY). It should be noted that although we have identified most of the major fragments, there are several minor proteolytic fragments that we have not characterized.

Expression and purification of the core domain

The portion of the human p53 gene encoding residues 102–292 was amplified in a PCR, and the amplified product was cloned into the pET3d expression vector (Novagen). The core domain peptide was produced in *E. coli* BL21(D3) cells at room temperature. Cells were harvested by centrifugation and lysed by sonication in 40 mM MES-Na at pH 6.0, 200 mM NaCl, and 5 mM DTT. The lysate was clarified by centrifugation, diluted fivefold, loaded onto a Mono S cation-exchange column (Pharmacia) in 40 mM MES-Na at pH 6.0, 5 mM DTT, and was eluted by a NaCl gradient to yield the core domain at >85% purity. For most of the experiments, the Mono S fraction was precipitated by 80% saturated ammonium sulfate and was purified further by gel-filtration chromatography on a Superdex 75 gel-filtration column (Pharmacia) in 50mM bistrispropane-HCl at pH 6.8, 100 mM NaCl, and 5 mM DTT to yield the core domain at >98% purity.

Gel mobility shift and methylation interference assays

Binding reactions (15 μl) contained 25 mM Tris-HCl (pH 7.2), 175 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 5% glycerol, 50 μg/ml of bovine serum albumin, 28 μg/ml of nonspecific DNA, and 1–2 nM radiolabeled RGC site DNA. The reactions were equilibrated at room temperature for 20 min, and were electrophoresed at 10 V/cm in a 4.5% polyacrylamide gel running 0.5× Tris-borate buffer. Typically, 50% binding was obtained with reactions that contained 8.5 μg/ml of intact p53 or 23 μg/ml of core domain. For the competition experiments, the radiolabeled RGC site was first mixed with the appropriate unlabeled DNA and was then equilibrated with the protein. The amount of bound and free DNA was quantitated using a PhosphorImager (Fuji). The methylation interference analysis was performed essentially as described (Baldwin 1993). Briefly, the core domain–RGC complex was formed under conditions similar to those of Figure 2A (except the protein concentration was adjusted to give ~30% bound DNA). The free and core domain-bound RGC fragments were separated on a 4.5% polyacrylamide gel, and the DNA was eluted in 0.5 M ammonium acetate and 1 mM EDTA at 50°C. The DNA was cleaved at the methylated sites using piperidine, and equal amounts of labeled DNA were electrophoresed on a denaturing polyacrylamide (12%) sequencing gel.

DNA fragments

The DNA fragments used in this study were constructed by annealing purified synthetic oligonucleotides. The sequences of these fragments, reading along one strand, are as follows (the pentamer sequences are indicated in bold letters): RGC (42 bp), GATCCGAT**TGCCTTGCCTGGACTTGCCT**GGCCTTGCCTTTTG; mutant RGC (42 bp), GATCCGAT**TCCCTTCCGTGCAGTTCCGT**GGCCTTCCCTTTTG; [The wild-type RGC site contains 9 guanine residues that appear to be important for binding as demonstrated by methylation interference experiments (Kern et al. 1991), and the mutant RGC site has 6 of these guanine residues mutated to cytosines.] nonspecific DNA (38 bp), CGTACTTATCGAGCGGGGCGTAGTGATAGTTCTCTAG; no-pentamer (36 bp), GATCCGTATTCTCTTTTCCTAATTACAATTCGATTG; one pentamer (41 bp), GATCCGTATTCTCTTTTCCT**AGACA**AATTACAATTCGATTG; two pentamers (46 bp), GATCCGTATTCTCTTTTCCT**AGACATGCCT**AATTACAATTCGATTG; four pentamers (56 bp), GATCCGTATTCTCTTTTCCT**TGACTTGCCCAGACATGTTT**AATTACAATTCGATTG.

Plasma emission spectroscopy

Plasma emission spectroscopy was performed at the Institute of Ecology, University of Georgia, (Athens), using 31-element simultaneous, inductively coupled plasma (EPA method 6010). A typical analysis used a 1.5 ml solution of 46 μM core domain peptide (concentration was determined by amino acid composition analysis) in gel-filtration buffer and yielded 1.4 parts per million (21.4 μM) zinc. (This analysis has a detection limit for zinc of 0.05 ppm.) The stoichiometry corresponds to 0.47 mole of zinc per mole of core domain. An analysis of buffer showed that there were no contaminating metals.

Expression, purification, and cross-linking of the carboxy-terminal peptides

The cloning and expression of the carboxy-terminal peptides was done according to the procedures used for the core domain peptide. The p53(311–367) peptide was purified as follows: The *E. coli* lysate was acidified with 100 mM Na-acetate (pH 4.6), clarified by centrifugation, loaded onto a Mono S cation-exchange column running 100 mM Na-acetate (pH 4.6), and eluted with a NaCl gradient. The p53(311–393) peptide was extracted from the insoluble fraction of the *E. coli* lysate using 6.4 M guanidine-HCl, and was purified by reverse-phase HPLC on a C4 column (the mobile phase contained 0.1% trifluoroacetic acid, and the peptide was eluted with an acetonitrile gradient). Gel-filtration experiments with the carboxy-terminal peptides were performed using a Superdex 75 column with a buffer of 100 mM Tris-HCl (pH 7.6), and 200 mM NaCl. No changes in the elution profiles were observed over a concentration range of 1 μM –1 mM of either peptide. Cross-linking reactions contained 167 μM of peptide and glutaraldehyde (Sigma) in 50 mM HEPES-Na (pH 8.0), and 150 mM NaCl. After 30 min, the reactions were stopped by the addition of 100 mM Tris-HCl.

Acknowledgments

This work was supported by the Howard Hughes Medical Institute (C.O.P.) and by National Institutes of Health grant (GM31471) to C.O.P. We thank Jill Bargonetti and Carol Prives for providing us with immunopurified human p53 protein, Lena Nekludova for her assistance in interpreting some of the microchemical analysis data, and Svetlana Gorina and Alicia Russo for technical assistance.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

References

Baldwin, A.S. 1993. DNA–protein interactions. In *Current protocols in molecular biology* (ed. K. Janssen), p. 12.3. Wiley Interscience, New York.

Bargonetti, J., P.N. Friedman, S.E. Kern, B. Vogelstein, and C. Prives. 1991. Wild-type but not mutant p53 immunopurified proteins bind to sequences adjacent to the SV40 origin of replication. *Cell* **65:** 1083–1091.

Bargonetti, J., J. Reynisdottir, P.N. Friedman, and C. Prives. 1992. Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. *Genes & Dev.* **6:** 1886–1898.

El-Deiry, W.S., S.E. Kinzler, J.A. Pietenpol, K.W. Kinzler, and B. Vogelstein. 1992. Definition of a consensus binding site for p53. *Nature Genet.* **1:** 45–49.

Farmer, G., J. Bargonetti, H. Zhu, P. Friedman, R. Prywes, and C. Prives. 1992. Wild-type p53 activates transcription in vivo. *Nature* **358:** 83–86.

Fields, S. and S.K. Jang. 1990. Presence of a potent transcriptional activating sequence in the p53 protein. *Science* **249:** 1046–1048.

Foord, O.S., P. Bhattacharya, Z. Reich, and V. Rotter. 1991. A DNA binding domain is contained in the C-terminus of wild type p53 protein. *Nucleic Acids Res.* **19:** 5191–5198.

Friedman, P.N., X. Chen, J. Bargonetti, and C. Prives. 1993. The p53 protein is an unusually shaped tetramer that binds directly to DNA. *Proc. Natl. Acad. Sci.* **90:** 3319.

Funk, W.D., D.T. Pak, R.H. Karas, W.E. Wright, and J.W. Shay. 1992. A transcriptionally active DNA-binding site for human p53 complexes. *Mol. Cell. Biol.* **12:** 2866–2871.

Hainut, P. and J. Milner. 1993. A structural role for metal ions in the "wild type" conformation of the tumor suppressor protein p53. *Cancer Res.* **53:** 1739–1742.

Hartwell, L. 1992. Defects in a cell cycle checkpoint may be responsible for the genomic instability of cancer cells. *Cell* **71:** 543–546.

Hollstein, M., D. Sidransky, B. Vogelstein, and C.C. Harris. 1991. p53 mutations in human cancers. *Science* **253:** 49–53.

Hupp, T.R., D.W. Meek, C.A. Midgley, and D.P. Lane. 1992. Regulation of the specific DNA binding function of p53. *Cell* **71:** 875–886.

Kastan, M.B., Q. Zhan, W.S. El-Deiry, F. Carrier, T. Jacks, W.V. Walsh, B.S. Plunkett, B. Vogelstein, and A.J. Fornace Jr. 1992. A mammalian cell cycle checkpoint utilizing p53 and GADD45 is defective in ataxia-telangiectasia. *Cell* **71:** 587–597.

Kern, S.E., K.W. Kinzler, A. Bruskin, D. Jarosz, P. Friedman, C. Prives, and B. Vogelstein. 1991. Identification of p53 as a sequence specific DNA-binding protein. *Science* **252:** 1708–1711.

Kern, S.E., J.A. Pietenpol, S. Thiagalingam, A. Seymour, K.W. Kinzler, and B. Vogelstein. 1992. Oncogenic forms of p53 inhibit p53-regulated gene expression. *Science* **256:** 827–830.

Kuerbitz, S.J., B.S. Plunkett, W.V. Walsh, and M.B. Kastan. 1992. Wild-type p53 is a cell cycle checkpoint determinant following irradiation. *Proc. Natl. Acad. Sci.* **89:** 7491–7495.

Lane, D.P. 1992. p53, guardian of the genome. *Nature* **358:** 15–16.

Levine, A.J., J. Momand, and C.A. Finlay. 1991. The p53 tumor supressor gene. *Nature* **351:** 453–456.

Livingstone, L.R., A. White, J. Sprouse, E. Livanos, T. Jacks, and T.D. Tlsty. 1992. Altered cell cycle and gene amplification potential accompany loss of wild-type p53. *Cell* **70:** 923–935.

Milner, J. and E.A. Medcalf. 1991. Cotranslation of activated mutant p53 with wild type drives the wild-type p53 protein into the mutant conformation. *Cell* **65:** 765–774.

Momand, J., G.P. Zambetti, D.C. Olson, D. George, and A.J. Levine. 1992. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* **69:** 1237.

Oliner, J.D., K.W. Kinzler, P.S. Meltzer, D.L. George, and B. Vogelstein. 1992. Amplification of a gene encoding a p53 associated protein in human sarcomas. *Nature* **358:** 80–83.

Oliner J.D., J.A. Pietenpol, S. Thiagalingam, J. Gyuris, K.W. Kinzler, and B. Vogelstein. 1993. Oncoprotein MDM2 conceals the activation domain of tumor suppressor p53. *Nature* **362:** 857–860.

Raycroft, L., H. Wu, and G. Lozano. 1990. Transcriptional activation by wild-type but not transforming mutants of the p53 anti-oncogene. *Science* **249:** 1049–1051.

Shaulian, E., A. Zauberman, D. Ginsberg, and M. Oren. 1992. Identification of a minimal transforming domain of p53: Negative dominance through abrogation of sequence-specific DNA binding. *Mol. Cell. Biol.* **12:** 5581–5592.

Shaulian, E., A. Zauberman, J. Milner, E.A. Davies, and M. Oren. 1993. Tight DNA binding and oligomerization are dispensable for the ability of p53 to transactivate target genes and suppress transformation. *EMBO J.* **12:** 2789–2797.

Stenger, J., G.A. Mayr, K. Mann, and P. Tegtmeyer. 1992. Formation of stable p53 homotetramers and multiples of tetramers. *Mol. Carcinog.* **5:** 102–106.

Stürzbecher, H.W., R. Brain, C. Addison, K. Rudge, M. Remm, M. Grimaldi, E. Keenan, and J.R. Jenkins. 1992. A C-terminal α-helix plus basic region motif is the major structural determinant of p53 tetramerazation. *Oncogene* **7:** 1513–1523.

Unger, T., M.M. Nau, S. Segal, and J.D. Minna. 1992. p53: A transdominant regulator of transcription whose function is ablated by mutations occuring in human cancer. *EMBO J.* **11:** 1383–1390.

Vogelstein, B. and K.W. Kinzler. 1992. p53 function and dysfunction. *Cell* **70:** 523–526.

Yin, Y., M.A. Tainsky, F.Z. Bischoff, L.C. Strong, and G.M. Wahl. 1992. Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. *Cell* **70:** 937–948.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots.

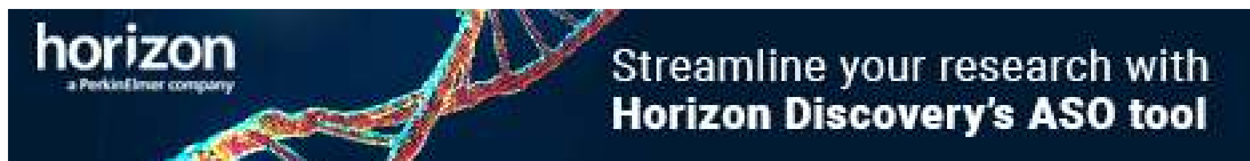
N P Pavletich, K A Chambers and C O Pabo

Genes Dev. 1993, **7:**
Access the most recent version at doi:10.1101/gad.7.12b.2556

References This article cites 31 articles, 11 of which can be accessed free at:
http://genesdev.cshlp.org/content/7/12b/2556.full.html#ref-list-1

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**



Copyright © Cold Spring Harbor Laboratory Press