Crystal Structure of a p53 Tumor Suppressor–DNA Complex: Understanding Tumorigenic Mutations

Yunje Cho, Svetlana Gorina, Philip D. Jeffrey, Nikola P. Pavletich

Mutations in the p53 tumor suppressor are the most frequently observed genetic alterations in human cancer. The majority of the mutations occur in the core domain which contains the sequence-specific DNA binding activity of the p53 protein (residues 102–292), and they result in loss of DNA binding. The crystal structure of a complex containing the core domain of human p53 and a DNA binding site has been determined at 2.2 angstroms resolution and refined to a crystallographic *R* factor of 20.5 percent. The core domain structure consists of a β sandwich that serves as a scaffold for two large loops and a loop-sheet-helix motif. The two loops, which are held together in part by a tetrahedrally coordinated zinc atom, and the loop-sheet-helix motif form the DNA binding surface of p53. Residues from the loop-sheet-helix motif interact in the major groove of the DNA, while an arginine from one of the two large loops interacts in the minor groove. The loops and the loop-sheet-helix motif consist of the conserved regions of the core domain and contain the majority of the p53 mutations identified in tumors. The structure supports the hypothesis that DNA binding is critical for the biological activity of p53, and provides a framework for understanding how mutations inactivate it.

Inactivation of the p53 tumor suppressor is a common event in the development of diverse types of human cancers. About half of all cancer cases involve missense mutations of one p53 allele coupled to the deletion of the second allele (*1*), and many of the remaining cases involve cellular or viral oncogenes that inactivate p53 (*2*). Wild-type p53 has been shown to inhibit neoplastic transformation (*3*). It can block transformation by activated oncogenes in cell culture (*4*), can inhibit the growth of tumor cells in vitro (*5*), and can prevent tumor formation in animal models (*6*). No such antitumor responses can be elicited from tumor-derived p53 mutants (*4–6*). Furthermore, transgenic mice lacking p53 are prone to the spontaneous development of tumors at a very early age (*7*), demonstrating the indispensable role p53 plays in preventing cancer.

Recent data suggest that p53 controls a cell cycle checkpoint responsible for maintaining the integrity of the genome. It has been shown that p53 induces cell cycle arrest in response to DNA damage (*8*). Mutations in p53 eliminate this response and result in an enhanced frequency of genomic rearrangements (*8*, *9*). It is thought that the resulting genetic instability increases the probability that the tumor cells escape the normal restrictions against excessive growth (*10*).

The wild-type p53 protein can bind to specific DNA sequences (*11*, *12*) and activate the transcription of genes containing p53 binding sites (*13–16*). Studies of tumor-derived p53 mutants have shown that they are defective in sequence-specific DNA binding, and consequently they cannot activate the transcription of genes (*14–16*). These studies have highlighted sequence-specific DNA binding and transactivation as key biochemical activities mediating the biological effects of p53. This hypothesis has gained further support by the demonstration that p53 induces the expression of the Cip1 cell cycle inhibitor, which results in growth arrest (*17*).

More than a thousand p53 mutations have been identified in human tumors (*1*, *18*). The majority of these mutations are found in the central 200–amino acid portion of p53, and they are particularly common in the four conserved regions (Fig. 1A). Recent biochemical studies have demonstrated that the core portion of p53 (residues 102–292) folds into a compact structural domain that contains the sequence specific DNA binding activity of the protein (*19–21*). These findings have identified the core domain as holding the key to understanding how p53 binds DNA and how tumorigenic mutations inactivate it. To address these questions, we have crystallized the core domain of human p53 with a consensus DNA binding site, determined the structure of this complex at 2.2 Å resolution, and refined it to an *R* factor of 20.5 percent. We now report the structure of this complex and discuss the implications for our understanding of the function of p53 and its inactivation in tumors.

Structure determination. The identification and characterization of the core domain of p53 have been described (*19*). Briefly, proteolytic digestion of the human p53 protein showed that residues 102 to 292 constitute an independently folded, compact structural domain. The isolated core domain lacks the oligomerization domain located near the carboxyl-terminal portion [residues 311–364 (*19*)], and it exists as a monomer in solution (*19*). Gel mobility shift and methylation interference experiments showed that the core domain expressed in *Escherichia coli* binds to DNA

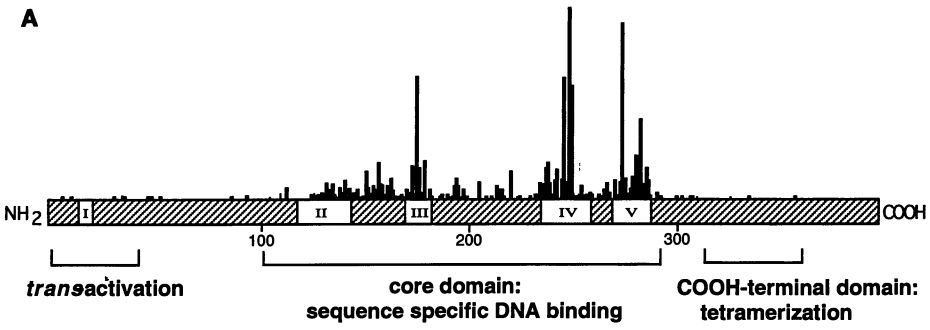

Fig. 1. The p53 core domain and the DNA binding site used in the cocrystallization. (**A**) Domains of p53 (*16*, *19*). Boxes with roman numerals indicate the five regions of the gene that are conserved across species (*1*), and the bar graph shows the approximate position and frequency of the tumor-derived mutations (*18*). (Conserved region I is outside the core domain.) (**B**) DNA duplex used in the cocrystallization. A p53 half-site containing two consensus pentamer sequences is in bold letters [defined according to (*12*)]. The sequences where two core domain molecules bind are underlined. One of the core domain molecules binds a site at the interface of two DNA fragments related by crystallographic symmetry. In the text, bases of the upper strand in the figure are referred to by base pair number, and bases of the lower strand are referred to by base pair number followed by the prime symbol (′).

The authors are in the Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA.

with sequence specificity and affinity similar to that of intact p53 (*19*). These studies also showed that p53 contains a tightly bound zinc atom which is necessary for the DNA binding activity of the protein (*19*).

Crystals of the core domain–DNA complex were grown from an *E. coli* expressed fragment of the human p53 protein corresponding to residues 94 to 312, and a 21–base pair (bp) DNA duplex containing a p53 half-site (Fig. 1B). A half-site, which was shown to be sufficient for high-affinity binding by the core domain (*19*), corresponds to two consensus pentamer sequences instead of the four typically found in p53 binding sites (*12*). Crystals of the complex were grown at 4°C in an anaerobic chamber by the hanging drop vapor diffusion method. The best crystals were obtained by mixing a 0.7 mM solution of the complex containing a 1 molar equivalent of the DNA fragment [in 50 mM bis-tris propane-HCl (BTP), 100 mM NaCl, 10 mM dithiothreitol (DTT), pH 6.8] with an equal volume of the buffer used in the crystallization well {12 to 15 percent polyethylene glycol (PEG) 400, 100 mM 2-[*N*-morpholino]ethanesulfonic acid-Na^+ (MES), 50 mM BTP-HCl, 10 mM DTT, pH 6.4}; the crystals grew over the course of 4 weeks. The crystals form in the space group C2 with a = 120.4 Å, b = 68.8 Å, c = 112.2 Å, and β = 107.1° and have three core domain molecules and one DNA fragment in the asymmetric unit. The crystals diffract to about 2.2 Å resolution, but radiation damage limits the useful resolution of data collected at 1°C to about 2.8 Å (Table 1). Multiple isomorphous replacement (MIR) phases were calculated with data from several mercury and iodine derivatives (5-iodouracil was substituted for thymine at specific positions), and these phases produced electron density maps with a mean figure of merit of 0.73 to 3.2 Å (Table 1 and Fig. 2). Partial models containing approximately 40 percent of the protein backbone atoms were used to determine the noncrystallographic symmetry operators relating the three core domain molecules. These operators were used to average the electron density with the program RAVE (*22*). The averaging showed that the structures of the three monomers were similar, with pairwise comparisons of MIR electron densities resulting in correlation coefficients of about 0.56 and R factors of about 16 percent. The averaged map was used in conjunction with the initial MIR map to complete the model of the asymmetric unit. This model was refined by simulated annealing [with the program X-PLOR (*23*)] to an R factor of 22.0 percent for data between 8.0 and 2.8 Å.

The 2.8 Å model was used as the starting point for the refinement against 2.2 Å resolution data collected with a crystal frozen at −175°C. The crystal was prepared for freezing by equilibrating it in a buffer containing 12.5 percent isopropanol, 25 percent PEG 400, 50 mM MES-Na^+, 50 mM BTP-HCl, and 10 mM DTT, pH 6.4. The crystal was then placed in a thin wire loop and was frozen in a stream of nitrogen gas at −175°C. Upon freezing, the unit cell changed to a = 117.5 Å, b = 67.9 Å, c = 108.8 Å, and β = 105.5°. A local rotation-translation search was used to optimize the

Table 1. Data collection and analysis. Diffraction data were collected at 1°C with the Rigaku R-AXISIIC imaging plate area detector mounted on a Rigaku RU200HB rotating anode x-ray generator. Initial MIR phases were calculated at 3.2 Å resolution with the program PHARE [CCP4 package of crystallographic programs (*46*)] and had a mean figure of merit of 0.73. The MIR map was improved by solvent flattening with Wang's protocol (*47*), and this map was used to build partial models of the three core domain molecules and the DNA by means of the programs O (*48*) and CHAIN (*49*). Noncrystallographic symmetry averaging with the program RAVE (*22*) improved the electron density for the core domain, especially in the β-sandwich portion of the structure, and the averaged map was used in conjunction with the initial MIR map to complete the model. After several cycles of refinement with the full model, X-PLOR omit maps were used to systematically check every part of the complex. About 2 to 5 percent of the structure was deleted in each calculation, and simulated annealing was used to reduce model bias in the omit maps. To correct for anisotropic diffraction and for absorption problems during data collection with the frozen crystal, a local scaling program that scales the observed and calculated structure factor amplitudes was used (*50*). The current model contains residues 94 to 289 of the core domain and 21 base pairs of the DNA fragment. The 23 COOH-terminal residues (290–312) do not have interpretable electron density, and we presume that these residues are disordered in the crystal. Base pairs 14 through 18 of the DNA, a region that is not bound by the core domains, have weak electron density in the MIR and omit maps. This region also has high-temperature factors in the refined model and thus appears to have static disorder in the crystal.

	Resolution (Å)	Reflections		Data coverage (%)	R_{sym} (%)	Heavy atom sites	MIR analysis (20–3.2 Å)	
		Measured	Unique				Mean isomorphous difference	Phasing power
Native	2.8	36838	20336	92.2	5.9			
Native (−175°C)	2.2	123663	36754	86.0	5.7			
$HgCl_2$	3.0	32588	16690	93.9	7.5	4	0.25	1.98
K_2HgI_4-1	3.0	18696	13978	78.0	9.7	2	0.17	1.14
K_2HgI_4-2	3.0	30443	15960	89.0	13.6	4	0.23	2.26
K_2PtCl_4*	3.0	26025	15535	85.6	10.9	1	0.15	0.50
IdU (2, 6, 4′, 20′)	2.8	36210	19515	88.0	6.6	4	0.11	1.87
IdU (4′, 20′)	2.5	42508	25553	86.0	6.5	2	0.08	1.14
IdU (3′, 21′)	2.8	34520	19325	92.3	7.3	2	0.10	1.09
IdU (21′)	2.5	25990	16572	76.6	5.5	1	0.09	0.91
IdU (3′)	2.5	44521	24795	89.6	5.9	1	0.08	0.84
IdU (6)	2.8	26490	18237	82.4	6.2	1	0.09	0.83

Refinement:

Resolution (Å)	Reflections ($F > 2\sigma$) (No.)	Total number of		R factor (%)	rms†		rms‡ B values (Å²)
		Atoms	Water molecules		Bond lengths (Å)	Bond angles (°)	
6.0–2.2	34275	5813	380	20.5	0.011	2.76	3.62

*The K_2PtCl_4 derivative data were used to 4.0 Å resolution in the MIR analysis. †The rms deviation in bond lengths and bond angles from ideal values. ‡The rms deviation in temperature factors between bonded atoms. The rms in $R_{sym} = \Sigma_h\Sigma_i|I_{h,i} - I_h|/\Sigma_h\Sigma_i I_{h,i}$, where I_h is the mean intensity of the i observations of reflection h. Mean isomorphous difference = $\Sigma|F_{PH} - F_P|/\Sigma F_{PH}$, where F_{PH} and F_P are the derivative and native structure factor amplitudes, respectively. Phasing power = $[(F_{H(calc)})^2/(F_{PH(obs)} - F_{PH(calc)})^2]^{1/2}$.

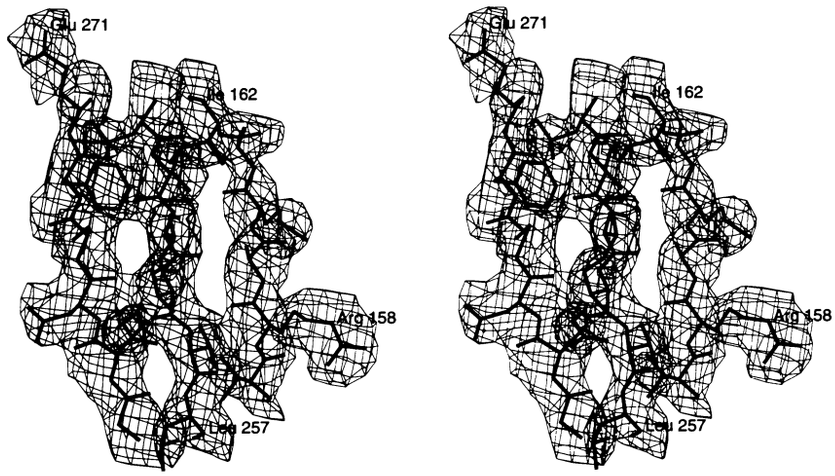

Fig. 2. Multiple isomorphous replacement (MIR) electron density map of the p53 core domain at 3.2 Å resolution. This view shows a portion of the structure from the β sandwich of the core domain. The map is contoured at 1.2σ, and the 2.8 Å resolution atomic model is shown as a stick figure. The 2.2 Å resolution model is not shown because it was refined against data collected at −175°C and its position is shifted with respect to the MIR map calculated with data collected at 1°C. For the MIR analysis, the mercury and platinum derivatives were prepared by incubating the crystals in a solution containing 15 percent PEG400, 50 mM MES-Na^+, 50 mM BTP-HCl, 1 mM β-mercaptoethanol, and one of the following heavy atom compounds: 0.4 mM $HgCl_2$ for 4 hours, 0.5 mM K_2HgI_4 for 6 hours (K_2HgI_4–1 in Table 1), 0.5 mM K_2HgI_4 for 12 hours (K_2HgI_4–2 in Table 1), and 0.7 mM K_2PtCl_4 for 6 hours. Three of the mercury atoms bind a site consisting of Cys^{124}, Cys^{135}, and Cys^{141} on each core domain, and a fourth mercury atoms binds Cys^{277} of the core domain in the consensus complex. Iodine derivatives were prepared by substituting 5-iodouracil for thymine at specific positions on the DNA (base pair numbers for the substituted thymines are indicated in Table 1). The diffraction data from the native and derivative crystals contained nonisomorphous differences in the high-resolution shells limiting the use of the derivative data to 3.2 Å resolution.

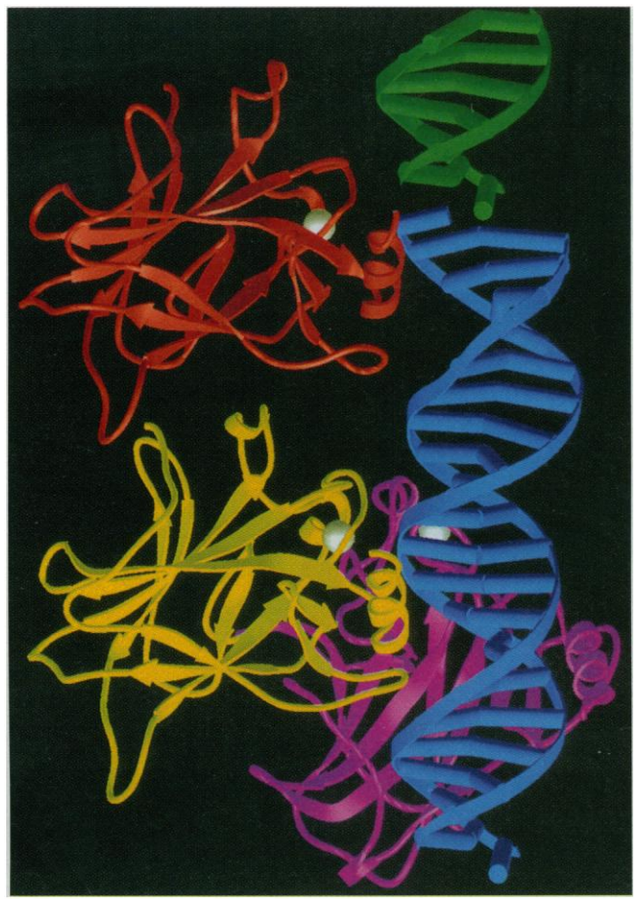

Fig. 3. Schematic ribbon drawing of the asymmetric unit, which contains three p53 core domain molecules and one DNA duplex. Two of the core domains bind DNA (blue); one (yellow) interacts extensively with a consensus binding site, and the other (red) binds at a nonconsensus site at the interface of DNA fragments related by crystallographic symmetry (a portion of the symmetry-related DNA fragment is shown in green). The third core domain molecule (purple) does not bind DNA, but makes protein-protein contacts stabilizing crystal packing. The zinc atoms are shown as white spheres.

2.8 Å model, and this was followed by several rounds of simulated annealing with X-PLOR and then least-squares refinement with the program TNT (*24*). The structure presented here has an *R* factor of 20.5 percent for data from 6.0 to 2.2 Å. The root-mean-square (rms) deviation from ideal bond lengths is 0.011 Å, and the rms deviation from ideal bond angles is 2.76°; the rms deviation in the temperature factors between bonded atoms is 3.62 $Å^2$.

Overall structure of the p53 core domain–DNA complex. The crystals contain one DNA duplex and three p53 core domain molecules in the asymmetric unit (Fig. 3). The DNA has one-base overhangs, and the overhang base pairing facilitates the end-to-end packing of the DNA fragment to form a pseudo-continuous double helix that runs through the crystal. One of the core domain molecules binds a consensus site near the center of the DNA duplex and makes extensive contacts to the bases and the phosphate backbone. We find another core domain molecule binding 11 bp away, in a region of weak homology to the consensus sequence (this site is at the interface of two DNA duplexes related by crystallographic symmetry). The two core domain molecules bind on the same face of the duplex and interact weakly with each other, forming a head-to-tail dimer. They have similar docking arrangements with the DNA, but the nonconsensus complex makes fewer contacts with the bases compared to the consensus complex. A third core domain molecule does not bind DNA, but makes protein-protein contacts that stabilize crystal packing.

The three core domain molecules have similar overall structures, and it appears that DNA binding does not result in any significant structural changes. The Cα atoms of the core domain in the consensus complex can be superimposed on the Cα atoms of the core domain that does not bind DNA with an rms deviation of 0.75 Å; similar rms deviations are obtained for the other pairwise superpositions. The structure of the core domain contains a sandwich of two antiparallel β sheets that have four and five β strands, and a loop-sheet-helix motif that packs tightly against one end of the β sandwich (Fig. 4, A to C). At this end of the β sandwich, there are two large loops that are held together in part by a tetrahedrally coordinated zinc atom. Three cysteines and a histidine are the ligands for the zinc atom.

Although the β sandwich comprises a major part of the core domain structure, it is not directly involved in DNA binding. Instead, the core domain uses the loop-sheet-helix motif and one of the two large loops to bind DNA. Recognition relies on both the major and minor grooves of the

DNA. The helix and the loop from the loop-sheet-helix motif fit in the major groove and make contacts to the edges of the base pairs, while the large loop provides an arginine that fits in the adjacent minor groove. Residues from both the loop-sheet-helix motif and the large loop make contacts to the DNA backbone that divides the major and minor grooves in this region.

The most striking observation is that the portion of the structure involved in DNA binding consists of the conserved regions of the core domain and contains the majority of the mutation hotspots observed in tumors (*18*). This portion of the structure consists of (i) the loop-sheet-helix motif and the large loop that interact with the DNA, (ii) the second large loop that packs against the DNA binding large loop, and (iii) the end of the β sandwich that acts as a scaffold for these structural elements. In general, mutations are most frequent in the regions of the core domain that are closest to the DNA; the frequency of mutations decreases substantially as we move away from the DNA, toward the other end of the β sandwich.

Structure of the p53 core domain. The major part of the core domain structure consists of two antiparallel β sheets that have four and five β strands (Fig. 4, A to C). The sheets pack face-to-face across an extended hydrophobic core, forming a β sandwich with a "Greek key" topology. At one end of the β sandwich, the sheets pack tightly against each other, forming a compact barrel-like structure. This region contains the NH_2-termini of β strands S1, S8, S6, S4, and S10, and the COOH-termini of β strands S3, S5, S7, and S9 connected by short loops of 5 to 11 residues. The β strands are highly twisted, and their directions diverge as they reach the other end of the β sandwich. Here, the sheets open up and give the impression of being frayed. This end of the β sandwich is decorated with two long loops of 15 and 32 residues, and a loop–β sheet–α helix motif that fits into an opening between the sheets (Fig. 4, A to C). Our description of the structure focuses on this region because it is the loop-sheet-helix motif and the two large loops that make up the DNA binding surface of the protein. This portion of the structure also provides crucial insights into the inactivation of p53 by mutations because it corresponds to the four conserved regions of the core domain where most of the mutations are found (Fig. 1A).

The loop-sheet-helix motif contains a short, three-stranded β sheet consisting of the S2-S2′ hairpin and the four COOH-terminal residues of the extended S10 β strand. The H2 α helix packs against the S2-S2′ hairpin, and backbone atoms in the L1 loop make hydrogen bonds with the NH_2-terminus of the α helix (Fig. 4, A to C). This motif is anchored to the β sandwich by (i) the S10 strand, which participates both in the small three-stranded sheet and in the β sandwich, and (ii) hydrophobic interactions between the back of the hairpin and a pocket formed by the S1, S3, and S8 strands. This arrangement leaves the hairpin buried and positions one face of the α helix and the loop on the surface of the core domain. The L1 loop (residues 112–124) and the S2-S2′ β hairpin that follows it (residues 124–141) correspond almost exactly to conserved region II (residues 117–142). The end of the S10 strand

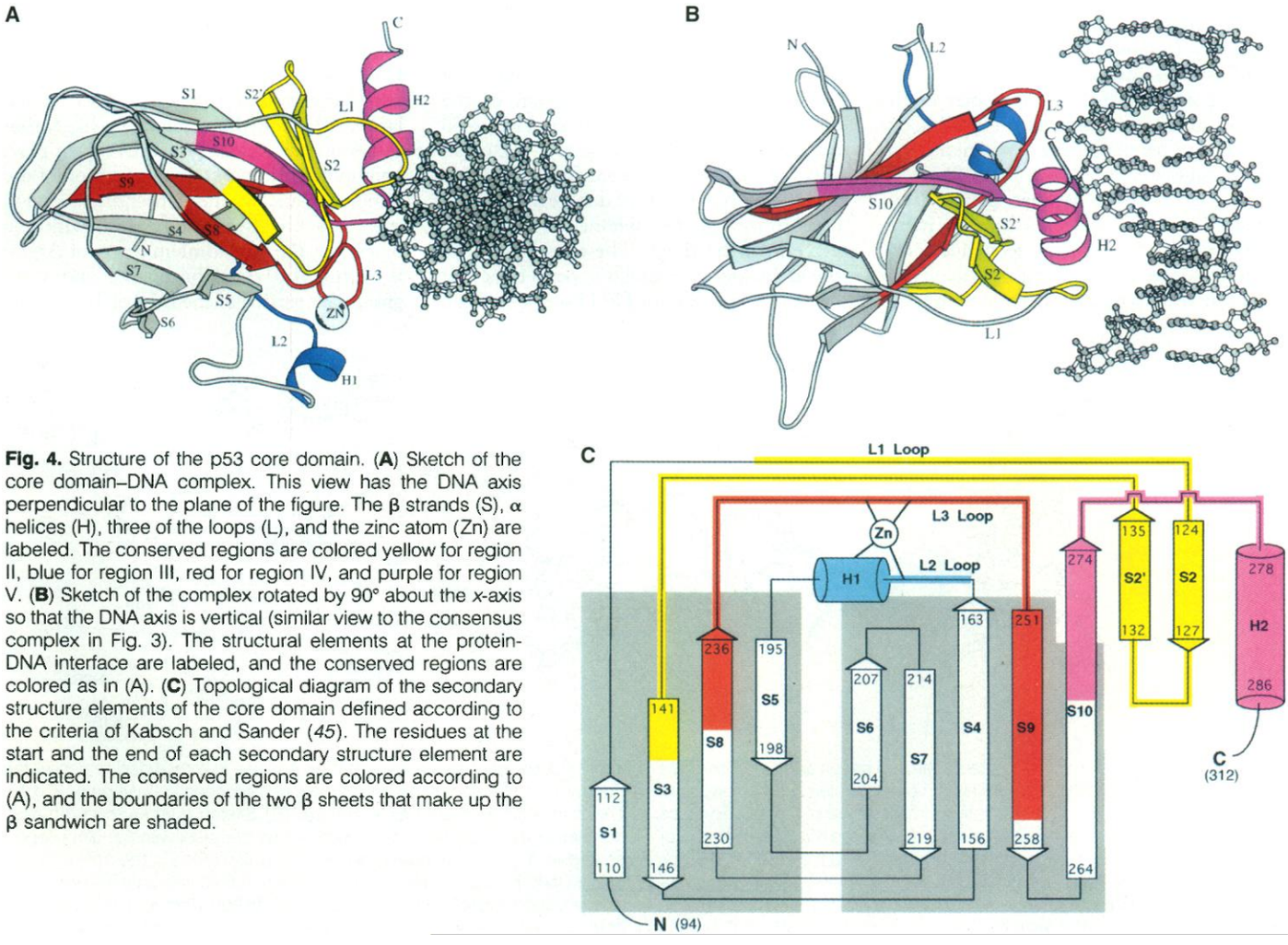

Fig. 4. Structure of the p53 core domain. (**A**) Sketch of the core domain–DNA complex. This view has the DNA axis perpendicular to the plane of the figure. The β strands (S), α helices (H), three of the loops (L), and the zinc atom (Zn) are labeled. The conserved regions are colored yellow for region II, blue for region III, red for region IV, and purple for region V. (**B**) Sketch of the complex rotated by 90° about the *x*-axis so that the DNA axis is vertical (similar view to the consensus complex in Fig. 3). The structural elements at the protein-DNA interface are labeled, and the conserved regions are colored as in (A). (**C**) Topological diagram of the secondary structure elements of the core domain defined according to the criteria of Kabsch and Sander (*45*). The residues at the start and the end of each secondary structure element are indicated. The conserved regions are colored according to (A), and the boundaries of the two β sheets that make up the β sandwich are shaded.

(residues 271–274) and the H2 α helix (residues 278–286) occur later in the amino acid sequence, but they too correspond to a conserved region, region V (residues 270–286).

Of the two large loops at the decorated end of the β sandwich, the longer L2 loop occurs between β strands S4 and S5, whereas the shorter L3 loop occurs between S8 and S9 (Fig. 4, A to C). Both loops have little regular secondary structure. L2 starts with a turn and continues in three extended segments interrupted by a short helix (H1) and another turn. L3 is less extended, and contains three turns. In these loops, the lack of extensive backbone hydrogen bonds appears to be compensated, in part, by (i) the binding of a zinc atom shared by both loops, and (ii) several side chain–side chain and side chain–backbone interactions in between the loops, and between the loops and the end of the β sandwich. The ligands for the tetrahedrally coordinated zinc atom are Cys^{176} of the L2 loop, His^{179} of the H1 helix (in the L2 loop), and Cys^{238} and Cys^{242} of the L3 loop. As was observed with the loop-sheet-helix motif in this region, the L2 and L3 loops also consist of conserved residues of p53. L2 (residues 163–195) contains region III (residues 171–181) and L3 (residues 236–251) coincides with region IV (residues 234–258).

p53-DNA interactions. Typically, p53 binding sites consist of four copies of the pentamer consensus sequence PuPuPuC(A/T) (*12*). The pentamers are oriented in alternating directions, with some sites containing up to 13 bp between pairs of pentamers (→ ← X → ←; each arrow represents a pentamer consensus sequence and X represents variable spacing). The crystal structure reveals that the core domain binds primarily a single pentamer consensus sequence, but base pairs outside the pentamer consensus also participate in binding. This observation suggests that part of an adjacent pentamer is also required for binding [PuPuPuC(A/T)-**(T/A)G**]. Of the two core domains, one binds a consensus site making extensive DNA contacts, while the other binds a nonconsensus site making fewer DNA contacts. In our discussion of DNA binding, we focus on the consensus complex, and point out differences with the nonconsensus complex as they occur.

Protein-DNA interactions can best be described in three principal parts: (i) major groove contacts in the pentamer sequence from the H2 helix and the L1 loop, (ii) minor groove contacts in the A·T–rich region of the binding site from the L3 loop, and (iii) phosphate contacts to the DNA strand flanked by the major and minor groove contacts.

In the major groove, two of the three variable bases [Pu**PuPu**C(A/T)-(T/A)G] of the consensus sequence make side chain contacts (Fig. 5, A and B): Lys^{120} from the L1 loop donates hydrogen bonds to the O6 and N7 of Gua^{8}, and Cys^{277}, which immediately precedes the α helix, accepts a hydrogen bond from N4 of $Cyt^{9'}$ on the pyrimidine-rich strand. Both of these contacts are consistent with the variability of the base sequence in this region of the consensus site. Thus, Lys^{120} could instead donate a hydrogen bond to the N7 of an adenine (in roughly 30 percent of the sites), and Cys^{277} could instead donate a hydrogen bond to the O4 of a thymine (in roughly 50 percent of the sites). The most critical of the major groove contacts appears to be one made to the invariant C·G base pair of the pentamer consensus [PuPuPu**C**(A/T)-(T/A)G]; Arg^{280} from the NH_2-terminal portion of the α helix donates a pair of hydrogen bonds to the N7 and O6 of $Gua^{10'}$ on the pyrimidine-rich strand. Furthermore, the Arg^{280} side chain is buttressed by a salt bridge with the carboxylate of Asp^{281}. None of these contacts are made in the nonconsensus complex. Here, the H2 helix still fits in the major groove, but is further away from the edges of the bases. In this region, the DNA is somewhat distorted, due in part to stacking with a symmetry-related DNA fragment.

The minor groove interaction occurs in the A·T rich region of the consensus sequence [PuPuPuC**(A/T)-(T/A)G**] and involves Arg^{248} from the L3 loop of the core domain reaching into the minor groove (Fig. 5, A and B). We presume that this interaction plays a critical role in DNA binding because Arg^{248} is the most frequently mutated residue of p53 in human cancers (*1*, *18*). A striking feature of this interaction is compression of the minor groove which allows for the tight packing of Arg^{248} against the sugar and phosphate groups inside the minor groove. The width of the minor groove is 9.3 Å (distance between the phosphate groups of $Thy^{11'}$ and Cyt^{15}) compared to 11.5 Å for B-type DNA (*44*). In this region, the base pairs have large propeller twists (up to 25°) and large buckle dihedral angles. It is likely that the A-T–rich sequence in this region plays an important role in the compression of the minor groove. In the consensus complex, there are no direct interactions with the bases, but the guanidinium group of Arg^{248} is within 4.0 Å of phosphate and sugar groups (in particular the O1P of $Thy^{11'}$, the

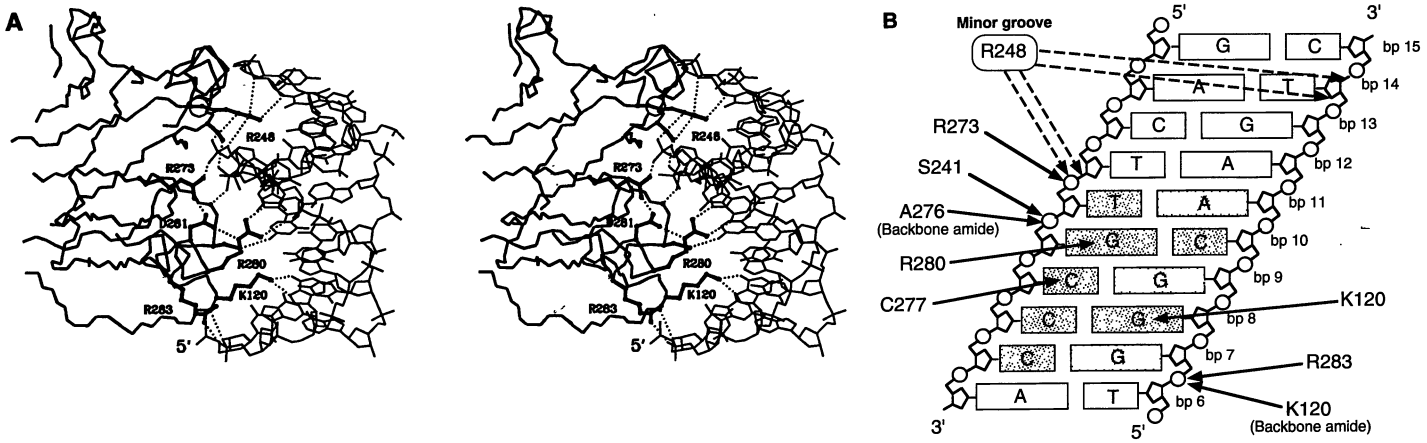

Fig. 5. The core domain makes contacts with the bases and the phosphate backbone of the DNA. (**A**) Stereo diagram of the protein-DNA interface of the consensus complex in an orientation similar to that of Fig. 4B. To make the contacts easier to see, the DNA is tilted. Backbone atoms are shown for part of the β sandwich, the loop-sheet-helix motif, and the L2 and L3 loops (residues 113–143, 160–184, 191–198, 213–214, 233–254, and 270–289), and side chains are shown for residues involved in DNA binding. The side chains Ser^{241} (in the vicinity of Arg^{248}) and Cys^{277} (behind Arg^{280}) could not be labeled because of their position in a crowded region of the diagram. The portion of the DNA shown corresponds to base pairs 6 to 15. The zinc atom is shown as a sphere. (**B**) Sketch summarizing the core domain–DNA interactions in the consensus complex with the DNA represented as a cylindrical projection. Arg^{248}, which binds in the minor groove of the DNA, is circled, and its long-range (<4.0 Å) interactions with sugar and phosphate groups are indicated by dashed arrows. The base pairs corresponding to the pentamer consensus sequence are shadowed (*12*).

O3′ of Thy$^{12'}$, and the O3′ and O4′ of Thy^{14}). We also observe a water molecule between the guanidinium group of the arginine and the N2 of Gua^{13} making hydrogen bonds that bridge the two.

In contrast to the major groove contacts, the minor groove interactions are more extensive in the nonconsensus complex, with the guanidinium group of Arg^{248} donating a pair of hydrogen bonds to the O2 of Thy^{2}, and to the N3 of Ade^{3} (all occurring on the symmetry-related DNA fragment). The differences in the minor groove interactions in the two complexes are correlated with the sequence of the base pair occurring immediately after the pentamer sequence [PuPuPuC(A/T)-**(T/A)**G]. It appears that a thymidine at this position (nonconsensus complex) allows more favorable interactions than an adenine (consensus complex).

DNA backbone contacts primarily involve the phosphate groups of Gua$^{10'}$ and Thy$^{11'}$ (Fig. 5, A and B). The phosphate group of Gua$^{10'}$ is bound by the Oγ group of Ser^{241} from the L3 loop and by the backbone amide of Ala^{276} from the loop-sheet-helix motif. The phosphate group of Thy$^{11'}$ is similarly involved, being bound by the Nη1 of Arg^{273} from the loop-sheet-helix motif. Arg^{273} is noteworthy because in addition to contacting a phosphate, it takes part in an extended network of interactions. The guanidinium group of Arg^{273} makes a salt bridge with the carboxylate of Asp^{281} which, as mentioned earlier, also interacts with Arg^{280}, the latter making a major groove base contact. This network of interactions bridges the phosphate group of Thy$^{11'}$; the side chains of Arg^{273}, Asp^{281}, and Arg^{280}; and the N7 and O6 groups of Gua^{10}. The Arg^{273} residue is frequently mutated in tumors (*1*, *18*), and thus we presume that it plays a critical role in DNA binding.

The core domain makes two more phosphate contacts several base pairs away (Fig. 5, A and B): the backbone amide of Lys^{120} contacts the O3′ of Thy^{6}, and the guanidinium group of Arg^{283} contacts the phosphate of Gua^{7}.

Overall, a similar set of phosphate contacts is maintained in the nonconsensus complex, except that Arg^{273} contacts an O3′ atom instead of a phosphate group (the phosphate it would contact is missing because the base is at the end of the DNA fragment).

Structures of mutation hotspots. A key finding that emerges from the structure is that the residues most frequently mutated in cancer (*1*, *18*) are at or near the protein-DNA interface, whereas the residues that are not mutated, or mutated the least frequently, are in general far from the DNA. Mutations are most frequent in (i) the L3 loop, where 30 percent of the tumor derived mutations are found (*18*), (ii) the H2 α helix and the COOH-terminal portion of the S10 β strand (from the loop-sheet-helix motif) with 25 percent of all mutations, and (iii) the L2 loop with 17 percent of the mutations (Fig. 6A). The L3 loop and the loop-sheet-helix motif provide the minor and major groove contacts, respectively, and they also make the critical phosphate contacts. The L2 loop, although it does not directly interact with the DNA, is involved in extensive interactions with the L3 loop.

Among the frequently mutated residues, six hotspots stand out (Fig. 6A) (*18*). These hotspots are Arg^{248} with 9.6 percent of the p53 mutations, Arg^{273} with 8.8 percent, Arg^{175} with 6.1 percent, Gly^{245} with 6.0 percent, Arg^{249} with 5.6 percent, and Arg^{282} with 4 percent (Fig. 6, A and B). The structure reveals that the two most frequently mutated residues directly contact the DNA. As discussed earlier, Arg^{248} (L3 loop) makes the minor groove contact (Figs. 5, A and B, and 6C), and Arg^{273} (from the loop-sheet-helix motif) contacts a backbone phosphate (Figs. 5, A and B, and 6D). The remaining four hotspot residues appear to play a critical role stabilizing the structure of the DNA binding surface of p53. Three are arginines, and their side chains participate in van der Waals, electrostatic, and hydrogen bonding interactions with other side chains and with backbone carbonyl groups. Structurally, these interactions are noteworthy because the arginines make full use of the hydrogen bond donor potential of their guanidinium groups, and many of these hydrogen bonds are made to carbonyl groups of the polypeptide backbone.

Arg^{175} occurs in the L2 loop in the vicinity of the zinc binding site, and for the most part is buried away from the protein surface (Fig. 6, B and E). It is surrounded by portions of the L2 and L3 loops and is involved in interactions bridging the two. The guanidinium group donates a pair of hydrogen bonds (Nε and Nη2) to the backbone carbonyl of Met^{237} on the L3 loop, and a bifurcated hydrogen bond (Nη1) to the backbone carbonyl of Pro^{191} and the Oγ of the Ser^{183} side chain which are on the L2 loop.

Arg^{249} occurs in the L3 loop, adjacent to the minor groove contact (Arg^{248}), and it is surrounded by portions of the L2 and L3 loops and the COOH-terminal end of the S3 strand from the β sandwich (Fig. 6, B and F). The Cδ of the arginine side chain makes a van der Waals contact with the His^{162} side chain (L2 loop), and the guanidinium group stacks against the Tyr^{163} side chain (S3 strand). The guanidinium group also donates a hydrogen bond (Nε) to the backbone carbonyl of Met^{246} (L3 loop), and another hydrogen bond (Nη1) to the backbone carbonyl of Gly^{245} (L3 loop). In addition, it makes a salt bridge with the carboxylate of Glu^{171} (L2 loop).

Arg^{282} from the H2 helix plays a structural role in the loop-sheet-helix motif, being involved in the packing of the H2 helix against the β hairpin and the L1 loop (Fig. 6, B and G). The aliphatic portion of the side chain makes a van der Waals contact with the Phe^{134} side chain, the Cγ of Thr^{125}, and the Cβ of Ser^{127} (all occurring on the β hairpin). The guanidinium group also donates a hydrogen bond (Nη1) to the backbone carbonyl of Tyr^{126} (β hairpin), a bifurcated hydrogen bond (Nε) to the carboxylate of Glu^{286} (helix H2) and the Oγ of Ser^{127} (β hairpin), and a hydrogen bond (Nη2) to the Oγ1 of Thr^{118} (L1 loop).

Among the six mutation hotspots, Gly^{245} (L3 loop) is the only non-arginine residue (Fig. 6, B and H). It appears to play a critical role because it allows the L3 loop to assume a conformation not favored with a residue containing a side chain, the phi and psi dihedrals being −118° and −113°, respectively. This backbone conformation, in turn, is important in allowing the formation of two hydrogen bonds: one between the backbone amide of Gly^{245} and the backbone carbonyl of Cys^{247} (one of the zinc ligands), and the other between the backbone carbonyl of Gly^{245} and the guanidinium group of Arg^{249} (one of the six hotspots discussed earlier). Furthermore, the Cα of Gly^{245} is in close proximity to the L2 loop, and thus this region cannot accommodate a side chain without disrupting the loop.

Apart from the six mutation hotspots, other frequently mutated residues include [Fig. 6A (*18*)] Arg^{280} on the H2 helix which contacts the invariant guanine in the major groove of the DNA (with 2.1 percent of the mutations), and the ligands for the zinc: Cys^{176} (1.5 percent) and His^{179} (1.9 percent) on the L2 loop, and Cys^{238} (1.8 percent) and Cys^{242} (1.4 percent) on the L3 loop.

In contrast to the DNA binding surface, the β sandwich is not a frequent target of mutations. The few mutations that are found in this region generally involve residues in the hydrophobic core of the β sandwich, and are more frequent in the portion of the sandwich closest to the DNA binding surface. This portion of the hydrophobic core contains Cys^{141}, Val^{143}, Val^{157}, Ile^{195}, Val^{197}, Tyr^{234}, Tyr^{236}, and Phe^{270}, and mutations in these residues correspond to roughly 6 percent of all p53 mutations [Fig. 6A (*18*)]. In contrast, the portion of the hydrophobic core farther away from the DNA binding surface contains less than 1 percent of all mutations

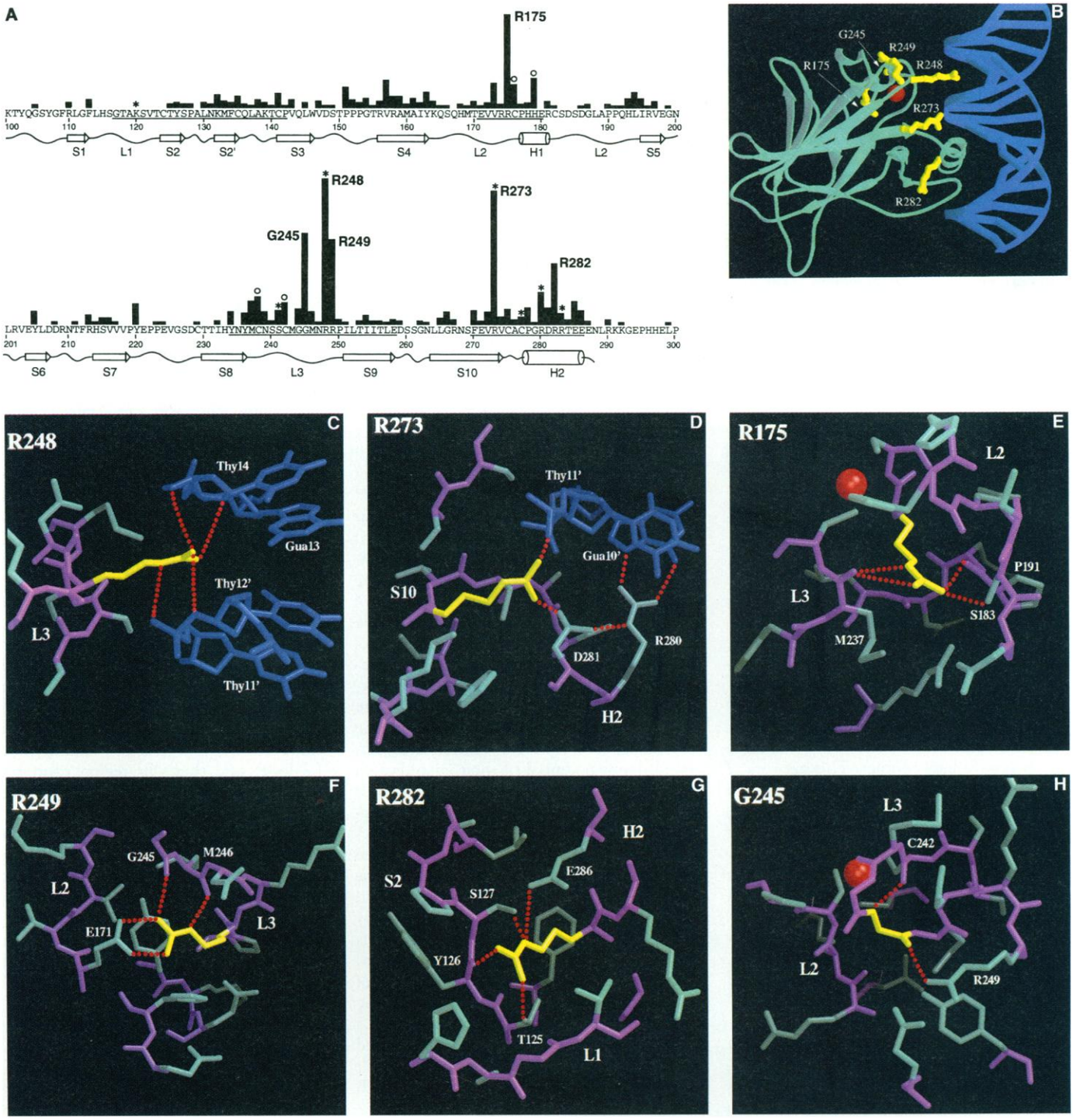

Fig. 6. The residues most frequently mutated in cancer are at or near the protein-DNA interface. (**A**) Sequence of the p53 core domain showing the conserved regions (underlined), and the secondary structure elements. The number of tumor-derived missense mutations at each residue are indicated by the bar graph and the six most frequently mutated residues are labeled (*18*). Residues involved in DNA binding are indicated by asterisks, and those involved in binding the zinc atom are indicated by circles. Single letter abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**B**) Ribbon drawing of the p53 core domain–DNA complex showing the six most frequently mutated residues of p53. The side chains of these residues are colored yellow, the core domain is light blue, and the DNA is dark blue. The zinc atom is shown as a red sphere. (**C** to **H**) Detailed views of the six most frequently mutated residues of p53 showing the protein and DNA residues that are within 7 Å of each mutation hotspot. The six most frequently mutated residues are colored yellow, the protein backbone atoms are purple, the amino acid side chains are light blue, and the DNA is dark blue. The zinc atom is shown as a red sphere. The hydrogen bond and electrostatic interactions that the mutation hotspot residues make with other side chains and the DNA are indicated by red-dotted lines. The residues involved in these interactions and the critical secondary structure elements are labeled.

[Fig. 6A (*18*)]. This portion of the hydrophobic core consists of Phe^{109}, Leu^{111}, Leu^{145}, Val^{218}, Thr^{230}, Ile^{232}, Ile^{255}, and Leu^{257}. One notable exception to this trend is a trio of residues, Pro^{151}, Val^{157}, and Tyr^{220}, which have mutation rates significantly higher than the rest of the β sandwich—4.5 percent for all three [Fig. 6A (*18*)]. It appears that they are important for the structure of the short loops at the end of the β sandwich opposite the DNA binding surface.

The structural basis for the distribution of p53 mutations observed in tumors has thus become evident. The majority occur at the L2 and L3 loops and the loop-sheet-helix motif because (i) structural shifts to accommodate mutations in this region will have more detrimental effects for DNA binding than those in the β sandwich, and (ii) this region and, in particular, the L2 and L3 loops have little regular secondary structure and rely extensively on side chain–side chain and side chain–backbone interactions for structural integrity.

Mechanisms of inactivation by mutations. The crystal structure of the core domain–DNA complex, in conjunction with biochemical studies of mutants, provides a solid basis for understanding how mutations affect the sequence-specific DNA binding activity of p53. One class of mutations involves residues that contact the DNA, and failure to bind DNA by these mutants can be attributed to loss of critical DNA contacts. Another class of mutations involves residues that appear to be important for the stable folding of the core domain, and loss of DNA binding by these mutants can be attributed to structural defects. Although the exact nature of the structural defects is not clear, biochemical studies of several of these mutants suggest that they have unfolded secondary and tertiary structure.

Arg^{175} provides one of the best examples of a structural p53 mutation because of its critical role in stabilizing the L2 and L3 loops. Several lines of evidence suggest that Arg^{175} mutants are unfolded. First, Arg^{175} mutants are found associated with the heat shock protein hsc70 (*25*), suggesting that they are at least partially denatured. Second, they bind the antibody PAb240 (*26*), which appears to be specific for denatured p53 (*27*), but not the antibody PAb1620, which recognizes the native state of p53 (*28*). Third, they are highly sensitive to proteolytic enzymes. Wild-type p53 yields a strikingly stable core domain after proteolytic digestion (*19, 20*), whereas the Arg^{175} mutants do not (*20*), suggesting that their core domain has unfolded regions accessible to proteases. Fourth, fusion proteins containing the full-length Arg^{175} mutant and the DNA binding domain of GAL4 cannot activate transcription from a GAL4 binding site (*29*), suggesting that the structural defects in these mutants are so extensive that they affect the transactivation domain at the NH_2-terminus.

Many other mutants have similar properties, and we now know that they predominantly involve structural residues. For example, Val^{143} mutants (in the β sandwich near the loop-sheet-helix motif) are found associated with hsc70 (*25*), and do not transactivate when fused to GAL4 (*29*); Arg^{249} mutants (in the L3 loop) are PAb240+/PAb1620− (*30*); and His^{179} mutants (one of the zinc ligands) do not transactivate when fused to GAL4 (*16*). Furthermore, metal chelating reagents result in wild-type p53 becoming PAb240+/PAb1620− (*31*), consistent with zinc playing a critical structural role.

In contrast, mutations at the DNA binding residue Arg^{273} (phosphate contact) do not significantly affect the structure of the protein. These mutants do not associate with hsc70 (*25*), they are PAb240−/PAb1620+ (*26*), they yield a stable core domain upon proteolytic digestion (*20*), and they transactivate when fused to the DNA binding domain of GAL4 (*29*). Mutants of other residues interacting with the DNA exhibit similar characteristics. Arg^{248} mutants (minor groove contact) transactivate when fused to GAL4 (*29*) and yield a stable core domain upon proteolytic digestion (*20*), and Arg^{280} mutants (major groove base contact) are PAb240−/PAb1620+ (*30*).

It has been suggested that mutant p53 molecules have a well-defined, common conformation, often referred to as the "mutant conformation," that is distinct from that of wild-type p53 (*32–34*). This hypothesis was based on the binding of monoclonal antibodies, primarily PAb240, which binds many of the mutants but not the wild-type protein (*26, 27*). In view of the crystal structure however, a reevaluation of these data suggests that this "mutant conformation" most likely represents denatured states of p53, rather than a well-defined, alternative conformational state of the protein.

First, we note that the peptide epitope for PAb240 [residues 212–217 (*26*)] is on the S7 β strand, part of which packs in the hydrophobic core of the β sandwich, and is thus mostly inaccessible to an antibody. This is consistent with PAb240 not binding the native state of wild-type p53 and further suggests that the mutants that bind PAb240 are at least locally unfolded in this region of the β sandwich. Second, the PAb240 epitope is far from the L3 loop and the loop-sheet-helix motif where many of the PAb240+ mutations occur. Thus, these mutations will have to produce unfolding that propagates a long distance on the structure and will most likely result in global denaturation. Third, the PAb240 antibody was produced from denatured p53 (*27*), and it can in fact bind to all types of p53 when they are denatured, including wild-type p53 on immunoblots (*27*). Thus, PAb240 binding most likely identifies p53 mutants that have unfolded regions rather than an alternate conformation.

In summary, some of mutations involve residues that bind DNA, and these mutations inactivate p53 by eliminating critical DNA contacts. Many of the remaining mutations involve residues that are important for the structural integrity of the core domain, and these mutations inactivate p53 most likely by unfolding of the structure. The extent of the unfolding will depend on the nature of the mutation, but biochemical data suggest that in many cases the unfolding is extensive and in fact may involve complete denaturation of the core domain.

Implications for understanding protein-DNA interactions. The p53 core domain does not contain any previously identified DNA binding motif. At a general level, the structure of the p53 core domain differs from other structurally characterized DNA binding domains in that it requires a comparatively large structural scaffold, the β sandwich, to correctly position and orient the structural elements that interact with the DNA.

At the level of protein-DNA interactions, however, it has several similarities to other protein-DNA complexes. First, p53 uses an α helix that fits in the major groove of the DNA to make contacts to the edges of the bases. This reemphasizes the central role that α helices play in DNA recognition. Previous structural studies of protein-DNA complexes have shown that α helices are used by most of the major families of DNA binding proteins, including the prokaryotic helix-turn-helix (*35, 36*), eukaryotic homeodomain (*37*), zinc finger (*38*), steroid receptor (*39*), and leucine zipper (*40*) motifs. Second, p53 uses a loop packing at the NH_2-terminal portion of the α helix to make additional contacts to the bases in the major groove. A similar helix-loop arrangement has been observed in the GATA-1–DNA complex (*41*).

The third similarity to protein-DNA interactions observed in other complexes involves the minor groove contact. In the p53 core domain–DNA complex, Arg^{248} binds in the minor groove adjacent to the major groove where the α helix binds. Two phosphate groups, which are in the section of the DNA backbone flanked by the α helix in the major groove on one side, and the arginine in the minor groove on the other side, make contacts to side chains and to a backbone amide group. The DNA

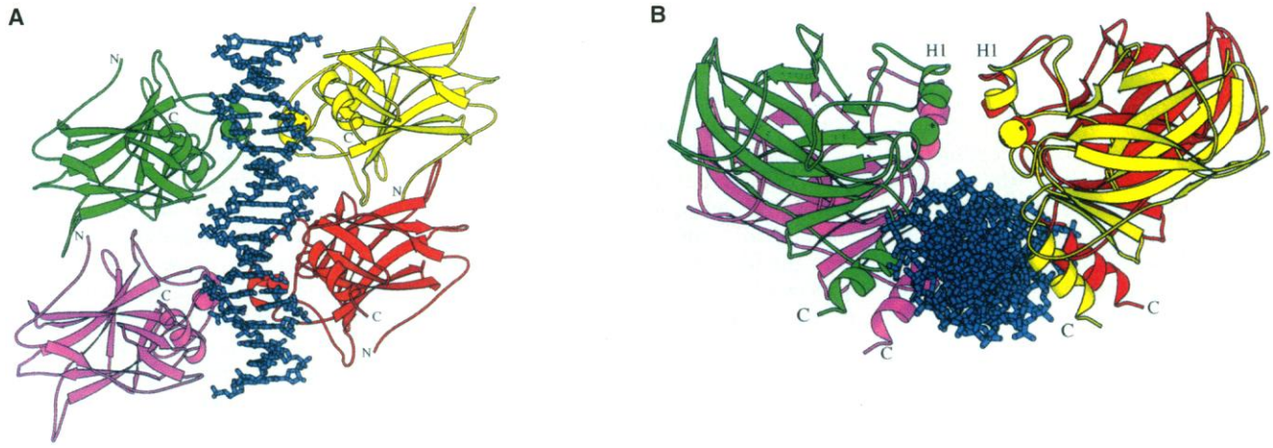

Fig. 7. Model of the core domain tetramer–DNA complex. The model was constructed by superimposing the core domain–DNA complex on a model of uniform B-DNA containing four pentamer consensus sequences (the superposition was done by aligning the phosphate atoms on the DNA). (**A**) Ribbon sketch of the tetramer model with the DNA in a vertical orientation. The four core domains are colored green, purple, yellow, and red, and the DNA is colored blue. The amino and carboxyl termini of the core domains are labeled. The head-to-tail dimer of the crystal structure corresponds to the green and purple core domains. (**B**) A view of the tetramer model with the DNA axis perpendicular to the plane of the figure. The H1 helices, which are in a position to make protein-protein contacts, are labeled.

sequence where the phosphate and arginine–minor groove contacts are made is A·T rich, and the minor groove is significantly narrower than in ideal B-DNA, resulting in a tight fit for the arginine.

Crystallographic studies of the phage 434 repressor (*36*), 434 cro (*42*), and hepatocyte nuclear factor-3 [HNF (*43*)]–DNA complexes have shown that they too use an arginine to interact with the minor groove of the DNA. Furthermore, in these complexes an α helix binds in the major groove adjacent to the arginine–minor groove contact. Two phosphate groups on the DNA backbone between the minor and major grooves make contacts to side chains and to backbone amide groups. The DNA sequences where the phosphate and arginine contacts are made are A·T rich, and the minor grooves are significantly narrower than in ideal B DNA. Thus, although the overall protein structures, arrangements of the α helices, and the precise DNA contacts differ in the four complexes, the similarities suggest that the phosphate contacts flanked by a helix in the major groove and an arginine in the adjacent minor groove, coupled to A·T–rich sequences, is an especially favorable arrangement.

Model of a p53 tetramer-DNA complex. p53 forms a tetramer via an oligomerization domain in the COOH-terminal portion of the protein [residues 312–365 (*19*)], and it binds DNA sites that typically contain four copies of a consensus pentamer sequence (*12*). These findings have suggested that p53 binds DNA as a tetramer with a stoichiometry of one p53 molecule per consensus pentamer. In the crystal structure, the core domain binds with an apparent stoichiometry of one molecule per pentamer consensus, but presumably due to the choice of binding site and to its lacking the oligomerization domain, it binds as a head-to-tail dimer. The dimer interface involves interactions between residues Ser^{96}, Ser^{99}, and Thr^{170} of the consensus complex, and Thr^{140}, Glu^{198}, Gly^{199}, and Glu^{224} of the nonconsensus complex. The DNA fragment used in the cocrystallization contains an additional consensus pentamer that could allow a symmetric dimer; however, we do not observe binding to this pentamer. This appears to be due to crystal packing forces because, in the crystal, the core domain that does not bind DNA excludes another core domain binding to this pentamer. [The core domain that does not bind DNA makes extensive protein-protein contacts that stabilize crystal packing.] To address the question of whether p53 can bind all four pentamers, we constructed a model of the tetramer-DNA complex based on the core domain–DNA complex and a DNA site containing four consensus pentamers (Fig. 7, A and B).

This model suggests that tetramer binding is possible because there are no steric clashes between the monomers (Fig. 7, A and B). Several additional features of the tetramer model are noteworthy. First, pairs of dimers can interact via their H1 helices, making protein-protein contacts that may contribute to tetramer binding (Fig. 7, A and B). Second, the four core domain molecules occur on the same face of the DNA, with their COOH-terminal H2 α helices extending through the major groove, and ending at the opposite face of the DNA (Fig. 7B). Third, the COOH-termini of the four core domains are within 35 to 40 Å of each other, a distance that can adequately be spanned by the roughly 25-residue flexible linkers that would connect the ends of the core domains to their oligomerization domains (*19*). Therefore, the crystal structure of the core domain dimer–DNA complex is consistent with biochemical data suggesting that intact p53 binds DNA as a tetramer.

The crystal structure of the core domain–DNA complex provides new insights into the DNA binding activity of p53 and its inactivation by tumorigenic mutations. The key finding of our study is that the majority of the tumorigenic mutations occur in the portion of the core domain structure involved in DNA binding. This finding solidifies the hypothesis that DNA binding and transactivation are the critical activities of p53 required for tumor suppression.

The crystal structure also provides a framework for understanding how mutations inactivate p53. A set of mutations appears to eliminate critical DNA contacts, while others appear to destabilize the two loops and the loop-sheet-helix motif involved in DNA binding. The structure thus suggests possible targets for the design of compounds to restore activity to mutant p53 proteins found in tumors.

REFERENCES AND NOTES

1. M. Hollstein, D. Sidransky, B. Vogelstein, C. C. Harris, *Science* **253**, 49 (1991); C. C. Harris, *ibid.* **262**, 1980 (1993); J. M. Nigro *et al.*, *Nature* **342**, 705 (1989); T. Takahashi *et al.*, *Science* **246**, 491 (1989).
2. J. D. Oliner, K. W. Kinzer, P. S. Meltzer, D. L. George, B. Vogelstein, *Nature* **358**, 80 (1992); M. Scheffner, B. A. Werness, J. M. Huibregtse, A. J. Levine, P. M. Howley, *Cell* **63**, 1129 (1990).
3. A. J. Levine, J. Momand, C. A. Finlay, *Nature* **351**, 453 (1991); B. Vogelstein and K. W. Kinzer, *Cell* **70**, 523 (1992).
4. C. A. Finlay, P. W. Hinds, A. J. Levine, *Cell* **57**, 1083 (1989); D. Eliyahu, D. Michalovitz, S. Eliyahu, O. Pinhasi-Kimhi, M. Oren, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8763 (1989); D. Michaloviitz,

O. Halevy, M. Oren, *Cell* **62**, 671 (1990).
5. S. J. Baker, S. Markowitz, E. R. Fearon, J. K. V. Willson, B. Vogelstein, *Science* **249**, 912 (1990).
6. P.-L. Chen, Y. Chen, R. Bookstein, W.-H. Lee, *ibid.* **250**, 1576 (1990).
7. L. A. Donehower *et al.*, *Nature* **356**, 215 (1992).
8. M. B. Kastan *et al.*, *Cell* **71**, 587 (1992).
9. Y. Yin, M. A. Tainsky, F. Z. Bischoff, L. C. Strong, G. M. Wahl, *Cell* **70**, 937 (1992); L. R. Livingstone *et al.*, *ibid.*, p. 923.
10. L. Hartwell, *ibid.* **71**, 543 (1992).
11. J. P. Bargonetti, N. Friedman, S. E. Kern, B. Vogelstein, C. Prives, *ibid.* **65**, 1083 (1991); S. E. Kern *et al.*, *Science* **252**, 1708 (1991); W. D. Funk, D. T. Pak, R. H. Karas, W. E. Wright, J. W. Shay, *Mol. Cell. Biol.* **12**, 2866 (1992).
12. W. S. El-Deiry, S. E. Kinzler, J. A. Pietenpol, K. W. Kinzler, B. Vogelstein, *Nat. Genet.* **1**, 45 (1992).
13. S. Fields and S. K. Jang, *Science* **249**, 1046 (1990).
14. L. Raycroft, H. Wu, G. Lozano, *ibid.*, p. 1049; G. Farmer *et al.*, *Nature* **358**, 83 (1992).
15. S. E. Kern *et al.*, *Science* **256**, 827 (1992).
16. T. Unger, M. M. Nau, S. Segal, J. D. Minna, *EMBO J.* **11**, 1383 (1992).
17. W. S. El-Deiry *et al.*, *Cell* **75**, 817 (1993); J. W. Harper, G. R. Adami, N. Wei, K. Keyomarsi, S. J. Elledge, *ibid.*, p. 805.
18. A. J. Levine *et al.*, *Br. J. Cancer* **69**, 409 (1994); the version of the database used in this study excludes mutations resulting in chain termination and contains 1060 mutations identified in human tumors; M. Hollstein *et al.*, *Nucleic Acids Res.*, in press.
19. N. P. Pavletich, K. A. Chambers, C. O. Pabo, *Genes Dev.* **7**, 2556 (1993).
20. J. Bargonetti, J. J. Manfredi, X. Chen, D. R. Marshak, C. Prives, *ibid.*, p. 2565.
21. Y. Wang *et al.*, *ibid.*, p. 2575; T. D. Halazonettis and A. N. Kandil, *EMBO J.* **12**, 5057 (1993).
22. T. A. Jones, in *Molecular Replacements*, E. J. Dodson, Ed. (SERC, Daresbury, United Kingdom, 1992), pp. 91–105.
23. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987); A. T. Brunger, *X-PLOR v2.1 Manual* (Yale Univ. Press, New Haven, CT, 1990).
24. D. E. Tonrud, L. F. Ten Eyck, B. W. Matthews, *Acta Crystallogr.* **A43**, 489 (1987).
25. P. W. Hinds *et al.*, *Cell Growth Diff.* **1**, 571 (1990).
26. C. W. Stephan and D. P. Lane, *J. Mol. Biol.* **225**, 577 (1992).
27. J. V. Gannon, R. Greaves, R. Iggo, D. P. Lane, *EMBO J.* **9**, 1595 (1990).
28. R. K. Ball, B. Siegl, S. Quellherst, G. Brander, D. G. Braun, *ibid.* **3**, 1485 (1984).
29. L. Raycroft, J. N. Schmidt, K. Yoas, M. Hao, G. Lozano, *Mol. Cell Biol.* **11**, 6067 (1991).
30. J. Bartek, R. Iggo, J. Gannon, D. P. Lane, *Oncogenes* **5**, 893 (1990).
31. P. Hainaut and J. Milner, *Cancer Res.* **53**, 1739 (1993).
32. J. Milner and E. A. Medcalf, *J. Mol. Biol.* **216**, 481 (1990).
33. ———, *Cell* **65**, 765 (1991).
34. T. D. Halazonetis, L. J. Davis, A. N. Kandil, *EMBO J.* **12**, 1021 (1993).
35. S. R. Jordan and C. O. Pabo, *Science* **242**, 893 (1988).
36. A. K. Aggarwal, D. W. Rodgers, M. Drottar, M. Ptashne, S. C. Harrison, *ibid.*, p. 899.
37. C. R. Kissinger, B. Liu, E. M. Blanco, T. B. Kornberg, C. O. Pabo, *Cell* **63**, 579 (1990).
38. N. P. Pavletich and C. O. Pabo, *Science* **252**, 809 (1991).
39. B. F. Luisi *et al.*, *Nature* **352**, 497 (1991).
40. T. E. Ellenberger, C. J. Brandl, K. Struhl, S. C. Harrison, *Cell* **71**, 1223 (1992).
41. J. G. Omichinski *et al.*, *Science* **261**, 438 (1993).
42. A. Mondragon and S. C. Harrison, *J. Mol. Biol.* **219**, 321 (1993).
43. K. L. Clark, E. D. Halay, E. Lai, S. K. Burley, *Nature* **264**, 412 (1993).
44. W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1984).
45. W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).
46. S.E.R.C. [U.K.] Collaborative Computing Project No. 4 (Daresbury Laboratory, Warrington, United Kingdom, 1979).
47. B. C. Wang, *Methods Enzymol.* **115**, 90 (1985).
48. T. A. Jones, J.-Y. Zou, S. W. Cowan, *Acta Crystallogr.* **A47**, 110 (1991).
49. J. S. Sack, *J. Mol. Graphics* **6**, 224 (1988).
50. Program DSCALEAD, M. R. Rould (1989).
51. Supported by NCI grant CA08748-29 and the Dewitt Wallace Foundation. We thank A. Levine for providing us with the database of p53 mutations; E. Kelly and S. Geromanos for synthesizing the DNA; and J. Hubbard for assistance with the color illustrations. Coordinates are being deposited with the Brookhaven Protein Data Bank.

9 May 1994; accepted 3 June 1994