The Three-Dimensional Structure of p53

Elena S. Stavridi, Yentram Huyen, Emily A. Sheston, and Thanos D. Halazonetis

Protein function is completely dependent on three-dimensional structure. Yet, molecular biologists often pay little attention to structural data. In a field, like p53, where numerous functions and protein-protein interactions have been proposed, the structural information can serve as a sieve to distinguish between credible models and models that are less likely to be physiologically relevant. Structural information can also help design meaningful experiments. In this chapter we will present the structural data that are available for p53 and consider their implications for p53 function. As will become evident, there are still major gaps in our knowledge of p53 structure.

2.1. p53 DOMAINS AND REGIONS

The human p53 protein is 393 amino acids long. Protein domains, defined as independently folding units of a protein, typically have a size of between 40 and 200 amino acids (Koonin et al., 2002). This suggests that p53 contains more than one protein domain, a prediction that has been confirmed by structural and functional studies (Vogelstein et al., 2000). Three domains are recognized in p53 (Fig. 2.1). A transactivation domain (residues 1–70), a sequence-specific DNA binding domain (residues 94–293) and a tetramerization domain (residues 324–355). These domains are flanked...
by linker regions. A proline-rich region (residues 71–93) links the transactivation and sequence-specific DNA binding domains; a second proline-rich region (residues 294–323) links the sequence-specific DNA binding and tetramerization domains; and a basic region (residues 356–393) forms the very C-terminus of the protein.

2.1.1. The Transactivation Domain

Functional studies indicate that the N-terminus of p53 (residues 1-70) can activate transcription either in the context of full-length p53 or when grafted to heterologous proteins (Fields and Jang, 1990; Raycroft et al., 1990; Lin et al., 1994; Candau et al., 1997). Accordingly, the N-terminus is referred to as the transactivation domain. Technically, the term domain to describe this region of p53 is inaccurate, since the N-terminus of p53 is not an independently folding unit. A better term would have been transactivating region; however, the term transactivating domain has been used so extensively that it is unlikely to be replaced.

The transactivation domain of p53 serves at least two roles. It activates transcription and it also regulates p53 function and stability. It does so by interacting with transcription factors, such as p300 and CBP, and also with Mdm2, a ubiquitin protein ligase that targets p53 for degradation in the proteasome (Avantaggiati et al., 1997; Gu et al., 1997; Lill et al., 1997; Scolnick et al., 1997; Lin et al., 1994; Honda et al., 1997). p300/CBP and Mdm2 have overlapping binding sites within the N-terminus of p53. In response to DNA damage phosphorylation of p53 N-terminal residues decreases the affinity of p53 for Mdm2 and, concomitantly, increases its affinity for p300/CBP (Shieh et al., 1997; Lambert et al., 1998; Chehab et al., 1999). This leads to increased p53 protein levels and increased p53 transcriptional activity.

A three-dimensional structure of the N-terminus of p53 bound to p300, CBP or any other transcription factor is not yet available. However, a three-dimensional structure of an N-terminal p53 peptide bound to the N-terminus of Mdm2 has been determined (Kussie et al., 1996) and explains how phosphorylation of residues Thr18 (see footnote for list of amino acid codes) and Ser20 of p53 decrease its affinity for Mdm2. The structure comprises residues 25–109 of human Mdm2 bound to a

---

Single and three-letter amino acid codes: A, Ala, alanine; C, Cys, cysteine; D, Asp, aspartic acid; E, Glu, glutamic acid; F, Phe, phenylalanine; G, Gly, glycine; H, His, histidine; I, Ile, isoleucine; K, Lys, lysine; L, Leu, leucine; M, Met, methionine; N, Asn, asparagine; P, Pro, proline; Q, Gln, glutamine; R, Arg, arginine; S, Ser, serine; T, Thr, threonine; V, Val, valine; W, Trp, tryptophan; Y, Tyr, tyrosine.
15-residue p53 peptide (Fig. 2.2). The region of Mdm2 that binds to p53 comprises an independently-folding domain consisting of a four helix bundle flanked on either side by β-sheets. The p53 peptide corresponds to amino acids 15–29 of the full-length protein. The N-terminal end of this peptide was not structured and was therefore invisible in the solved structure. The rest of the peptide corresponding to residues 17–29 of full-length p53 was structured and visible; most of this segment folds as an α-helix that interacts with a deep cleft on the Mdm2 surface. The interaction between the two proteins is mediated by hydrophobic interactions; residues Phe19, Trp23 and Leu26 of p53 interact with multiple conserved Mdm2 hydrophobic residues, including Leu54, Met62, Tyr67 and Val93, whose side chains line the cleft on the Mdm2 surface. Thus, the amphipathic nature of the p53 helix (hydrophobic on the side that interacts with Mdm2 and hydrophilic on the side exposed to solvent) is critical for the interaction.

The p53-Mdm2 structure explains how various phosphorylation events, induced in response to DNA damage, decrease the affinity between the two proteins (Chehab et al., 1999; Craig et al., 1999; Sakaguchi et al., 2000). The hydroxyl group of Ser20 is close to the side chain of Met62 of Mdm2; thus, Ser20 phosphorylation would lead to a steric clash, as well as position a negatively charged phosphate group next to the Met hydrophobic side chain. Phosphorylation of Thr18 is also predicted to weaken the interaction between p53 and Mdm2. The hydroxyl group of the Thr18 side chain forms a hydrogen bond with the carboxyl group of Asp21; this hydrogen bond stabilizes the amphipathic α-helix and would be disrupted by phosphorylation of Thr18. The structure further suggests that phosphorylation of Ser15 is unlikely to directly affect the interaction between p53 and Mdm2, because Ser15 is not involved in the p53-Mdm2 interface. However, Ser15 phosphorylation may have an effect in vivo, because it enhances the affinity of p53 for p300 and CBP, which compete with Mdm2 for binding to p53 (Shieh et al., 1997; Lambert et al., 1998).

The p53-Mdm2 structure further explains how amino acid substitutions that replace Leu22 and Trp23 of p53 with Gln and Ser, respectively, abolish the interaction...
between p53 and Mdm2 (Lin et al., 1994). In vivo these two substitutions make p53 refractory to Mdm2-dependent regulation (Chehab et al., 1999). The same substitutions also compromise the transcriptional activity of p53 (Lin et al., 1994) suggesting that the interaction of p53 with p300 and CBP may also involve an amphipathic helix from p53 binding to a hydrophobic cleft on the surface of these transcription factors.

The p53 amphipathic helix that interacts with Mdm2 (residues 18–26) spans only a small part of the p53 transactivation domain (residues 1–70). In some transactivation assays substitutions targeting Leu22 and Trp23 compromise, but do not abolish p53 transcriptional activity. Complete loss of transcriptional activity requires additional substitutions targeting residues Trp53 and Phe54 of p53 (Candau et al., 1997), suggesting that the p53 transactivation domain encompasses multiple elements with which it can interact with transcription factors and other proteins (such as Mdm2).

While the structure of the p53-Mdm2 complex displays a specific p53 conformation, in the absence of Mdm2 the p53 transactivation domain adopts mostly an unstructured conformation (Botuyan et al., 1997; Lee et al., 2000). The α–helical fold is stabilized when p53 binds to Mdm2, because Mdm2 provides a hydrophobic cleft that favors partitioning of the p53 hydrophobic and hydrophilic residues on distinct surfaces (the surfaces towards Mdm2 and solvent, respectively). The spacing of the hydrophobic residues on the p53 primary sequence is such that they partition on one surface when they adopt an α–helical fold. Ligand-receptor interactions, in which binding is associated with conformational changes, are referred to as induced-fit (Koshland, 1958). The conformational flexibility of the p53 transactivation domain may have been selected to allow p53 to interact with multiple proteins, since each p53-protein complex may stabilize a specific p53 conformation suitable for that specific interaction. Further, because of the conformational flexibility, the p53 transactivation domain residues can adapt to the active sites of various enzymes (kinases, acetylases, etc.), which explains why many post-translational modifications of p53 map to the transactivation domain (Vogelstein et al., 2000). As discussed above, these modifications can favor or disfavor specific p53-protein interactions and therefore regulate p53 function.

2.1.2. The DNA Binding Domain

2.1.2.1. Overall Structure and DNA Contacts

Probably the most interesting domain of the p53 tumor suppressor protein is its sequence-specific DNA binding domain, which encompasses amino acids 94–293 of the full-length protein and which is targeted by the vast majority of cancer-associated p53 mutations (Vogelstein et al., 2000). The structure of the p53 DNA binding domain has been solved in complex with DNA and also in the absence of DNA (Cho et al., 1994; Zhao et al., 2001).

The scaffold of the p53 DNA binding domain is a β–sandwich formed by two antiparallel β–sheets that pack against each other (Fig. 2.3). One end of the β–sandwich is formed by evolutionarily weakly conserved loops that join β–strands from the opposing β–sheets. In contrast, the other end of the β–sandwich is characterized by the presence of conserved secondary structure elements interspersed within the loops.
Figure 2.3. Three-dimensional structure of the DNA binding domain of human p53 in complex with DNA. Select secondary structure elements are marked (H1, H2, L1, L2, and L3). The side chains of residues that are frequently substituted in human cancer are shown. Of the residues corresponding to mutation hotspots, Arg175, Arg249, and Arg282 (red side chains) stabilize the native structure; whereas, Arg248 and Arg273 (purple side chains) contact DNA. The DNA backbone is shown as a ribbon. A, B, and C are orthogonal views.
that join the β–sheets. The first such element is a β–hairpin formed by two antiparallel β–strands (S2 and S2′). The second such element is a short α–helix (H1) that contains two Zn-chelating residues, which together with two other Zn-chelating residues contributed by loop L3, coordinate a Zn atom. The third element is a conserved α–helix (H2) that extends to the C-terminal end of the DNA binding domain. The presence of conserved secondary structure elements decorating one end of the β–sandwich clearly suggests that this is the functionally important part of the domain. Indeed, this is the side of the domain that contacts DNA, as revealed by the structure of the DNA binding domain in complex with DNA (Fig. 2.3).

The DNA sequence recognized by p53 consists of four tandem copies of the pentamer consensus sequence GGGCA arranged head-to-tail (El-Deiry et al., 1992). Two such pentamers comprising a p53 half-site are shown in Fig. 2.4A. To facilitate describing the structure of p53 bound to DNA, we assigned a number to the position of each nucleotide in the pentamer. The complementary nucleotides are indicated by the same number and an apostrophe (Fig. 2.4A). Each pentamer within the 20-nucleotide binding site is recognized predominantly by a single DNA binding domain (since p53 is a homotetramer). The contacts with the pentamer sequence are mediated by loops L1 and L3, strand S10, helix H2 and the loop that connects S10 to H2 (hereafter referred to as loop L4). Loops L1, L4, strand S10 and helix H2 contact the major groove of DNA (Fig. 2.4). From loop L1, the side chain of Lys120 makes a sequence-specific contact with guanine G4 and the amide nitrogen of the same residue contacts the phosphate backbone. From S10, the side chain of Arg273 contacts the phosphate backbone. From loop L4, the amide nitrogen of Ala276 contacts the phosphate backbone and the side chain of Cys277 makes a specific contact with cytosine C3′. From H2, the side chain of Arg280 makes a specific contact with guanine G2′, whereas the side chain of Arg283 contacts the phosphate backbone. Loop L3 interacts with the minor groove of DNA. The side chain of Ser241 contacts the phosphate backbone, whereas the side chain of Arg248 makes four contacts with DNA backbone sugar and phosphate atoms. The ability of Arg248 to make multiple contacts with the DNA backbone requires significant compression of the minor groove. In turn, this might explain the preference for adenine at position 1 of the consensus binding site, since adenine:thymidine pairs favor compression of the minor groove of double-stranded DNA (Yoon et al., 1988).

Taking into account the preference for adenine:thymidine bases at position 1 of the GGGCA pentamer to allow compression of the minor groove, it becomes evident that the structure explains binding specificity for positions 1, 2, 3 and 4 of the pentamer. This is in fact consistent with mutagenesis data suggesting that the nucleotide at position 5 contributes very little to the sequence-specificity of p53 DNA binding (El-Deiry et al., 1992; Halazonetis et al., 1993). Additional mutagenesis data further supports the p53–DNA structure solved by N. Pavletich (Cho et al., 1994). Extensive mutagenesis of the mouse p53 DNA binding domain identified the same residues as being important for DNA binding as the crystal structure (Halazonetis and Kandil, 1993). Further, mutagenesis of Lys120 changes the specificity of p53 for the nucleotide at position 4 of the pentamer (Freeman et al., 1994), whereas mutagenesis of Arg273 decreases the affinity of p53 for DNA, but does not change sequence-specificity (Wieczorek et al., 1996), consistent with the structure, which
Figure 2.4. Human p53–DNA contacts. A: Diagram of a p53 half-site consisting of two head-to-tail pentamer repeats. Two tandem half-sites constitute a complete p53 binding site. The nucleotides in the GGGCA pentamer are numbered 5, 4, 3, 2, and 1, respectively. The consensus pentamer has pyrimidine at position 1, cytosine at position 2, and purines at positions 3, 4, and 5. The lines indicate contacts of p53 residues with specific atom groups of the p53 DNA binding site. The green circles represent the bases, the yellow circles represent the sugars and the red circles the phosphate groups. Contacts mediated by amide nitrogens are indicated by the p53 residue number preceded by an N- prefix. B: Three-dimensional structure of the p53 DNA binding domain in complex with DNA highlighting the p53–DNA contacts. The orientation is similar to the one shown in Figure 2.3A. The p53 residues (side chains or amide nitrogen) are shown in the same color as the DNA atoms they contact. Select secondary structure elements are marked (L1, L2, L3, L4, and S10).

shows Lys120 in contact with the base at position 4 and Arg273 in contact with the DNA phosphate backbone. In fact, the Arg273 substitution can be rescued by substituting Thr284 in helix H2 with Arg. Thr284 does not contact DNA in the solved p53–DNA structure, but its substitution with Arg is predicted to allow a new contact to be established with the DNA backbone (Wieczorek et al., 1996).
2.1.2.2. Mapping Cancer-Associated Mutations on the DNA Binding Domain

Most of the tumor-associated \( p53 \) mutations target the sequence-specific DNA binding domain (Vogelstein et al., 2000). Certain residues within this domain are targeted much more frequently than others (Hollstein et al., 1991). The structure of the \( p53 \) DNA binding domain bound to DNA provides an explanation: the so-called “hotspot” residues contribute critically to DNA binding activity. Generally, \( p53 \) mutants have been divided into two classes (Cho et al., 1994). Class I mutants substitute residues that directly contact DNA, such as Arg248 and Arg273, whereas Class II mutants substitute residues, such as Arg175, Arg249 and Arg282, that stabilize the native structure of the \( p53 \) DNA binding domain (Fig. 2.3).

The nature of the amino acid substitutions explains the different properties of Class I and II mutants. Since Class I mutants target surface residues, the proteins encoded by these mutants are typically able to adopt the native fold. In contrast, Class II mutants, by definition, cannot adopt the native fold and are either partially (i.e. locally) or completely unfolded. Accordingly Class II mutants tend to be sequestered in the cytoplasm in complex with protein chaperones, such as the heat shock protein hsc70 (Finlay et al., 1988). Furthermore, Class II mutants react with antibodies, such as PAb240, that do not react with wild-type \( p53 \). These antibodies recognize epitopes that are buried in native \( p53 \), but become exposed in the unfolded \( p53 \) mutants (Gannon et al., 1990). The realization that Class II tumor-derived \( p53 \) mutants are actually unfolded proteins needs to be considered when interpreting their various activities. Unfolding exposes “sticky” hydrophobic residues through which Class II mutants can bind non-specifically to a wide array of proteins.

There are at least three reasons to explain why the tumor-associated mutations target predominantly the DNA binding domain of \( p53 \). The first reason is that mutations in the DNA binding domain generate dominant negative mutants. This is because \( p53 \) binds DNA as a homotetramer and tetramers containing both mutant and wild-type subunits are defective in DNA binding (Bargonetti et al., 1992; Halazonetis and Kandil, 1993). A second reason is the low melting temperature of the sequence-specific DNA binding domain of \( p53 \), which for human \( p53 \) is just a few degrees above 37°C (Bullock et al., 1997). The low melting temperature of the \( p53 \) DNA binding domain means that practically every non-conservative substitution in its core will decrease the melting temperature below 37°C preventing the protein from folding. The third reason is that many of the functionally important residues of this domain are arginines. Arginine codons are particularly susceptible to mutagenesis, as they contain CG dinucleotides, which, when damaged, are repaired with lower fidelity than other dinucleotides (Pfeifer et al., 2002).

2.1.3. The Oligomerization Domain

2.1.3.1. Overall Structure

The oligomerization domain of \( p53 \) resides within its C-terminus between residues 324 to 355. Its three-dimensional structure has been solved by X-ray
Figure 2.5. Three-dimensional structure of the human p53 oligomerization domain. The four subunits are labeled A, B, C, and D. The side chains of hydrophobic residues that mediate important inter-subunit interactions are shown. A, B, and C represent three different views.

crystallography and NMR spectroscopy (Lee et al., 1994; Jeffrey et al., 1995; Clore et al., 1995). The four monomer subunits adopt identical conformations and are related to each other by three perpendicular axes of symmetry (dihedral symmetry). Each monomer subunit of the p53 oligomerization domain consists of a β–strand (residues 326–333) and an α–helix (residues 335–354) that together form a V-like shape (Fig. 2.5). A conserved glycine residue (G334) allows a tight turn to be formed between these two secondary structure elements.

The interactions between the four subunits are extensive and are primarily hydrophobic in nature. The subunits are typically labeled A, B, C and D. In the dimer formed by subunits A and B (or the symmetrically equivalent subunits C and D) the β–strands pack antiparallel to each other and are stabilized by inter-subunit hydrophobic interactions between Phe328 and Phe338 and between the two Leu330 residues (Fig. 2.5). Because of the V-like shape of each subunit, the α–helices of subunits A and B also pack antiparallel to each other. Packing of the helices against the β–sheet is stabilized primarily by hydrophobic interactions mediated by Phe341.
In the dimer formed by subunits A and C (or the symmetrically equivalent subunits B and D) there are no interactions between the $\beta$–strands, but there are extensive interactions between the C-termini of the $\alpha$–helices mediated primarily by Leu348 and Leu350 (Fig. 2.5). Interestingly, in this part of the structure the helices pack parallel to each other. Thus, the $\alpha$–helices of the p53 oligomerization domain have the capacity to form both antiparallel (between subunits A and B) and parallel (between subunits A and C) interactions.

The central part of the $\alpha$–helices is involved in stabilizing the interaction between the A-B and C-D dimers. Residues Met340 and Leu344 from all four subunits interact with each other at the very center of the domain and are responsible for the tetramer stoichiometry (Fig. 2.5).

2.1.3.2. Mapping Cancer-Associated Mutations on the Oligomerization Domain

The frequency of mutations targeting the p53 oligomerization domain in human cancer is at least 100-fold lower than the frequency of mutations targeting the DNA binding domain (Hollstein et al., 1991). There are several explanations for this observation. First, the p53 oligomerization domain is thermodynamically very stable, since its melting temperature is about 70°C (Johnson et al., 1995). Given this high melting temperature, most single amino acid substitutions cannot lead to unfolding of the domain. In fact, the most common cancer-associated substitution in the p53 oligomerization domain targets Gly334; this residue is critical for the tight turn between the $\beta$–strand and $\alpha$–helix and its substitution with practically every other residue is sufficient to unfold the domain (Lee et al., 1994). The second reason for the low frequency of mutations targeting the oligomerization domain is that inactivating the oligomerization domain does not lead to dominant negative mutants (Shaulian et al., 1992). In contrast, p53 mutants with inactive DNA binding domains retain the ability to form heterooligomers with wild-type p53 and, consequently, have dominant negative activity.

2.1.3.3. Determinants of Oligomerization Stoichiometry

The structure of the p53 oligomerization domain allows us to visualize the residues that mediate inter-subunit interactions and therefore the residues that determine the stoichiometry of this domain as a tetramer. Because of the dihedral symmetry it is possible to disrupt interactions between subunits A and C (and the equivalent B and D) without affecting the interactions between subunits A and B (or the equivalent C and D). This has the effect of dissociating a p53 tetramer into A-B and C-D dimers. Amino acid substitutions that have this effect include Leu344 to Ala and a double substitution of Met340 to Gln and Leu344 to Arg (Waterman et al., 1995; Davison et al., 2001).

A more interesting way to change the stoichiometry of the p53 oligomerization domain from tetramer to dimer involves substitution of two hydrophobic residues, Phe341 and Leu344, at the core of the domain with other hydrophobic residues of different side chain size (McCoy et al., 1997). Whenever the side chain of the residue at position 344 becomes larger than the side chain of the residue at position 341, the stoichiometry of the domain switches from tetramer to dimer (Fig. 2.6). In this case
the change in stoichiometry is accompanied by a change in the fold of the domain; the dimer retains the antiparallel packing of the $\beta$–strands (compare Fig. 2.6 to Fig. 2.5B), but the $\alpha$–helices pack parallel to each other, as the overall shape of each subunit changes from V-like to L-like. Interestingly, the packing of the C-termini of the $\alpha$–helices in this mutant p53 dimer is practically identical to the way the C-termini of the $\alpha$–helices of subunits A and C pack in the wild-type structure (compare Fig. 2.6 to Fig. 2.5C). To our knowledge this is the only example of designed amino acid substitutions that change the fold of a protein domain. One implication of these findings is that the side chain size of hydrophobic residues is an important determinant of protein fold.

2.1.3.4. Determinants of Oligomerization Specificity

p53 belongs to a family of proteins with conserved sequence-specific DNA binding and oligomerization domains. While under certain conditions p53 can hetero-oligomerize with its family members, under physiological conditions the majority of p53 is thought to exist as homotetramers (Davison et al., 1999; Mateu et al., 1999). In turn, this implies oligomerization specificity. Specificity in the interaction between various subunits is generally thought to be mediated by charged and polar residues, which are capable of forming electrostatic interactions and hydrogen bonds. The p53 oligomerization domain has very few electrostatic interactions between the subunits. In fact, the oligomerization specificity of p53 can be changed by amino acid substitutions that target its hydrophobic residues. Substitutions that weaken existing strong hydrophobic interactions between the subunits coupled with substitutions that enhance existing weak hydrophobic interactions create variant p53 oligomerization domains that form homotetramers, but interact very weakly with the wild-type domain (Mateu et al., 1999; Stavridi et al., 1999). These results suggest that hydrophobic
interactions are critical determinants of both the fold and oligomerization specificity of the p53 oligomerization domain.

### 2.1.4. The N-terminal Proline-Rich Region

As shown in Fig. 2.1 the p53 domains discussed above are separated from each other by short linker regions. The most N-terminal of these regions, spans residues 71–93 and links the transactivation and DNA binding domains. This region has ten prolines and is often referred to as the polyproline region.

Functional experiments suggest that the polyproline region contributes to the functional activity of p53 and especially to its ability to induce apoptosis (Walker et al., 1996; Sakamuro et al., 1997; Ruaro et al., 1997; Venot et al., 1998). There are at least two ways to explain the functional significance of this region. Regions rich in prolines are often ligands for proteins containing SH3 domains (Yu et al., 1994). Thus the polyproline region may mediate an interaction between p53 and a protein containing an SH3 domain. While definitive evidence for a protein that binds to the p53 polyproline region is not yet available, 53BP2, an SH3 domain-containing protein, is known to bind p53 (Iwabuchi et al., 1994). The interaction between 53BP2 and p53 will be described further below in this Chapter. A second way by which the polyproline region could affect p53 function might be indirect through an effect on p53 folding. Proline-rich regions have conformationally-constrained backbones and may affect the kinetics of protein folding by acting as folding barriers separating two protein domains. As discussed above the DNA binding domain of p53 has a very low melting temperature; it is therefore possible that the polyproline region may be needed to facilitate its folding. Such a role has not yet been experimentally demonstrated for p53, but in other proteins polyproline regions have been shown to facilitate folding (Kusano et al., 2001; Wang et al., 2002).

### 2.1.5. The C-terminal Proline-Rich Region

The C-terminal proline-rich region spans residues 294–323 and links the DNA binding and tetramerization domains. This region is not described in the literature as being proline-rich, but has seven prolines. The primary function of this region is to accommodate the different symmetries of the p53 DNA binding and oligomerization domains in p53 tetramers bound to DNA (discussed below). However, as discussed above, a second function of the C-terminal proline-rich region might be to facilitate folding of the p53 DNA binding domain. As in the case of the N-terminal proline-rich region, such a function needs to be experimentally documented for p53.

### 2.1.6. The C-terminal Basic Region

The C-terminal basic region of p53 spans residues 356–393. As its name indicates this region is rich in basic amino acids. Like the N-terminal transactivation domain, this region is conformationally flexible (Ayed et al., 2001) and is subject to multiple post-translational modifications that regulate p53 function (Appella and Anderson,
Because of its basic nature, this region resembles histone tails. As such it appears to be modified by the same enzymes that modify histone tails and is subjected to similar post-translational modifications (phosphorylation and acetylation). The functional significance of these modifications is beyond the scope of this chapter. In vitro, they enhance the sequence-specific DNA binding activity of wild-type p53 (Appella and Anderson, 2000; Brooks and Gu, 2003; Xu, 2003) and a similar effect can be achieved by deleting this region or masking it with a monoclonal antibody (Hupp et al., 1992; Halazonetis et al., 1993). It is unclear, however, whether these modifications enhance the sequence-specific DNA binding activity of p53 in vivo. Instead, in vivo the C-terminal post-translational modifications may enhance the transcriptional activity of p53 by facilitating the recruitment of transcriptional coactivators to p53-target genes; these coactivators, many of which are histone acetyltransferases, can modify both p53 and histones leading to “opening up” of the chromatin structure, which is thought to be required for efficient transcription (Barlev et al., 2001; Espinosa and Emerson, 2001; Wang et al., 2001).

The structure of the p53 C-terminal basic region has been solved in complex with two proteins: Sir2, a protein deacetylase (Avalos et al., 2002), and S100B, a calcium-binding protein (Rustandi et al., 2000). Since the p53 C-terminal basic region on its own is conformationally flexible its interaction with Sir2 and S100B is governed by the “induced-fit” model. Thus, the p53 C-terminal basic region adopts conformations that fit the ligand-binding sites in Sir2 and S100B. As shown below these two conformations are completely different from each other.

Sir2 is a member of the evolutionarily highly conserved family of sirtuin NAD+-dependent deacetylases. This family of proteins is implicated in a variety of functions related to DNA, such as transcriptional silencing, DNA repair and chromosomal stability (Guarente, 2000; Denu, 2003). Members of this family catalyze the removal of acetyl groups from the ε-amino group of lysines in reactions that also yield nicotinamide and O-acetyl-ADP-ribose (Sauve et al., 2001) as products. Acetylated histones are physiological substrates of yeast Sir2 and their deacetylation leads to transcriptional silencing. As mentioned above, the C-terminal basic region of p53 is acetylated in vivo by the same enzymes that acetylate histone tails. This observation suggested that Sir2 proteins might also deacetylate p53, a prediction now confirmed by three laboratories (Luo et al., 2001; Vaziri et al., 2001; Langley et al., 2002).

The structure of a peptide corresponding to residues 372–389 of human p53 with an acetylated lysine at position 382 was solved in complex with an archaeal Sir2 protein (Avalos et al., 2002). In this complex, residues 379–387 of p53 were structured and the p53 peptide adopts the secondary structure of a β--strand, which together with two β--strands of Sir2 (strands 7 and 9) participates in a three-strand β–sheet (Fig. 2.7). Remarkably, other than for the side chain of acetylated Lys382, the contacts between Sir2 and the p53 peptide are not sequence-specific; rather the interaction is mediated by backbone atom hydrogen bonds of the kind present in all β–sheets. Thus, the structure suggests that Sir2 enzymes can deacetylate practically any acetylated peptide, a prediction confirmed by enzymatic assays involving several Sir2 proteins, including human SIRT2 (Avalos et al., 2002). One would have to assume
that substrate specificity for Sir2 enzymes \textit{in vivo} will be mediated by interactions between the substrate and enzyme that are further away from the enzyme active site; such interactions are not evident in structures containing short peptides as substrates. Thus, a structure of a longer p53 polypeptide bound to a human Sir2 protein is needed to understand how p53 is targeted by human Sir2 \textit{in vivo}.

S100B belongs to the so-called EF-hand family of calcium binding proteins. This family includes well-known proteins such as calmodulin and troponin and its members change conformation in response to changes in the intracellular concentration of calcium ions (Yap et al., 1999; Lewit-Bentley and Rety, 2000; Donato, 2001). Many EF-hand proteins, including S100B, are involved in various cytoskeletal functions, such as contractile activity of muscle cells and in the dynamics of actin filaments. The three-dimensional structure of S100B has been solved in the unbound and calcium-bound forms and shows the protein as a homodimer with each subunit consisting of four \( \alpha \)-helices. Upon calcium binding helix 3 changes orientation exposing a patch of hydrophobic residues that serves as binding site for various proteins (Drohat et al., 1996; Kilby et al., 1996; Matsumura et al., 1998). S100B binds to the actin capping protein CapZ and the structure of calcium-bound S100B has been solved in complex with a 12-residue peptide corresponding to residues 265–276 of CapZ (Inman et al., 2002). In this structure the CapZ peptide adopts an \( \alpha \)-helical conformation and its interaction with S100B involves several CapZ hydrophobic residues, including Ile269, Trp271, Ile274 and Leu275 (Fig. 2.8A). Trp271 is located in a deep hydrophobic cleft of S100B and is conserved in the majority of S100B binding peptides.

S100B was discovered to bind p53 as part of studies examining the phosphorylation of the p53 C-terminal basic region by the calcium and phospholipid-dependent protein kinase PKC. As a result of these studies binding between p53 and S100B was demonstrated \textit{in vitro} and was shown to involve the C-terminal basic region of p53 (Baudier et al., 1992). The structure of S100B bound to a p53 peptide containing residues 367-388 of human p53 was subsequently solved by NMR spectroscopy (Rustandi et al., 2000). In this structure the p53 peptide adopts an \( \alpha \)-helical conformation.
Figure 2.8. Three-dimensional structure of S100B in complex with a CapZ (A) or a human p53 (B) peptide. S100B is a symmetric dimer, but the view shown is not along an axis of symmetry, so the two S100B subunits and their bound peptides are viewed from different angles.

(Fig. 2.8B). However, the orientation of the α-helix is different from the orientation of the CapZ helix in the S100B-CapZ structure (compare Fig. 2.8A to Fig. 2.8B). Interestingly, only a single p53 hydrophobic residue, Leu383, is present in the S100B hydrophobic cleft and p53 lacks the tryptophan residue that is present in the consensus sequence of S100B-binding peptides. Thus, this structure does not provide strong support for the hypothesis that S100B is a physiologically important regulator of p53 function in vivo. Further studies are needed to address this question; such studies can now be designed with help from the S100B-p53 structural data.

2.2. MODELS FOR THE STRUCTURE OF FULL-LENGTH p53 HOMOTETRAMERS

As described above the three-dimensional information regarding p53 derives from analysis of isolated domains and not from analysis of full-length protein. In principle, it should be straightforward to model the structure of a full-length p53 tetramer bound to DNA using the available structures of the isolated domains. However, more careful analysis reveals that a model of a p53 tetramer cannot be easily derived from the available structures. There are two problems that need to be considered. The first problem relates to the discrepancy between the symmetry of the various p53 subunits relative to each other in a p53 homotetramer and the symmetry of the p53 DNA binding sites in DNA. This discrepancy, which is explained below, suggests that p53 undergoes global changes in its conformation as it binds DNA. The second problem relates to whether four p53 sequence-specific DNA binding domains can bind DNA without steric clashes, when the crystal structure of the p53 monomeric DNA binding domain bound to DNA is used to model the structure of a p53 tetramer bound to DNA. As described below, computer modeling reveals steric clashes implying that binding of p53 tetramers to DNA is associated with local conformational changes either in the DNA or in the p53 DNA binding domains or in both.
2.2.1. Global (Symmetry-Related) p53 Conformational Changes Associated with DNA Binding

The need for global p53 conformational changes associated with DNA binding arises because the p53 homotetramer and its DNA binding site have different symmetries. The symmetry of p53 is dictated by the symmetry of its oligomerization domain. As mentioned above this domain has dihedral symmetry, which means that the four subunits are related to each other by three axes of symmetry that are perpendicular to each other (Fig. 2.5). Assuming that all four p53 subunits adopt the same conformation (the one that is energetically most favored), then the entire p53 tetramer, including the DNA binding domains, will adopt dihedral symmetry (Fig. 2.9A).

The p53 DNA binding site consists of four tandem pentamers arranged head-to-tail. The symmetry relating these pentamers can be described as cyclic-translation: within each half-site the repeats are related by cyclic symmetry via a single rotation axis and the two half-sites are related to each other by translation (Fig. 2.9B).

Because p53 and its DNA binding site have different symmetries, DNA binding must be associated with a conformational switch that allows the DNA binding domains to adopt the cyclic-translation symmetry of the DNA binding site. The proline-rich
region that links the DNA binding domain and oligomerization domains is critical for resolving the symmetry discrepancy (Waterman et al., 1995). Thanks to its length and conformational flexibility, this region can adopt distinct conformations in each subunit, such that the DNA binding domains exhibit cyclic-translation symmetry, while the tetramerization domain retains dihedral symmetry (Fig. 2.9C). In support of this model, deletions within the C-terminal proline-rich region prevent p53 from binding to full-length DNA sites, although binding to DNA half-sites is still possible (Waterman et al., 1995; Fig. 2.9D).

The extent to which the DNA binding domains of full-length tetrameric p53 adopt a stable dihedrally-symmetric state in the absence of DNA (Fig. 2.9A) is not known. Based on in vitro observations that deletion of the C-terminal basic region of p53 enhances DNA binding, it was proposed that this region participates in inter-subunit interactions that stabilize the dihedrally-symmetric state of p53 (Hupp et al., 1995; Waterman et al., 1995). To address this model, the conformations of p53 molecules that have or do not have the C-terminal basic region were compared by NMR spectroscopy (Ayed et al., 2001). One of the p53 fragments that was examined included the sequence-specific DNA binding domain, the C-terminal proline-rich region and the oligomerization domain, while the other included in addition the C-terminal basic region. Both p53 fragments had amino acid substitutions within the oligomerization domain that disrupted the interactions between the A-B and C-D dimers; thus these p53 fragments assembled as dimers. Dimeric stoichiometry was necessary, because the sensitivity of NMR spectroscopy experiments decreases as protein size increases. The conformations of these two p53 fragments were compared, but not determined. This is possible, because the magnetic resonance frequencies of atoms are dependent on their local environment. Thus, similar magnetic resonance frequencies indicate similar conformations. The analysis revealed that the two p53 fragments adopted similar conformations, suggesting that the C-terminal basic region does not affect the conformation or symmetry of p53. However, these experiments do not completely rule out the possibility that the C-terminal basic region stabilizes a dihedrally-symmetric state of p53 for the following reasons. First, the C-terminal basic region of p53 may stabilize the dihedrally-symmetric state by interacting with the N-terminus, which was not present in the p53 constructs examined by NMR spectroscopy. Second, the interactions that stabilize the dihedrally symmetric state may occur only in the context of p53 tetramers and not in the context of p53 dimers. Third, p53 dimers do not need to resolve differences in p53 and DNA symmetries to bind DNA half-sites, since p53 dimers and DNA half-sites both have cyclic symmetry (Fig. 2.9E). Thus, the use of dimeric p53 proteins, while necessary for technical reasons, may not have been optimal to study the effects of the C-terminal basic region on p53 conformation.

2.2.2. Local (Sequence-Specific DNA Binding Domain and/or DNA) Conformational Changes Associated with p53 DNA Binding

The structure of the human p53 DNA binding domain was solved from crystals that contained in the asymmetric unit three DNA binding domains and one oligonucleotide containing two tandem pentamers (Cho et al., 1994). Only one of the three
DNA binding domains was bound to DNA in a sequence-specific manner, yet the structures of all three domains were virtually identical suggesting that binding of isolated DNA binding domains to DNA is not associated with any significant conformational change in the p53 protein structure (Cho et al., 1994). This conclusion is further supported by the structure of the mouse p53 DNA binding domain solved in the absence of DNA (Zhao et al., 2001); the mouse p53 structure is essentially identical to that of the human p53 DNA binding domain bound to DNA.

The absence of significant conformational changes associated with binding of isolated human p53 DNA binding domains to DNA does not preclude the possibility of conformational changes in the p53 DNA binding domain, when full-length p53 binds DNA as a homotetramer. In fact, attempts to model four p53 sequence-specific DNA binding domains bound to linear B-form DNA using the available crystallographic data reveals steric clashes between the four p53 sequence-specific DNA binding domains (Nagaich et al., 1997). One way to resolve the steric clashes is by bending the p53 DNA site and experimental data indeed suggests that p53 bends its target DNA upon binding (Balagurumoorthy et al., 1995; Cherny et al., 1999; Nagaich et al., 1999). However, it is also possible that the steric clashes are resolved by changes in the conformation of the p53 DNA binding domains themselves in a way that changes slightly their position relative to DNA. It is hard to predict whether the steric clashes are resolved by DNA bending, changes in the conformation of the p53 DNA binding domain or both. Interestingly, NMR analysis of isolated human p53 DNA binding domains examined in the absence and presence of specific p53 DNA sites reveals that DNA binding is accompanied by changes in p53 magnetic resonance frequencies of residues that are far away from the bound DNA (Klein et al., 2001; Rippin et al., 2002). These frequency shifts might be indicative of conformational changes in the p53 DNA binding domain, although other interpretations cannot be excluded. The definitive way to understand how p53 homotetramers resolve the steric clash problem would be to solve the structure of a p53 homotetramer bound to DNA. This has proven to be technically very difficult.

2.3. STRUCTURES OF p53 WITH 53BP1 AND 53BP2

53BP1 and 53BP2 (p53 binding proteins 1 and 2, respectively) were identified by a yeast two-hybrid screen for proteins that bind to p53 (Iwabuchi et al., 1994). Both 53BP1 and 53BP2 bind to overlapping surfaces of the p53 sequence-specific DNA binding domain, yet utilize different structural motifs to do so. Functional data link p53 to 53BP1 and 53BP2, but it still remains to be proven whether the functional interactions are due to direct protein-protein interactions.

2.3.1. The p53-53BP1 Interaction

53BP1 localizes rapidly to sites of DNA double-strand breaks after exposure of cells to ionizing radiation, where it is thought to activate ATM (Schultz et al., 2000; Mochan et al., 2003). Thus, 53BP1 functions in the same DNA damage
The Three-Dimensional Structure of p53 checkpoint pathway as p53. Orthologs of 53BP1 are present in all eukaryotes (Weinert and Hartwell, 1988; Willson et al., 1997). The common structural feature among these orthologs are highly conserved C-terminal BRCT repeats. Deletion mapping analysis indicates that the interaction between p53 and 53BP1 is mediated by the sequence-specific DNA binding domain of p53 and the BRCT repeats of 53BP1 (Iwabuchi et al., 1994).

BRCT repeats are protein-protein interaction motifs. They are present in proteins that function in the cellular response to DNA damage (Bork et al., 1997). 53BP1 and its orthologs have two tandem BRCT repeats connected by a short linker region. The structure of the two tandem BRCT repeats of 53BP1 in complex with the sequence-specific DNA binding domain of 53 has been solved by X-ray crystallography (Joo et al., 2002; Derbyshire et al., 2002).

In the three-dimensional structure, the two BRCT repeats and the linker region between the repeats pack tightly against each other to form a single globular domain (Fig. 2.10A). Each BRCT repeat consists of a β-sheet surrounded by α-helices on

Figure 2.10. Three-dimensional structure of the p53 DNA binding domain in complex with the BRCT repeats of 53BP1. A: View of the protein domains with the p53 DNA binding domain shown in the same orientation as in Figure 2.3A. The evolutionarily conserved cleft at the interface of the two BRCT repeats of 53BP1 maps to the helices colored in red. B: p53–53BP1 interface showing the side chains of residues involved in the protein–protein interaction. The orientation is not the same as that shown in panel A.
either side. The interface between the two repeats is formed by conserved hydrophobic and charged interactions that allow helix 2 from the first repeat to form a three-helix bundle with helices 1 and 3 from the second repeat. This leads to a cleft on the surface of 53BP1 at the interface of the two BRCT repeats. The surface of this cleft is highly conserved in evolution. In BRCA1, another DNA damage checkpoint protein with two C-terminal tandem BRCT repeats, this cleft is targeted by most cancer-associated mutations (Williams et al., 2001; Joo et al., 2002).

In the p53-53BP1 structure the DNA binding domain of p53 does not contact the evolutionarily conserved cleft at the interface of the two BRCT repeats (Fig. 2.10A). Instead, the p53 DNA binding domain contacts a surface of 53BP1 formed by the N-terminal BRCT repeat and the linker region between the two BRCT repeats. From the p53 side, the interaction with 53BP1 is mediated primarily by the L3 loop and to a lesser extent by the L2 loop and H1 helix. Specifically, from loop L3 of p53, Met243 contacts Val1829 and Tyr1846 of 53BP1; Asn247 contacts the 53BP1 main chain; Arg248 contacts Asp1861; and Arg249 contacts Asn1845 (Fig. 2.10B). From loop L2 of p53, Gln167 contacts Gln1863 and from helix H1, Arg181 contacts Asp1833 of 53BP1 (Fig. 2.10B). The p53 residues that are key to its interaction with 53BP1, Arg248 and Arg249, are highly conserved in evolution; however, the evolutionary conservation may reflect the importance of these residues for sequence-specific DNA binding (Cho et al., 1994). Of the 53BP1 residues that are key to the interaction with p53 several are conserved in Xenopus and C. elegans. Nevertheless, as mentioned above, the p53 binding surface of 53BP1 is less well-conserved in evolution than the cleft at the interface between the two BRCT repeats.

The functional significance of the p53-53BP1 physical interaction is as yet unclear. Comparison of the surfaces of p53 involved in DNA and 53BP1 binding indicates significant overlap such that it would be impossible for a p53 sequence-specific DNA binding domain to contact simultaneously both DNA and 53BP1. This could imply that p53 interacts with 53BP1 transiently prior to or during the process of its activation by DNA damage. A transient interaction between p53 and 53BP1 might explain why an interaction between endogenous p53 and 53BP1 proteins has not yet been reported in human cells. Furthermore, a transient interaction might allow 53BP1 to recruit p53 to activated ATM and Chk2, where p53 is phosphorylated; activated p53 could then be released from 53BP1 and relocalize to the promoters of its target genes. The availability of the 53BP1-p53 structure will allow this putative model to be tested by facilitating the construction of p53 mutants that lose the capacity to interact with 53BP1, while retaining DNA binding activity. These p53 mutants can then be tested for activation in response to irradiation.

2.3.2. The p53-53BP2 Interaction

53BP2 was the second protein identified in the original yeast two-hybrid screen for proteins that interact with p53 (Iwabuchi et al., 1994). 53BP2 turns out to be the C-terminal fragment of a full-length protein termed ASPP2. In turn, ASPP2 belongs to a family of proteins that in humans includes 3 members, ASPP1, ASPP2 and
iASPP (Samuels-Lev et al., 2001; Bergamaschi et al., 2003). All three proteins are characterized by the presence of 4 tandem ankyrin repeats followed by an SH3 domain in their C-terminus. In addition, they all bind the DNA binding domain of human p53, but have different activities; ASPP1 and ASPP2 promote the apoptotic activity of p53, whereas iASPP inhibits p53 activity. In C. elegans there is only one member of this family that most closely resembles human iASPP. Suppression of C. elegans iASPP leads to apoptosis in germ cells, which can be rescued by suppression of p53. Thus, in C. elegans, which lack Mdm2, the activity of p53 is curtailed by iASPP (Bergamaschi et al., 2003).

The structure of the C-terminus of ASPP2 (53BP2) has been solved in complex with the sequence-specific DNA binding domain of p53 (Gorina and Pavletich, 1996). The 53BP2 C-terminus contains four ankyrin repeats (Michaely and Bennett, 1992) and an SH3 domain (Cohen et al., 1995) (Fig. 2.11A). Each ankyrin repeat consists of a β–hairpin followed by two antiparallel β–helices arranged perpendicular to the

![Figure 2.11](image)

**Figure 2.11.** Three-dimensional structure of the p53 DNA binding domain in complex with the ankyrin and SH3 domains of 53BP2 (ASPP2). **A:** View of the protein domains with the p53 DNA binding domain shown in the same orientation as in Figure 2.3B. **B:** Interface of p53 with 53BP2 showing the side chains of residues involved in the protein–protein interaction.
The four ankyrin repeats are tightly packed against each other, such that the \( \beta \)–hairpins form a continuous \( \beta \)–sheet and the \( \beta \)–helices form helix bundles. The C-terminal ankyrin repeat packs against the SH3 domain, which adopts the same \( \beta \)–strand-rich fold as previously described SH3 domains in other proteins.

The interaction between 53BP2 and p53 involves loops L2 and L3 and helix H1 of p53. In that regard, the interaction surface in p53 overlaps pretty well with the 53BP1 interaction surface and partially with the sequence-specific DNA binding surface. From the 53BP2 side, the interaction involves the SH3 domain, which interacts with loop L3 of p53, and the fourth ankyrin repeat, which interacts with loop L2 and helix H1 of p53.

The 53BP2 SH3 domain contains the characteristic peptide-binding groove present in SH3 domains (Fig. 2.11B). This groove is occupied by the L3 loop of p53, which makes electrostatic, hydrogen bond and hydrophobic interactions with 53BP2. Specifically, Arg248 of human p53 interacts with Asp475 and Glu495 of 53BP2; Asn247 of p53 interacts with Tyr469 of 53BP2 and Met243 of p53 interacts with Leu514 of 53BP2. The latter contact appears to be important for the specificity of the p53/53BP2 interaction. Most SH3 domains have a tyrosine at the position corresponding to Leu514 of ASPP1, which would not be able, due to its bulky side chain, to accommodate the Met243 side chain of human p53.

The fourth ankyrin repeat of 53BP2, via its \( \beta \)–hairpin, interacts with the L2 loop and helix H1 of p53 (Fig. 2.11B). This interaction involves hydrogen bonds mediated by residues Ser183 and His178 of p53 with backbone amides of 53BP2. The interaction is facilitated by the presence of a tyrosine insertion in the \( \beta \)–hairpin of 53BP2; this tyrosine (Tyr424) allows steric complementarity between p53 and 53BP2 and most likely contributes to the specificity of the p53/53BP2 interaction.

As with the interaction of p53 with 53BP1, the genetic evidence in human cells and in C. elegans suggests that ASPP proteins and p53 functionally interact (Samuels-Lev et al., 2001; Bergamaschi et al., 2003). However, the significance of the direct protein-protein interaction between p53 and ASPP2 (53BP2) remains to be established. Analysis of p53 mutants that fail to bind 53BP2, but retain DNA and 53BP1-binding activities, will help address this question.

### 2.4. STRUCTURE OF THE p73 C-TERMINAL SAM DOMAIN

As discussed elsewhere in this volume, p53 is a member of a protein family that in humans includes two other members, p63 and p73 (Yang and McKeon, 2000). The similarity between p53, p63 and p73 at the amino acid sequence level is high and all three proteins have similar domain organizations with a DNA binding domain at the center of the protein and an oligomerization domain at the C-terminus. One difference, however, is the presence of a C-terminal extension in certain splice variants of p63 and p73 that contains a Sterile Alpha Motif (SAM) domain (Thanos and Bowie, 1999).

SAM domains are protein-protein interaction motifs, primarily found in signaling molecules and transcription factors involved in developmental regulation (Schultz
The structure of the SAM domain of human p73 was determined by NMR spectroscopy (Chi et al., 1999) and consists of 5 α–helices that form a tight globular structure (Fig. 2.12). The overall structure is very similar to that reported for the SAM domain of the ephrin receptor tyrosine kinase with the only difference being that whereas the ephrin receptor SAM domain homo-oligomerizes, the p73 SAM domain is a monomer (Smalla et al., 1999; Stapleton et al., 1999; Thanos et al., 1999). It is very likely that the SAM domains of p63 and p73 mediate protein-protein interactions with, as yet, unknown proteins. The identification of these proteins will greatly facilitate the understanding of the function of the p63 and p73 SAM domains.

### 2.5. CONCLUSIONS AND FUTURE DIRECTIONS

The p53 tumor suppressor is one of the best studied human proteins. As described in this chapter, significant progress has been made towards understanding its three-dimensional structure. The structural information has helped our understanding of p53 function and regulation and has also explained how tumor-associated mutations inactivate p53. The challenge now is to extend the structural studies to visualize multidomain fragments of p53 and even full-length p53. Structures of p53 with important partner molecules, such as transcriptional coactivators, are also needed. In the end the structural information is likely to form the basis for pharmacologic rescue of p53 function in human cancer.
ACKNOWLEDGMENTS

The authors thank Nikola Pavletich, Phil Jeffrey, Cheryl Arrowsmith, Stanley Opella, Mark McCoy, Rodney Harrington and Victor Zhurkin for helpful discussions.

REFERENCES

The Three-Dimensional Structure of p53


The Three-Dimensional Structure of p53


The p53 Tumor Suppressor Pathway and Cancer
Zambetti, G.P. (Ed.)
2005, X, 246 p., Hardcover