

## In vitro evolution of thermostable p53 variants

ICHIRO MATSUMURA AND ANDREW D. ELLINGTON

University of Texas at Austin, Austin, Texas 78712

(RECEIVED July 15, 1998; ACCEPTED December 9, 1998)

### Abstract

The tumor suppressor p53 is conformationally unstable at physiological temperature. Even the activated p53 $\Delta$ 30 variant, which lacks the self-inhibiting carboxy terminal domain, has a half-life of only 8 min at 37 °C in vitro. We have developed a genetic approach to identify p53 variants that stabilize the active conformation. The human p53 $\Delta$ 30 gene was randomly mutated, and the resulting library was expressed in *Escherichia coli* under conditions that apparently denatured the parental protein. Stable p53 variants were identified based on their ability to specifically bind a p53 consensus site. The initial thermostable variants were randomly recombined by DNA shuffling, and substitutions that were functionally additive or synergistic were identified in a second more stringent round of screening. The DNA binding activity of N239Y/N268D/E336V p53 $\Delta$ 30 variant has a half-life of 100 min at 37 °C, 12 times longer than that of the parental protein. The thermostable variants should be more amenable to crystallographic studies and more effective in gene therapies than the wild-type protein.

**Keywords:** directed evolution; DNA shuffling; in vitro evolution; p53; thermostability; tumor suppressor

The level of p53 activity within a cell is an important determinant of its tumorigenic fate. Over half of all tumors contain mutated p53 genes, making it by far the most frequently mutated gene found in human cancers (Hollstein et al., 1991). Transgenic mice homozygous for inactive p53 alleles develop tumors much more frequently than their wild-type counterparts (Donehower et al., 1992). Expression of p53 in cultured cells can lead to cell cycle arrest, apoptosis or senescence, all pathways that could potentially prevent tumor development (Levine, 1997). These findings have led to p53 being touted as a potential gene therapy for the treatment of a variety of cancers (Roth et al., 1996).

The structure and function of the p53 gene product are not completely understood. The p53 protein is a 393 amino acid transcription factor that activates or represses the transcription of many downstream target genes (Levine, 1997). p53 binds double-stranded DNA with two adjacent copies of the consensus sequence: 5'-PuPuPuC(A/T)(T/A)GPyPyPy-3' (El-Deiry et al., 1992), and its transcription activation activity in cultured cells is dependent on DNA binding (Kern et al., 1992). The protein contains an acidic activation domain (amino acids 1–43), a sequence specific DNA binding domain (100–300), and a carboxy terminal domain (300–

393). The latter domain includes a linker (300–320), a tetramerization domain (320–360), and a regulatory domain at the extreme C-terminus (363–393) (Ko & Prives, 1996).

The full-length unmodified protein does not bind its DNA targets in vitro until the carboxy terminal 30 amino acids are bound by an antibody, deleted, or phosphorylated (Hupp et al., 1992; Hupp & Lane, 1994). The sequence specific DNA binding activity is also activated in vitro when this regulatory domain is bound by small nucleic acids (Ko & Prives, 1996) and heat shock proteins (Hansen et al., 1996). The activation of p53 function in cultured mammalian cells has been correlated with phosphorylation (Ko & Prives, 1996), acetylation (Sakaguchi et al., 1998), glycosylation (Shaw et al., 1996), and proteolytic removal (Okorokov et al., 1997) of the carboxy terminal domain. These modifications are thought to activate p53 by causing it to undergo a conformational change (Hupp et al., 1992).

Perhaps because of the variety of regulatory roles and functions performed by p53, the protein is conformationally unstable (Milner, 1995). For example, the full-length protein can no longer be activated in vitro by antibody binding to its carboxy terminal regulatory domain after a 10 min incubation at 37 °C. Similarly, p53 $\Delta$ 30, the activated form of p53 lacking the carboxy terminal 30 amino acids (Hupp et al., 1992), loses DNA binding activity in vitro within 40 min at 37 °C (Hansen et al., 1996). The conformational instability of p53 complicates efforts to manipulate the protein for experimental or therapeutic purposes. Therefore, we developed a high-throughput genetic screen for p53 variants. A library of randomly mutated p53 $\Delta$ 30 genes was expressed in *Escherichia coli*, and those variants that retained DNA binding activity after extended incubations at 37 °C were isolated. The most thermostable variant, N239Y/N268D/E336V, was 12 times more resistant to inactivation at 37 °C than the parental protein.

Reprint requests to: Andrew D. Ellington, ICMB A4800/MBB 3.424, 26th and Speedway, University of Texas at Austin, Austin, Texas 78712; e-mail: andy.ellington@mail.utexas.edu.

**Abbreviations:** amp, ampicillin; BSA, bovine serum albumin; chl, chloramphenicol; EDTA, ethylenediaminetetraacetate; -his, construct containing a carboxy terminal histidine tag; IPTG, isopropyl beta-D-thiogalactopyranoside; LB, Luria Broth; PCR, polymerase chain reaction; p53 $\Delta$ 30, p53 gene or protein lacking the nucleotides encoding the carboxy terminal 30 amino acids or its gene product; p53CON, consensus p53 binding site; Pu, purine; Py, pyrimidine.

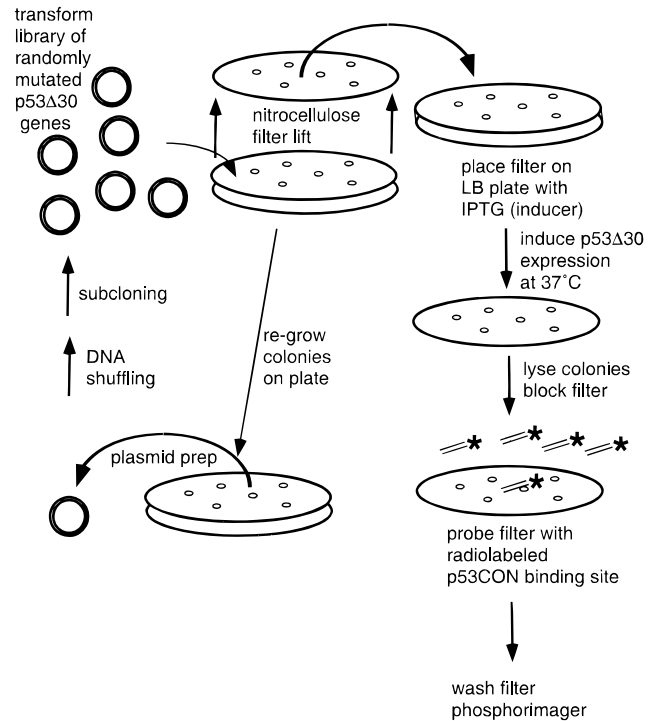
## Results

### Rationale for and development of a high-throughput screen for thermostability

The objective of this study was to identify sequence substitutions that stabilize the active conformation of p53. The likely existence of stabilizing substitutions in p53 had been predicted by Hansen et al. (1996) on both experimental and theoretical grounds, and several different strategies could potentially be pursued. Partial knowledge of the structure of p53 (Cho et al., 1994) has allowed rational alterations in p53 sequence and function. For example, Wiczorek et al. (1996) increased the affinity of p53 for its DNA ligand by introducing an additional basic residue in the sequence-specific DNA binding domain. Unfortunately, the rational design approach is limited because the structure of the full-length protein has not been solved, and because the relationships between sequence, structure, and function are often speculative. Random mutagenesis has often proven more effective than rational design in stabilizing protein structures (Alber & Wozniak, 1985; Zhao & Arnold, 1997), but individual amino acid substitutions in isolation have generally conferred only modest increases in thermostability (Zhao & Arnold, 1997). For these reasons, we have attempted to develop a sensitive and high-throughput screen for thermostable p53 function that can be iteratively applied. This screen should promote the accumulation of numerous stabilizing sequence variations that collectively will confer greatly enhanced thermostability.

Since p53 must bind DNA to function as a transcription factor (Kern et al., 1992), we developed an assay based on a screen originally developed to clone DNA binding proteins (Singh et al., 1989). In short, the *p53Δ30* gene was randomly mutated and subcloned into an *E. coli* expression vector (Fig. 1). The screen for p53 function was developed in *E. coli* for three reasons: (1) the use of bacterial rather than yeast or mammalian cells allows larger numbers of variants to be transformed and rapidly screened for activity; (2) *E. coli* does not post-translationally modify p53 (Hupp & Lane, 1994), so stabilizing mutations are likely to affect the protein structure in general, rather than the protein structure in the context of a particular set of modifications; (3) the colony-lift method permits the minute adjustment of reaction conditions such as temperature, pH, ionic strength, redox potential, and ligand concentrations. Previously, the *p53* gene was overexpressed in yeast in tandem with a p53-activated reporter gene (Scharer & Iggo, 1992). The fine-tuning of the activation and expression conditions necessary to identify thermostable variants would have been extremely difficult in the yeast system. Moreover, the colony-lift method is potentially more quantitative than those relying on reporter genes, since signal detection with a Phosphorimager (Molecular Dynamics, Sunnyvale, California) is linear over four orders of magnitude.

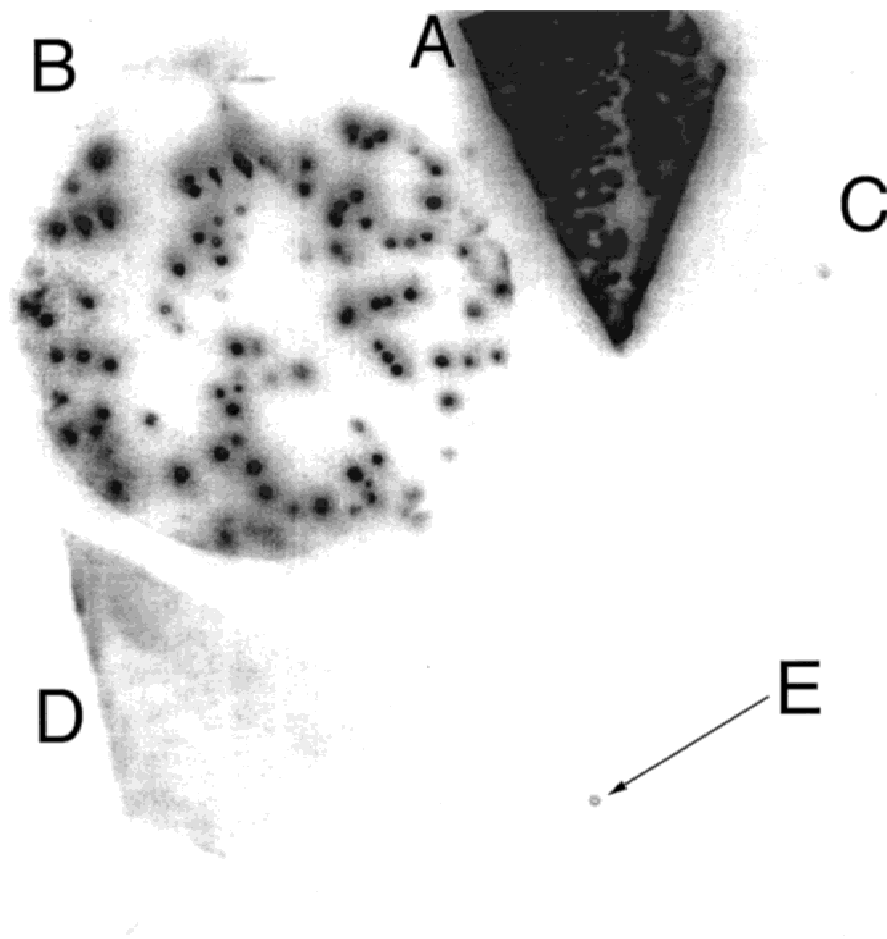
We chose to use mutagenic polymerase chain reaction (PCR), which is less strongly biased toward transitions than standard PCR conditions (Cadwell & Joyce, 1992), to randomly introduce sequence substitutions into *p53*. For example, Leung et al. (1989) found that 95% of the mutations in standard PCR were transitions. In contrast, the variants that we eventually isolated following mutagenic PCR and screening contained 37 transitions and 22 transversions. Bacteria were transformed with the library of randomly mutated *p53Δ30* genes and propagated under noninducing conditions. Colonies were adsorbed to a nitrocellulose filter and transferred to an agarose plate containing an inducer. The *p53Δ30*-expressing colonies were then lysed and probed with a radio-



**Fig. 1.** Screen for thermostable p53 DNA binding function. (From top left) A library of randomly mutated *p53Δ30* genes is subcloned into an expression vector and transformed into *E. coli*. The resulting colonies are transferred to a nitrocellulose filter, which is overlaid upon an agarose plate containing an inducer. The genes are induced under conditions in which the parental p53Δ30 protein is inactive and presumably denatured. The colonies are then lysed and probed with a radiolabeled (double-stranded) oligonucleotide encoding the consensus p53 binding site, p53CON (Funk et al., 1992). Colonies that retain DNA binding activity are isolated from the original plate and randomly recombined by DNA shuffling for a second round of screening.

labeled double-stranded oligonucleotide encoding the consensus p53 binding site, p53CON (Funk et al., 1992). Those colony remnants that bound the most label as measured by a Phosphorimager were picked from the original (noninducing) plate. Thermostable variants from the first round were randomly recombined and further mutated by DNA shuffling (Stemmer, 1994), and the chimeric variants were screened in a second round for further improvements in thermostability.

The screen developed by Singh et al. (1989) proved to be adaptable to screening p53 libraries for thermostable variants. In a control experiment, colonies were transformed with vectors that expressed the *p53Δ30* gene. When these colonies were grown in inducing conditions for 4 h at 23 °C, they could readily bind the p53CON probe (Fig. 2A). Sixteen percent (97/610) of the randomly mutated p53Δ30 proteins exhibited detectable DNA binding activity under these nonselective conditions (Fig. 2B), suggesting that most of the mutations are neutral or deleterious. In contrast, colonies without the expression vector did not bind the probe (Fig. 2C). Finally, when expression of the *p53Δ30* gene was induced for 2 h at 23 °C, followed by 2 h at 37 °C, the colonies bound an undetectable quantity of probe (Fig. 2D). These results strongly suggested that DNA binding activity in the filter-bound extracts could be attributed solely to the p53Δ30 protein, and that the conformation of this protein was not stable at 37 °C, as previously shown by Hansen et al. (1996).



**Fig. 2.** Sample screen for thermostable p53 $\Delta$ 30 variants. (A) Colonies carrying the parental p53 $\Delta$ 30 gene, (B) a library of randomly mutated p53 $\Delta$ 30 genes, or (C) no plasmid were induced for 4 h at 23 °C, then lysed and probed with radiolabeled p53CON oligonucleotide as described in the text. Colonies (D) transformed with the parental gene or (E) the mutant library were induced for 2 h at 23 °C, then 2 h more at 37 °C, then similarly lysed and probed. One colony on filter E expresses an apparently thermostable p53 $\Delta$ 30 mutant (this colony was eventually confirmed to contain a thermostable variant, and is clone 10 in Table 1).

#### Screen for thermostable variants

Ten thousand transformants were plated and induced under the selective regime described above (2 h at 23 °C followed by 2 h at 37 °C). Colonies that retained DNA binding activity were picked, re-plated, induced under selective conditions, and probed. Thirteen of the 28 colonies originally picked exhibited reproducibly greater DNA binding activity than the parental strain. These variants were sequenced and are presented in Table 1 (clones 4–16). It should be noted that these sequence substitutions do not cluster in three-dimensional space, and would have been difficult or impossible to identify through rational protein engineering (see Discussion).

The 13 p53 $\Delta$ 30 genes encoding thermostable variants were then randomly recombined by DNA shuffling (Stemmer, 1994). Two thousand recombinants were screened for DNA binding activity after a 12 h incubation at 37 °C. Three variants demonstrated reproducibly greater activity than the most thermostable variant from the first round of screening (Table 1). All of these variants (clones 1–3) contained the sequence substitutions N239Y and N268D, while two of the three variants (clones 1–2) also contained the sequence substitution E336V. The best explanation of the pheno-

types observed in our screen was that the sequence substitutions conferred thermostability to the known DNA binding function of p53. While the screen was theoretically sensitive to alterations in the level of expression of p53, the evolved N239Y/N268D/E336V and parental proteins were expressed at similar levels when the bacteria are propagated and induced in liquid culture at 23 °C (Fig. 3).

#### *In vitro* characterization of a thermostable p53 variant

Given the exceptional convergence of sequence substitutions observed in the second round of screening, the N239Y/N268D/E336V p53 $\Delta$ 30 variant (clone 2) was characterized in greater detail. The triple substitution and parental p53 $\Delta$ 30 genes were subcloned and expressed as fusion proteins with a carboxy terminal hexahistidine tag. The terminal histidines did not affect the DNA binding activity of p53 $\Delta$ 30 in the colony-lift assay (Fig. 4). The fusion proteins were purified by metal affinity chromatography and assayed for thermostability.

The observed improvements in DNA binding activities are specific for the p53 consensus site. In addition to p53CON

**Table 1.** Sequences of thermostable p53 mutants

Clone	Round <sup>a</sup>	Inferred amino acid changes <sup>b</sup>
1	2	M133V/ <b>N239Y</b> / <b>N268D</b> / <b>E336V</b> / <b>L(A)364P</b> <sup>c</sup>
2	2	<b>N239Y</b> / <b>N268D</b> / <b>E336V</b>
3	2	<b>E62V</b> / <b>N239Y</b> / <b>N268D</b>
4	1	<b>N268D</b> / <b>E336V</b> / <b>L(A)364P</b>
5	1	S116T/S166P/ <b>N268D</b> /A270
6	1	A88S/V157I/ <b>L344V</b>
7	1	D42G/ <b>N268S</b>
8	1	<b>N239Y</b>
9	1	E51K/E326G
10	1	<b>D207E</b>
11	1	<b>D207E</b> /F212S/ <b>L(A)364H</b>
12	1	P80L/ <b>V203A</b>
13	1	N30A/E56K/S106N/H115R/ <b>V203A</b>
14	1	C277S/ <b>L344M</b>
15	1	L45Q/T102A/P191S/ <b>D207E</b> /I332T
16	1	V31A/D49G/S183T/L264I/E346V

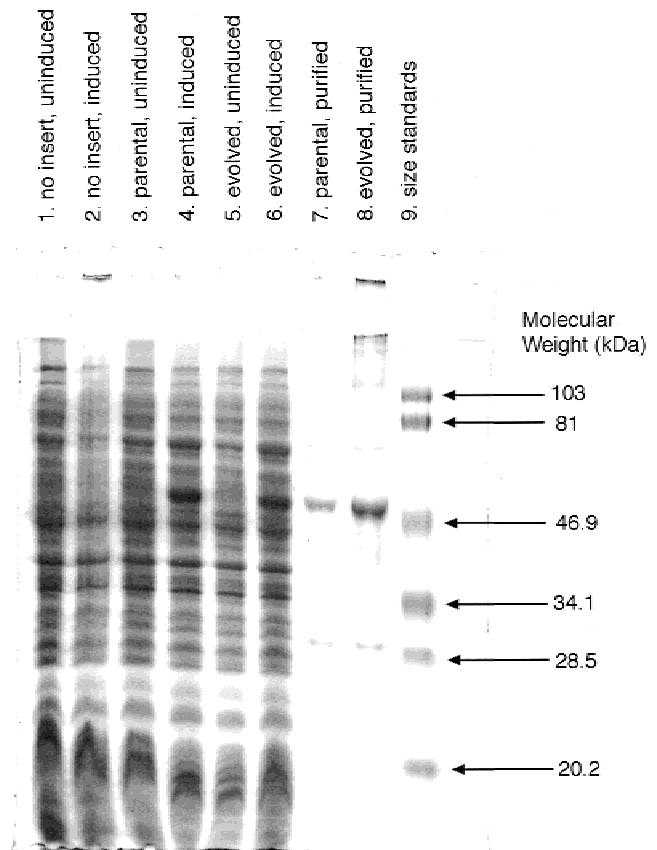
<sup>a</sup>The first round library was generated by a mutagenic polymerase chain reaction. The second was made by randomly recombining the first round clones by DNA shuffling.

<sup>b</sup>Repeatedly mutated loci are printed in bold letters.

<sup>c</sup>Amino acids 364 and 365 are alanine and histidine, respectively, in the full-length wild-type p53. The parental p53Δ30 protein used in these studies has 363 wild-type amino acids, plus a carboxy terminal leucine and glutamate.

(5'-GGACATGCCCGGGCATGTCC-3'), both the parental and evolved proteins bind other sequences that fit the p53 consensus: the p53 binding site upstream of the *p21* gene (5'-GAACATGTC CCAACATGTTG-3'; 85% identity with p53CON), and a half-site (5'-AACGTTGGACATGCCCGGGGC-3'; 45% identity with p53CON (data not shown)). In contrast, neither the parental nor evolved p53Δ30 proteins show detectable affinity for a noncognate, double-stranded control oligonucleotide (5'-GGCACGTGA GCCTGCAGGGG-3'). This is not surprising, given that nonspecific DNA binding activity has been shown to reside in the carboxy terminal tail of p53 (Reed et al., 1995), which was removed in the p53Δ30 construct. Furthermore, the screen was carried out in the presence of a large excess (5 μg/mL) of unlabeled salmon sperm DNA, which would be expected to compete with any nonspecific interactions between the p53Δ30-his protein and the oligonucleotide probe.

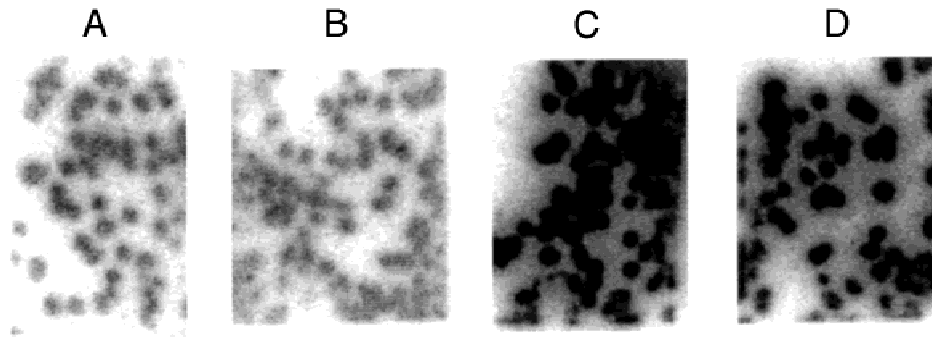
While the variants identified in our screen were more thermostable than wild-type p53, they nonetheless still lost activity after extended periods at 37 °C (Fig. 5), indicating that physiological temperature was likely still well above the apparent  $T_m$  of the protein. We therefore attempted to follow the kinetics of inactivation. The parental and triple substitution variant proteins were incubated at 37 °C for varying lengths of time and DNA binding activity was assessed in a band-shift assay (Fig. 6). The wild-type protein:DNA complex migrated primarily as two bands. These results are consistent with previous band-shift analyses of p53Δ30, in which complexes with a similar consensus binding site are frequently distributed over a range of mobility (see Hansen et al., 1996; Anderson et al., 1997). The N239Y/N268D/E336V p53Δ30-his protein:DNA complex did not readily enter the gel (see also Fig. 8), indicating that novel quaternary structural interactions may have formed.



**Fig. 3.** Expression of p53Δ30 protein by *E. coli* in liquid culture. *E. coli* carrying expression vectors with no insert (lanes 1–2), the parental p53Δ30-his (“his” indicates a carboxy terminal hexahistidine tag; lanes 3–4) or the evolved N239Y/N268D/E336V p53Δ30-his (clone 2; lanes 5–6) were propagated to mid-log phase at 37 °C. The cultures were incubated for 4 h more at 23 °C, with (lanes 2, 4, and 6) and without (lanes 1, 3, and 5) 0.5 mM of the inducer, IPTG. Whole cell extracts from 250 μL of the indicated cultures were run on SDS-polyacrylamide gels, along with the parental (lane 7) and evolved (lane 8) proteins purified from 2 mL of the appropriate induced cultures (see Materials and methods). The parental and evolved proteins are expressed at the similar levels in the induced cells (lanes 4 and 6), but the purification yield of the evolved variant is higher (compare lanes 7 and 8).

Labeled DNA molecules with altered mobility were assumed to be complexed with p53 protein. To determine the rate of decay, the amounts of DNA:protein complexes were quantitated as a function of time using a Phosphorimager. These kinetic data were then fit to a model of exponential decay (Fig. 7). The purified N239Y/N268D/E336V p53Δ30-his variant was inactivated much more slowly than the parental; the calculated half-lives of the triple substitution variant and the parental proteins at 37 °C are  $100 \pm 20$  and  $8.3 \pm 0.4$  min, respectively. The inactivation kinetics of the purified parental p53Δ30-his protein were similar to that of the impure protein adsorbed to a nitrocellulose filter, validating our initial assessments of p53 DNA binding function in bacterial lysates (compare Figs. 5, 6).

To the extent that the triple substitution variant was stable for longer at 37 °C, it might be expected that it should also be stable at higher temperatures than the parental protein. To determine whether this was the case, the triple substitution variant and wild-type proteins were incubated at increasing temperatures for 10 min and their residual DNA binding activities assayed (Fig. 8). The tri-



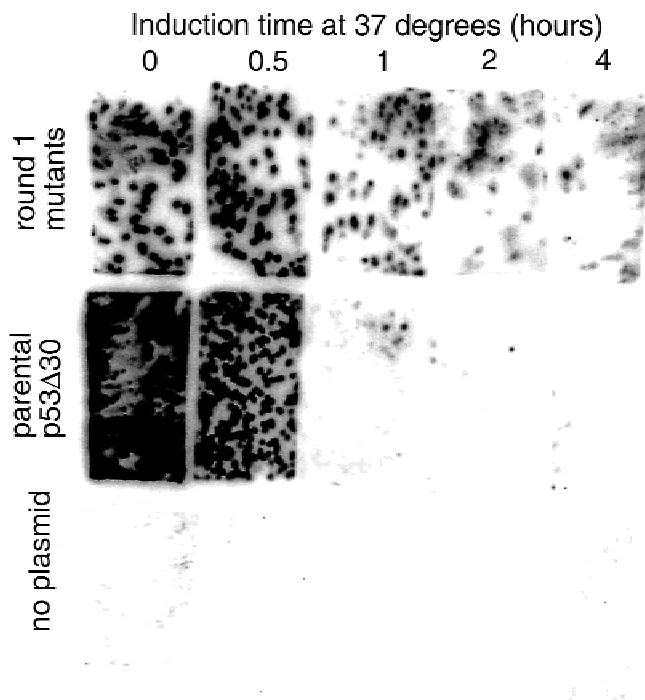
**Fig. 4.** DNA binding activity of p53 $\Delta$ 30 variants. (A) Colonies expressing parental p53 $\Delta$ 30, (B) p53 $\Delta$ 30-his, (C) evolved p53 $\Delta$ 30, or (D) evolved p53 $\Delta$ 30-his, proteins were induced for 4 h at 23 °C, lysed, and probed with radiolabeled p53CON oligonucleotide. The addition of the hexahistidine tag does not affect the DNA binding activity of either the parental or evolved forms.

ple substitution variant retained activity at significantly higher temperatures than the parental protein, indicating again that the amino acid substitutions generally stabilize the folded protein structure.

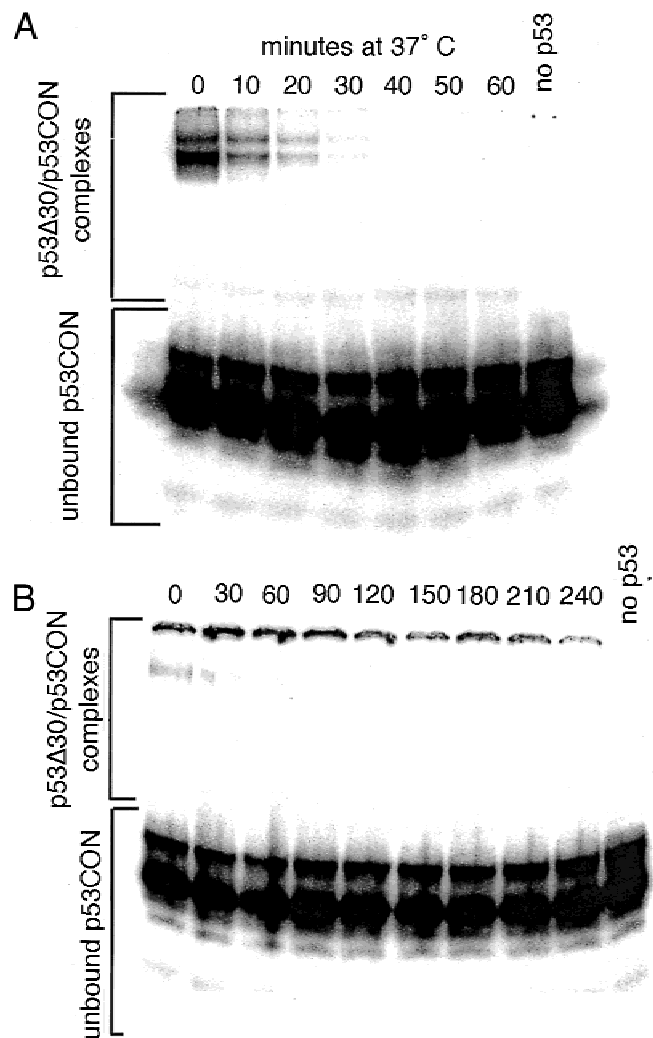
## Discussion

### Potential mechanisms of stabilization

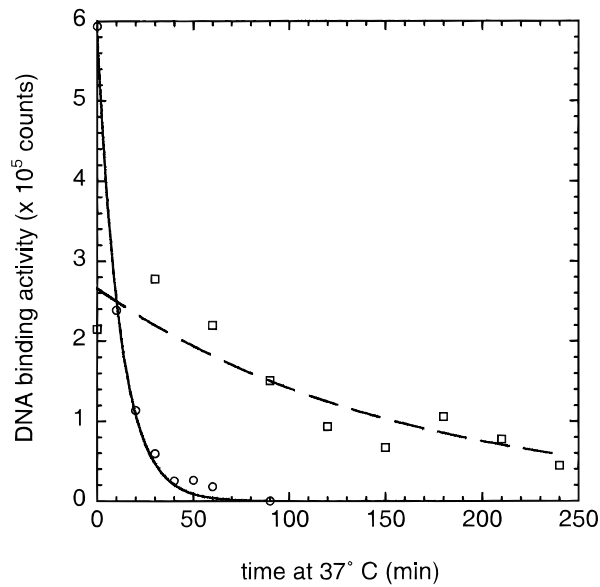
Two of the variants isolated, D207E (clone 10) and N239Y (clone 8), contained only single amino acid changes. These substitutions



**Fig. 5.** Confirmation of thermostability of p53 variants. (Top) Colonies carrying a pool of round 1 mutants, (middle) the parental p53 $\Delta$ 30 allele, or (bottom) no p53 were induced at 23 °C for varying lengths of time and then shifted to 37 °C for a total of 4 h of induction. For example, the samples on the left labeled “0” were at 23 °C for 4 h, while the samples on the right labeled “4” were at 37 °C for 4 h. The samples in the middle labeled “1” were at 23 °C for 3 h and 37 °C for 1 h. The colonies were lysed and probed as described in Figure 2 and in Materials and methods.

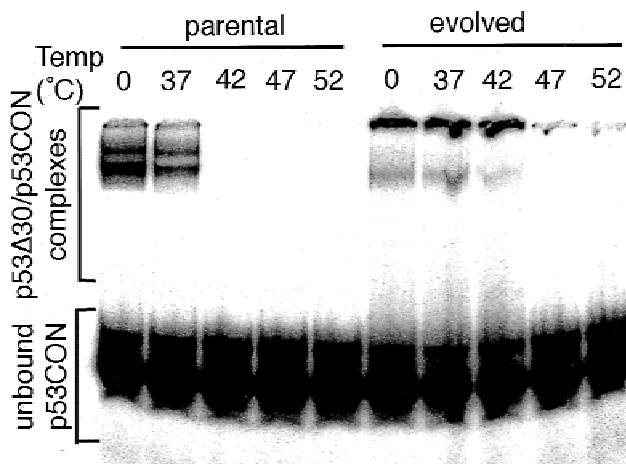


**Fig. 6.** Inactivation kinetics of a thermostable p53 variant. (A) The parental p53 $\Delta$ 30 and (B) N239Y/N268D/E336V p53 $\Delta$ 30 variant (clone 2) were expressed as histidine-tagged fusion proteins and purified (see Fig. 3). Twenty nanograms of each protein in 10  $\mu$ L binding buffer were pre-incubated at 37 °C for the times indicated and then mixed with 4 ng of radiolabeled DNA encoding the p53CON binding site. After a 30 min binding reaction, the samples were electrophoresed on 4% nondenaturing acrylamide gels and analyzed using a Phosphorimager. The positions of the shifted bands are indicated.



**Fig. 7.** Calculation of half-lives. The inactivation kinetic data for the p53 $\Delta$ 30-his (circles) and N239Y/N268D/E336V p53 $\Delta$ 30-his thermostable variant (squares) proteins from Figure 6 are plotted. The data were fitted to an exponential decay model as described in Materials and methods. The minimizing parameters for the parental protein (solid line) are  $y_0 = 5.9 \times 10^5$  counts and  $k_1 = 0.084 \text{ min}^{-1}$ , and  $y_0 = 2.7 \times 10^5$  counts and  $k_1 = 0.0063 \text{ min}^{-1}$  for the thermostable variant (dashed line). Given the derived  $k_1$  values, half-times are the value for  $t$  when  $y = \{y_0/2\}$ . The calculated half-life of the parental protein is  $8.3 \pm 0.4 \text{ min}$  while the half-life of N239Y/N268D/E336V p53 $\Delta$ 30-his is  $100 \pm 20 \text{ min}$ .

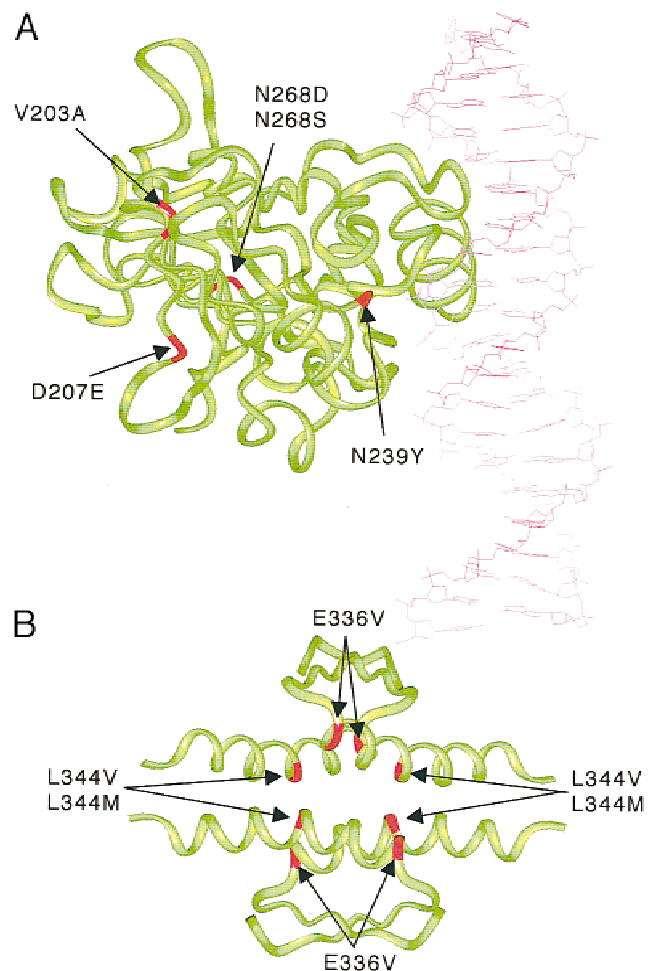
must therefore confer thermostability on their own. Other isolated variants contain reiterated sequence substitutions, N268D, N268S, V203A, E336V, L344V, L344M, L364H, and L364P; the bold sequence substitutions are in Table 1. These substitutions are also likely to be stabilizing in isolation.



**Fig. 8.** Temperature sensitivity of a thermostable p53 variant. Twenty nanograms of the parental p53 $\Delta$ 30-his or N239Y/N268D/E336V p53 $\Delta$ 30-his protein were pre-incubated for 10 min at the indicated temperature prior to carrying out band-shift analyses as in Figure 6 and Materials and methods. The positions of the shifted bands are again indicated.

The conformation of the sequence specific DNA binding “core” domain is likely to be conserved between the parental and evolved N239Y/N268D/E336V p53 $\Delta$ 30 proteins. Both p53 $\Delta$ 30 proteins bind DNAs that fit the p53 binding site consensus (p53CON, the p21 promoter and a half-site), but do not bind to a nonconsensus sequence (data not shown). Therefore, the amino acid substitutions that impart thermostability can be mapped onto the crystal structure of the core domain (Fig. 9A; Cho et al., 1994), and rationalized in terms of their contributions to structure, stability, and function.

Sequence substitutions in the core domain by and large appear to stabilize the beta-sandwich that forms the backbone of the domain. The aspartate at position 207 is at the end of beta-strand six, and its side chain makes hydrogen bond contacts with the side-chain nitrogen of the His214 in beta-strand seven, and the main-chain carbonyl oxygen of Phe212. The glutamate sequence substitution we have isolated at position 207 (clones 10, 11, and 15) would extend the side chain by a methylene group, possibly forming more stable versions of the hydrogen bonds found in the wild-type, and thereby stabilizing the beta-sandwich. The main chain of Val203, which is at the other end of beta-strand 6, forms a hydrogen bond



**Fig. 9.** Thermostabilizing mutations. (A) The main chains of the crystal structure of the sequence-specific DNA binding domain and (B) the NMR structure of the oligomerization domain are shown in green. Loci that were mutated more than once (see Table 1) are shown in red.

with Asp200. The alteration of Val203 to alanine (clones 12 and 13) should eliminate a steric constraint and optimize the position of the proton-donating main-chain nitrogen. The sequence substitution N268D (clones 1–5) might form alternate hydrogen bonds to the main-chain carbonyl of Arg267, as previously suggested by Brachmann et al. (1998). Our variant N268S (clone 7) would form a similar hydrogen bond, potentially more energetically favorable than in the wild-type. The sequence substitution N239Y (clones 1, 2, 3, and 8) could introduce an additional contact to the DNA ligand (Brachmann et al., 1998).

The sequence substitutions L344V (clone 6), L344M (clone 14), and E336V (clones 1, 2, and 4) are in the tetramerization domain (Fig. 7B; Lee et al., 1994). The interactions between the subunits of the wild-type tetramerization domain are largely hydrophobic (Lee et al., 1994; Jeffrey et al., 1995), and sequence substitutions that increase the hydrophobicity of this region may enhance the overall stability of the tetramer or higher order structures. The side chain of Leu344 participates in a hydrophobic interaction with its counterpart in one of the other monomers. Alteration of this residue to proline causes Li-Fraumeni syndrome, a genetic predisposition toward cancer (Varley et al., 1996). The more conservative changes to valine or methionine might instead optimize packing between the monomers. Similarly, the wild-type Glu336 side chain does not form any hydrogen bonds or salt bridges in the tetramer, and the change to valine may generate new hydrophobic interactions with adjacent side chains. Alternatively, such interactions may occur with the side chains of other tetramers. The formation of enhanced or novel quaternary structural interactions is consistent with experimental data. The evolved p53 $\Delta$ 30 protein forms specific complexes with DNA that are of higher molecular weight than the parental complexes (Figs. 6, 8).

An obvious caveat in this interpretation is that the higher molecular weight protein:nucleic acid complexes represent aggregates rather than defined quaternary structures. While this is possible, several facts argue against the possibility that nonspecific aggregates form, or that such aggregates are alone responsible, for enhanced thermostability. First and foremost, the evolved protein retains DNA binding specificity, which seems unlikely if some of the protein were denatured or unnaturally aggregated. Second, since the full-length, wild-type p53 is sometimes found to nonspecifically aggregate and remain in the well during gel-shift analyses, it is possible that the evolved, thermostable variant also nonspecifically aggregates. However, the shortened p53 $\Delta$ 30 variant that is a starting point for our experiments, shows no tendency to aggregate (see Hupp et al., 1992; Anderson et al., 1997; Figs. 6, 8). Thus, it is unlikely that the altered gel mobility of the evolved, thermostable variant is mechanistically similar to the nonspecific aggregation of the full-length p53. Indeed, while the nonspecific aggregation of full-length p53 seems to be related to its thermolability (Hansen et al., 1996), the altered gel mobility of the shortened, evolved p53 $\Delta$ 30 is instead correlated with its thermostability. Third, it is possible that DNA binding rather than quaternary structural interactions mediates aggregation and altered mobility. Since the oligonucleotide probe p53CON contains four binding sites for p53 monomers, models can be envisioned in which p53 variants, with improved DNA binding abilities, form oligonucleotide cross-linked “networks” with enhanced thermal stability. To test this possibility, the thermostable p53 variants were probed with a half-site that contained only two monomer binding sites. The half-site should not have been able to mediate network formation according to extant models for p53:DNA interactions (Cho et al., 1994; Jef-

frey et al., 1995). Nonetheless, the evolved variant still specifically bound the half-site, and still showed altered gel mobility. Overall, to the extent that thermostability and altered gel mobility are correlated, the most parsimonious explanation for the correlation is the formation of enhanced or novel quaternary structural interactions.

Finally, we isolated sequence changes at position 364 at the carboxy terminal end of our p53 $\Delta$ 30 construct (L364P, clones 1 and 4; L364H, clone 11). These sequence substitutions further alter a seemingly innocuous carboxy terminal sequence addition (L364/E365) that was originally added to the “wild-type” p53 $\Delta$ 30 gene to generate a novel restriction site. One explanation for the selection of further alterations at position 364 is that the A364L substitution may have destabilized the parental protein. However, we do not find our parental protein (+L364/E365) to be significantly more thermolabile than the “wild-type” p53 $\Delta$ 30 protein (which terminates after 363 amino acids). The in vitro inactivation rate of our parental protein at 37 °C (Figs. 3, 4) is apparently similar to that of the “wild-type” p53 $\Delta$ 30 protein as determined by Hansen et al. (1996). Moreover, variants that contain the L364P and L364H substitutions are not significantly more thermostable than variants that do not contain these substitutions (clone 1 vs. clone 2, or clone 10 vs. clone 11) (data not shown). Nonetheless, the fact that a “rationally” engineered substitution has been further adapted for function suggests that our screen for thermostability is extremely sensitive.

The validity of these structural hypotheses can only be determined by structural studies. However, the full-length, wild-type protein has not yet been crystallized and structural studies with the sequence substitutions that we have identified may prove to be much easier. Thermostable mutants are inherently less prone to aggregation and are conversely more amenable to crystallization. Moreover, to the extent that our variants are frozen in the active conformation, the solution of a full-length structure that contains these variants should provide a better target for drug design than even the structure of the full-length, wild-type protein, which could potentially prefer one of several conformers in a crystal.

#### Implications for gene therapies

By relying solely on the DNA binding ability of p53 as an assay, there was no assurance that the variants recovered would retain other cellular functions of p53. However, two lines of evidence indicate that the thermostable variants we have isolated are likely fully functional. First, three sequence substitutions at loci identified multiple times in our screens, V203A, D207E, and N268S, also appear during the natural evolution of p53. Ala203 and Glu207 are found in mouse p53 (among many others), while Ser268 is found in sheep (Soussi & May, 1996). Second, it is gratifying to note that two of the variants identified in our screen, N268D and N239Y, were previously identified in a yeast-based screen for second site suppressors of p53 mutations that promoted oncogenesis (Brachmann et al., 1998). The N268D and N239Y variants were each able to suppress subsets of p53 cancer mutations in reporter gene and apoptosis assays carried out in cultured mammalian cells. Thus, out of the thousands of possible stabilizing substitutions that could have been identified in our screen, the substitutions that were most frequently identified and that yielded the most thermostable proteins also retained p53 activities for which the mutants were not directly screened or selected. These results strongly argue that p53 variants recovered in *E. coli* may prove useful in mammalian cells as well. The p53 protein has already been shown to be a successful

reagent for gene therapies for human cancers (Roth et al., 1996). The longer lifetimes and greater DNA binding activities of the thermostable variants should make them even more useful in gene therapies based on *p53*.

The congruence of sequence substitutions resulting from different screens suggests that there may be relatively few mutational pathways for the recovery of improved function in *p53*. In fact, the most likely explanation for the recovery of mutations that can both promote thermostability and suppress *p53*-mediated oncogenesis is that these mutations alter the equilibrium between different *p53* conformers, and preferentially stabilize the active conformation. If this hypothesis is true, then the screen for thermostability we have developed can potentially be used to sequentially identify *p53* variants with a variety of improved, linked properties. It has been shown that stabilizing mutations tend to be additive in effect (Zhang et al., 1995). Thus, by iteratively identifying *p53* variants with more sequence substitutions that are successively more thermostable, we may also generate *p53* variants with progressively longer lifetimes *in vivo*, which can be more readily activated by tumor suppressors, and which can suppress a wider variety of oncogenic alleles. We are in the process of assessing these possibilities.

## Materials and methods

### Materials

DNA modifying enzymes were from New England Biolabs (Beverly, Massachusetts), except Taq polymerase, which was expressed and purified as described by Grimm and Arbuthnot (1995). Sequitherm Excel Long-Read DNA sequencing kits were from Epicentre (Madison, Wisconsin). Expression vector pET20b(+) and *E. coli* strain BL21(DE3)/pLysS were from Novagen (Madison, Wisconsin). *E. coli* strain InvaF' was from Invitrogen (Carlsbad, California). DNA purification columns were purchased from Qiagen (Chatsworth, California). Gamma labeled P-32 ATP was from ICN (Costa Mesa, California), and butterfly nitrocellulose membranes were from Schleicher and Schuell (Keene, New Hampshire). Plasmid pBluescript KS(+) was from the Stratagene Corp. (La Jolla, California). HiTrap chelating columns were from Pharmacia Biotech (Piscataway, New Jersey). SpectraPor dialysis tubing was from Spectrum (Laguna Hills, California), GelCode Blue staining reagent was from Pierce (Rockford, Illinois). The wild-type human *p53* cDNA gene (ATCC number 57254) was obtained from the ATCC (Rockville, Maryland). Other chemicals were from Sigma Chemicals (St. Louis, Missouri).

### Methods

#### Construction of *p53Δ30*

The human *p53Δ30* gene was amplified by the PCR using primers 5'-TTCATATGGAGGAGCCGCGAGTCAGATC-3' and 5'-CCGGTACCTACTCGAGCCTGCTCCCCCTG-3', and blunt-end ligated into the Eco RV site of pBluescript using standard methods (King & Blakesley, 1986). A Nde I site introduced at the 5' end of the gene and a Xho I site introduced at the 3' end of the gene are underscored; the Nde I site does not alter the amino acid coding sequence of the gene, but the Xho I site introduces the A364L/H365E substitutions mentioned in Discussion. *E. coli* cells were transformed by the method described by Inoue et al. (1990). The subcloned genes were sequenced at the Indiana University

Core Facility (Lincoln, Nebraska) using a LI-COR protocol. The *p53Δ30* gene was subcloned into the prokaryotic expression vector pET20b(+) via Nde I and Bam HI restriction sites. Random mutations were introduced into the gene by mutagenic PCR (Cadwell & Joyce, 1992) using primers 5'-AGATCTCGATCCCCGCGAAA TTAATACGA-3' and 5'-CGGGCTTTGTTAGCAGCCGGATCTC-3', which anneal to the pET vector outside the *p53* insert (100 nM primers, 60 mM Tris-HCl pH 8.5, 15 mM (NH<sub>4</sub>)SO<sub>4</sub>, 4.9 mM MgCl<sub>2</sub>, 0.25 mM MnCl<sub>2</sub>, 0.2 mM dGTP, 0.2 mM dATP, 0.6 mM dTTP, and 0.6 mM dCTP; 35 cycles of 94 °C, 30 s; 65 °C, 30 s; 72 °C, 2 min). For DNA shuffling, the *p53Δ30* variants were PCR amplified using the same primers under standard conditions (100 nM primers, 60 mM Tris-HCl pH 9.0, 15 mM (NH<sub>4</sub>)SO<sub>4</sub>, 2.0 mM MgCl<sub>2</sub>, and 0.2 mM dNTPs; 35 cycles of 94 °C, 30 s; 67 °C, 30 s; 72 °C, 1 min), digested with DNase I and reassembled as described by Stemmer (1994), then subcloned back into pET20b(+) for the second round of screening. The carboxy terminal hexahistidine tag (-his) was added by subcloning the *p53Δ30* gene into pET20b(+) using the Nde I and Xho I restriction sites.

#### Thermostability screen

The screen for *p53* function is based on a method developed by Singh et al. (1989). *E. coli* strain BL21(DE3) cells carrying the T7 lysozyme expression vector, pLysS, were transformed with the *p53Δ30* expression construct and plated on Luria broth (LB) plates containing 34 μg/mL chloramphenicol and 100 μg/mL ampicillin (LB-chl/amp). After 16 h of growth at 37 °C, the colonies were adsorbed onto a nitrocellulose filter and transferred colony-side up to LB-chl/amp plates containing 0.5 mM IPTG to induce expression of *p53Δ30*. The cells remaining on the original plate were regrown into full colonies by a further 8 h incubation at 37 °C. The colonies adsorbed to the nitrocellulose filter were induced at 23 °C for 4 h.

The cell membranes of the *p53Δ30*-expressing colonies were disrupted with chloroform gas for 15 min, giving the intracellular T7 lysozyme access to the peptidoglycan. The remaining manipulations were carried out in *p53* binding buffer (10 mM Tris-HCl pH 7.6, 50 mM NaCl, 10% glycerol, 0.1% Triton X-100, and 5 mM beta-mercaptoethanol) at 23 °C. The filters were treated with 2.5 U/mL DNase I (in *p53* binding buffer plus 10 mM MgCl<sub>2</sub>) for 15 min, blocked with 5% nonfat dry milk (in *p53* binding buffer plus 40 mM Tris-HCl pH 7.6) for 1 h, washed three times in binding buffer for 5 min each and probed with 20 nM radiolabeled oligonucleotide containing the *p53*CON sequence (underscored): 5'-GGACATGCCCGGGCATGTCCAT (Funk et al., 1992) (plus 5 μg/mL denatured salmon sperm DNA). The filters were each washed four times more for 7.5 min, and the quantity of probe bound to each filter was detected using a Phosphorimager (Molecular Dynamics, Sunnyvale, California).

#### Protein purification and characterization

*E. coli* BL21(DE3) cells were transformed with constructs that expressed the parental or N239Y/N268D/E336V *p53Δ30* genes fused to carboxy terminal six histidine tags (*p53Δ30*-his). The transformants were grown in LB-chl/amp at 37 °C to mid-log (Abs<sub>600</sub> = 0.3), then induced with 0.5 mM IPTG and shaken at 23 °C for 4 h. The cells were spun down and stored as a pellet at -20 °C. The protein was purified as described in the pET manual (Novagen) except that the binding, wash, and elution buffers were replaced with *p53* binding buffer without beta-mercaptoethanol,

plus 450 mM NaCl, and 0, 60, and 1,000 mM imidazole, respectively. In short, the cells were lysed by sonication and the insoluble fraction removed by centrifugation. The protein was loaded onto HiTrap chelating columns, washed, and eluted. The eluted p53 $\Delta$ 30-his proteins were dialyzed in p53 binding buffer (plus 450 mM NaCl) and stored in p53 binding buffer (plus 45% glycerol and 450 mM NaCl) at  $-20^{\circ}\text{C}$ . The protein was of the expected size and >95% pure as judged by denaturing polyacrylamide electrophoresis and staining with GelCode Blue staining reagent (Fig. 3). The concentration of the triple substitution variant was determined by amino acid analysis. The concentration of the parental protein was determined by a Bio-Rad protein assay, using the previously quantitated triple substitution variant as a standard.

The DNA binding assays were performed as described by Hupp et al. (1992). Twenty nanograms of p53 $\Delta$ 30 in 10  $\mu\text{L}$  of p53 binding buffer were pre-incubated for the temperatures and times indicated in Figures 6 and 8. The inactivation reactions were placed on ice, and 100 ng acetylated BSA, 90 ng pBluescript (Stratagene, La Jolla, California), and 4 ng radiolabeled p53CON all in two microliters total volume were added. The purified proteins were also incubated with radiolabeled, double-stranded oligonucleotides that either fit the DNA binding consensus (5'-GAACATGTCCCAACATGTTG-3' paired with 5'-CAACATGTTGGGACATGTTTC-3', and 5'-AACGTTGGACATGCCCCGGGGC-3 paired with 5'-GCCCCGGGGCATGTCCAACGTT-3') or did not fit the DNA binding consensus (5'-GGCACGTGAGCCTGCA GGGG-3' paired with 5'-CCCCTGCAGGCTCACGTGCC-3) (data not shown). After a 30 min incubation, the mixtures were loaded onto 4% acrylamide gels (19:1 acrylamide to bis-acrylamide, 0.33  $\times$  TBE, 0.1% Triton X-100) and run at 200 V for 1 h at  $4^{\circ}\text{C}$ . The gels were dried and the binding assessed using the Phosphorimager. The data were fit to an exponential decay model:

$$y = y_0 * e^{(-k_1 * t)} \quad (1)$$

by the least-squares method using the program KaleidaGraph 3.0.5 (Adelbeck Software, Reading, Pennsylvania), where  $y$  is the DNA binding activity at any given time (in units of pixel intensity measured by the Phosphorimager);  $y_0$ , the DNA binding activity at time zero; and  $k_1$  is the rate constant. Given the derived  $k_1$  values, half-times are the value for  $t$  when  $y = \{y_0/2\}$ . Error values refer to the standard error of the fit.

#### Note added in proof

While this manuscript was in press, Nikolova et al. (1998) reported that they followed a "semirational" approach to make and characterize some of the mutations identified in this study.

#### Acknowledgments

This work was supported by the Office of Naval Research (N000149610341). I.M. was supported in part by a Walther Cancer Research Center postdoctoral research fellowship, and in part by a National Science Foundation/Alfred P. Sloan Postdoctoral Research Fellowship in Molecular Evolution (DBI-9750002). We are grateful to Sabine Bell for synthesizing oligonucleotides and to members of the Ellington group for helpful discussion.

#### References

Alber T, Wozniak JA. 1985. A genetic screen for mutations that increase the thermal stability of phage T4 lysozyme. *Proc Natl Acad Sci USA* 82:747-750.

Anderson ME, Woelker B, Reed M, Wang P, Tegtmeyer P. 1997. Reciprocal interference between the sequence-specific core and nonspecific C-terminal DNA binding domains of p53: Implications for regulation. *Mol Cell Biol* 17:6255-6264.

Brachmann RK, Yu K, Eby Y, Pavletich NP, Boeke JD. 1998. Genetic selection of intragenic suppressor mutations that reverse the effect of common p53 cancer mutations. *EMBO J* 17:1847-1859

Cadwell RC, Joyce GF. 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl* 2:28-33.

Cho Y, Gorina S, Jeffrey PD, Pavletich NP. 1994. Crystal structure of a p53 tumor suppressor-DNA complex: Understanding tumorigenic mutations. *Science* 265:346-355.

Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA Jr, Butel JS, Bradley A. 1992. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumors. *Nature* 356:215-221.

El-Deiry WS, Kern SE, Pietenpol JA, Kinzler KW, Vogelstein B. 1992. Definition of a consensus binding site for p53. *Nat Genet* 1:45-49.

Funk WD, Pak DT, Karas RH, Wright WE, Shay JW. 1992. A transcriptionally active DNA binding site for human p53 protein complexes. *Mol Cell Biol* 12:2866-2871.

Grimm E, Arbutnot P. 1995. Rapid purification of recombinant tag DNA polymerase by freezing and high temperature thawing of bacterial expression cultures. *Nucleic Acids Res* 23:4518-4519.

Hansen S, Hupp TR, Lane DP. 1996. Allosteric regulation of the thermostability and DNA binding activity of human p53 by specific interacting proteins. *J Biol Chem* 271:3917-3924.

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

Hupp TR, Lane DP. 1994. Allosteric activation of latent p53 tetramers. *Curr Biol* 4:865-875.

Hupp TR, Meek DW, Midgley CA, Lane DP. 1992. Regulation of the specific DNA binding function of p53. *Cell* 71:875-886.

Inoue H, Nojima H, Okayama H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96:23-28.

Jeffrey PD, Gorina S, Pavletich NP. 1995. Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms. *Science* 267:1498-1502.

Kern SE, Pietenpol JA, Thiagalingam S, Seymour A, Kinzler KW, Vogelstein B. 1992. Oncogenic forms of p53 inhibit p53-regulated gene expression. *Science* 256:827-830.

King PV, Blakesley RW. 1986. Optimizing DNA ligations for transformations. *Focus* 8:30-32.

Ko LJ, Prives C. 1996. p53: Puzzle and paradigm. *Genes Dev* 10:1054-1072.

Lee W, Harvey TS, Yin Y, Yau P, Litchfield D, Arrowsmith CH. 1994. Solution structure of the tetrameric minimum transforming domain of p53. *Nat Struct Biol* 1:877-890.

Leung DW, Chen E, Goeddel DV. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* 1:11-15.

Levine AJ. 1997. p53, The cellular gatekeeper for growth and division. *Cell* 88:323-331.

Milner J. 1995. Flexibility: The key to p53 function? *Trends Biochem Sci* 20:49-51.

Nikolova PV, Henckel J, Lane DP, Fersht AR. 1998. Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proc Natl Acad Sci USA* 95(25):14675-14680.

Okorokov AL, Ponchel F, Milner J. 1997. Induced N- and C-terminal cleavage of p53: A core fragment of p53, generated by interaction with damaged DNA, promotes cleavage of the N-terminus of full-length p53, whereas ssDNA induces C-terminal cleavage of p53. *EMBO J* 16:6008-6017.

Reed M, Woelker B, Wang P, Wang Y, Anderson ME, Tegtmeyer P. 1995. The C-terminal domain of p53 recognizes DNA damaged by ionizing radiation. *Proc Natl Acad Sci USA* 92:9455-9459.

Roth JA, Nguyen D, Lawrence DD, Kemp BL, Carrasco CH, Ferson DZ, Hong WK, Komaki R, Lee JJ, Nesbitt JC. 1996. Retrovirus-mediated wild-type p53 gene transfer to tumors of patients with lung cancer. *Nat Med* 2:985-991.

Sakaguchi K, Herrera JE, Saito S, Miki T, Bustin M, Vassilev A, Anderson CW, Appella E. 1998. DNA damage activates p53 through a phosphorylation-acetylation cascade. *Genes Dev* 12:2831-2841.

Scharer E, Iggo R. 1992. Mammalian p53 can function as a transcription factor in yeast. *Nucleic Acids Res* 20:1539-1545.

Shaw P, Freeman J, Bovey R, Iggo R. 1996. Regulation of specific DNA binding by p53: Evidence for a role for O-glycosylation and charged residues at the carboxy-terminus. *Oncogene* 12:921-930.

- Singh H, Clerc RG, LeBowitz JH. 1989. Molecular cloning of sequence-specific DNA binding proteins using recognition site probes. *Biotechniques* 7:252–261.
- Soussi T, May P. 1996. Structural aspects of the p53 protein in relation to gene evolution: A second look. *J Mol Biol* 260:623–637.
- Stemmer WPC. 1994. DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. *Proc Natl Acad Sci USA* 91:10747–10751.
- Varley JM, McGown G, Thorncroft M, Cochrane S, Morrison P, Woll P, Kelsey AM, Mitchell EL, Boyle J, Birch JM. 1996. A previously undescribed mutation within the tetramerization domain of TP53 in a family with Li-Fraumeni syndrome. *Oncogene* 12:2437–2442.
- Wieczorek AM, Waterman JL, Waterman MJ, Halazonetis TD. 1996. Structure-based rescue of common tumor-derived p53 mutants. *Nat Med* 2:1143–1146.
- Zhang XJ, Baase WA, Shoichet BK, Wilson KP, Matthews BW. 1995. Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. *Protein Eng* 8:1017–1022.
- Zhao H, Arnold FH. 1997. Functional and nonfunctional mutations distinguished by random recombination of homologous genes. *Proc Natl Acad Sci USA* 94:7997–8000.