Long-Range Effects of Mutating R248 to Q/W in the p53 Core Domain

Sergey Yu Noskov,[†,§,||] Jon D. Wright,[†,||] and Carmay Lim*[,†,‡]

Institute of Biomedical Sciences, Academia Sinica, 11529 Taipei, Taiwan, and Department of Chemistry, National Tsing-Hua University, 300 Hsinchu, Taiwan

Received: September 25, 2002

The mutations of R248 to Trp and Gln in the core domain (CD) of the p53 protein are some of the most common mutations found in human cancer. Although the mutant 248Q and 248W p53-CDs retain the wild-type conformation and stability, they lack sequence-specific DNA binding and transactivation functions and the ability to suppress cell growth. The structural and energetic bases for the observed loss of DNA binding are unclear as the DNA-free and DNA-bound mutants are not available. Hence, we have generated three-dimensional models of the wild type, 248Q, and 248W p53-CDs, free and bound to DNA, using molecular dynamics simulations in the presence of explicit water molecules. Based on the simulation structures, the free energies of wild type and mutant p53-CDs binding to DNA have been computed and decomposed into component energies (electrostatic vs van der Waals vs cavity) and contributions from the interface residues. The DNA-free mutant structures were found to be consistent with antibody-binding and NMR data. The predicted DNA binding losses of both mutants were also in accord with experimental data. The calculations revealed that mutating R248 in the minor groove yielded long-range changes in the major groove DNA-binding interface, and the extent of these changes differs depending on the mutation type. The DNA-binding loss of the 248Q p53-CD mutant is due mainly to the loss of *major* groove contacts from K120 (located ∼20 Å from the mutation site) as well as unfavorable interactions of D281. In contrast, the DNA-binding loss of the 248W p53-CD mutant is due mainly to the loss of *minor* groove contacts from the mutant residue itself. It is also due, to a lesser extent, to the loss of major groove contacts from loop L1 and to the poorer packing of the protein−DNA interface in the 248 W mutant complex relative to the wild-type. The results obtained here are important in deciding the rescue strategy of mutant p53 DNA binding.

Introduction

The DNA-binding protein, p53, functions as a tumor supressor in human cells by activating transcription of genes that mediate cell cycle arrest and DNA damage repair,[1−4] and by inducing apoptosis in response to DNA damage, or other genotoxic stress such as hypoxia, oncogene activation, and viral infection.[5−7] Mutations of this genome guardian are associated with about 50% of human cancers,[8] 95% of lung cancers,[9] and has also been associated with resistance to therapy.[10,11] The human p53 protein comprises 393 amino acids that can be grouped[4] into (i) a transcription/transactivation region (residues 1−42), (ii) a sequence-specific DNA-binding core domain (residues 102−292), (iii) a tetramerisation domain (residues 323−356), and (iv) a C-terminal region (residues 363−393) that acts as a negative regulator of p53 sequence-specific DNA binding[12] and can also weakly bind DNA/RNA.

The most common type of p53 mutations are missense mutations that result in full-length, albeit mutant, proteins. This fraction of mutations is much higher than in other tumor supressor genes such as APC, ATM, and BRCA-1.[13] The majority of the p53 mutations found in human cancers map to the p53 core domain[14] (referred to as p53-CD) and ∼40% of them occur at R175, G245, R248, R249, R273, and R282.[8] Based on the 2.2 Å X-ray structure of a 21-base pair DNA duplex and three p53-CDs,[15] mutations of R175, G245, R249, and R282 have been classified as *conformational* because these four residues do not contact the DNA directly, whereas mutations of R248 and R273 have been classified as *contact* because the side chains of these two residues contact (within 4 Å) phosphate oxygens in the minor and major DNA groove, respectively. R248 is located in loop L3 (residues 237−250), while R273 is near the C-terminus of β-strand S10 (residues 264−274), which is part of the loop-sheet-helix motif consisting of loop L1 (residues 112−124), a three-stranded β sheet (residues 124−141 and 271−274), and helix H2 (residues 278−286).

Because the R248 and R273 contact mutants account for about one fifth of all p53 mutations and are assumed to retain a native-like three-dimensional (3D) structure, they have attracted considerable interest as targets for mutant p53 rescue.[8,16,17] Here, we have chosen to study the R248 → W and R248 → Q p53-CDs because they are two of the most common tumor-derived mutants[15] and have been found in brain, breast, colon, esophageal, head and neck, skin, lung, and liver cancer.[13] The p53-CD·DNA X-ray structure shows that on the DNA side, the R248 side chain has "*close*" contacts (defined by a heavy donor−heavy acceptor distance >3.5 Å but ≤4 Å) to the Thy 12′ and Thy 14 phosphate oxygens in the minor groove, while on the protein side, it is hydrogen bonded to S240, which is flanked by two zinc ligands, C238 and C242. When R248 is mutated to Gln or Trp in full-length p53 the wild-type

* Corresponding author. E-mail: carmay@gate.sinica.edu.tw. Phone: 886-2-2652-3031. Fax: 886-2-2788-7641.
† Academia Sinica.
‡ National Tsing-Hua University.
§ Permanent address: Institute of Solution Chemistry, Russian Academy of Sciences, Akademicheskaya str., 1, 153045, Ivanovo, Russia.
|| The first two authors should be regarded as joint first authors.

10.1021/jp022140w CCC: $22.00
Published on Web 11/19/2002

conformation[18] (i.e., PAb1620+, PAb240−, Hsp 70−) is retained. However, both 248Q and 248W p53 mutants lack sequence-specific DNA binding and sequence-specific transactivation function and growth suppression.[3,19−21]

Our goal here is to identify the key structural changes and residues governing the loss of DNA binding upon mutating R248 to a neutral Gln or Trp in the p53-CD. Because the DNA-free and DNA-bound 3D structures of these two 248 mutants have not been reported, their 3D solution structural models, free and DNA bound, were obtained using molecular dynamics (MD) simulations with explicit water molecules. The simulation protocol (see Methodology) had been calibrated in our previous work,[22] which showed that the free and DNA-bound wild-type p53-CD simulations preserved the key X-ray protein−protein and protein−DNA interactions as well as the qualitative features of the X-ray B-factors. As another validation check of the simulation protocol, albeit indirect, our previous work also showed that the DNA binding affinity and specificity of the p53-CD were lost upon mutating R273 → H, but restored by a second mutation of T284 → R, in accord with experimental findings.[22]

Here the computed DNA-free mutant structures were verified by their consistency with antibody-binding and NMR data. Based on the average MD structures of wild-type and mutant p53-CDs, the DNA-binding free energy of the 248 mutant relative to that of the wild-type protein was computed and compared to experiment. By decomposing the free energy change upon mutation into changes in the component energies and individual residue contributions, and correlating the energy changes with changes in the hydrogen bonding interactions and solvent accessibility, the key contributions to the loss of p53-CD binding upon mutating R248 → W/Q were identified. The implications of the results obtained here to the strategy of identifying potential linker molecules that may restore the DNA binding affinity and specificity of the two p53-CD mutants studied are discussed.

Methodology

MD Simulations. The core domains of wild-type p53, mutant 248Q, and 248W, free and bound to a 15-base-pair DNA duplex, 5′-^{3}AATTGGGCAAGTCTA^{17}-3′, were subjected to MD simulation using the CHARMM[23] program at a mean temperature of 300 K. The total numbers of atoms in the simulations of the core domain of wild-type p53, mutant 248Q, and 248W p53, free and DNA-bound, are 37615, 37618, 37626, 58389, 58383, and 58385, respectively. The net charge of the wild-type protein (+5) bound to zinc (+2), 20 counterions (+20) (see below), and the 15-base-pair DNA duplex (−30) is −3.

Starting Structures. The initial 3D structure for the wild-type p53-CD-DNA simulation was taken from chain B of PDB entry 1TSR, in which the core domain (residues 96−289) is bound to a consensus DNA binding site, GGGCA AGTCT; the second p53-CD is bound to a nonconsensus DNA binding site, while the third p53-CD is DNA-free.[15] Initial 3D structures for the mutant p53-CD·DNA simulations were prepared by a side chain rotamer library method[24] using a locally modified version of the SCWRL program.[25,26] The most common $\chi 1$ and $\chi 2$ angles for the mutant residue with respect to the wild-type backbone Φ and Ψ angles at that position were identified and the side chain without hydrogen atoms was built using the AMBER94 force field.[27] Subsequently, the side chain was checked for steric clashes against the backbone using a simple distance-dependent potential that approximates the van der Waals (vdW) potential. If there were unfavorable interactions with the backbone, the rotamer was rejected and the next most common rotamer was used. This was repeated until no clashes or no more rotamers were left. The built rotamer was then checked for side chain−side chain interactions using the same distance-dependent potential as before. If there were side chain to side chain clashes, then more rotamers were searched and the one with the lowest conflict was employed. For the DNA-free proteins, the starting structure was extracted from the average MD structure of the respective complex.

Simulation Protocol. Each simulation was carried out at physiological pH (pH 7) using the all-hydrogen CHARMM27 force field[28] and the TIP3P model of water[29] with spherical boundary conditions.[30] The protonation state of the ionizable residues and the treatment of nonbonded vdW and electrostatic interactions are as described in our previous work.[22] Hydrogen atoms were added to the initial protein-DNA complex and the resulting structure was minimized in the presence of strong harmonic constraints on all heavy atoms to remove bad contacts. Next, the structure was immersed in the center of a previously equilibrated TIP3P water sphere of 50 Å radius. The size of the water sphere was chosen such that it provided at least three layers of solvation around the complex. Water molecules within 2.6 Å of any of the protein or DNA heavy atoms were deleted. The water sphere was rotated four times and overlaid on the previous system to ensure that no cavities were formed in the system. Twenty sodium ions were added to the protein−DNA system by replacing water molecules with the highest electrostatic energies on the oxygen atom and located ≥7 Å apart from each other and ≥4 Å apart from the protein or DNA atoms. The fully solvated protein−DNA complex was then further minimized (with the constraints on heavy atoms reduced to zero in successive rounds of minimization) to relieve bad vdW contacts. It was then equilibrated for 60 ps with heavy constraints on only the terminal DNA residues to prevent the DNA strands from unwinding.[31] These constraints were reduced and each simulation was continued for 250 ps, during which coordinates were saved every 100 fs.

As the X-ray structure shows that DNA binding does not alter the protein structure significantly,[15] the p53 protein was extracted from the average MD structure of each complex and subjected to 250 ps of MD simulation, assuming the same protonation state of the ionizable residues, force field, and boundary conditions as the bound-state simulations. The length of each trajectory appears to be sufficiently long to obtain equilibrium structures (see Results, also Figures 1 and 2). Furthermore, the same simulation length was employed in our previous work,[22] which showed that the free and DNA-bound wild-type p53-CD simulations generally preserved the key features of the respective X-ray structures and B-factors, while the DNA-bound R273H + T284R double mutant simulation could restore the native-like conformation of the protein−DNA interface, in accord with experimental findings.

Analyses. For each trajectory the root-mean-square deviations (RMSDs) of the backbone protein and DNA atoms (excluding the terminal DNA residues) from the initial structure as well as the B-factors for the protein backbone atoms were computed. The B-factors for the protein backbone atoms were computed over the sampled phase of the simulations with error bars being calculated by splitting the sampled phase into three parts of 50 ps each and recomputing the B-factors for each part. Hydrogen bond analyses were performed by computing the average distance between a light/heavy donor and heavy acceptor from the last 150 ps portion of the trajectory. A hydrogen bond was defined by a heavy donor−heavy acceptor distance ≤3.5 Å, a

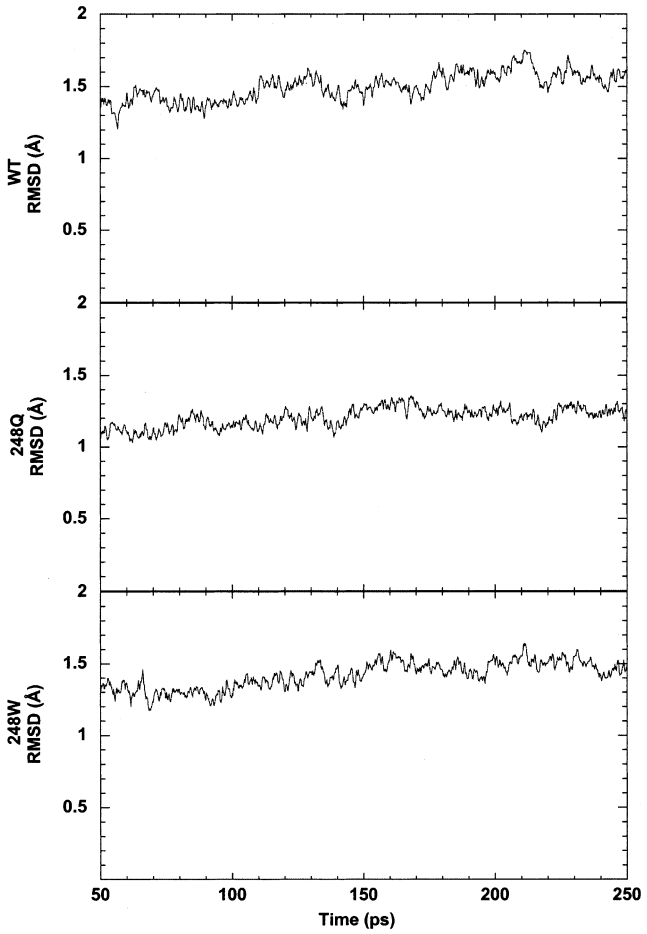

Figure 1. Plot of the RMSD of the backbone atoms from the initial structure as a function of simulation time in (a) wild-type, (b) R248→Q, and (c) R248→W p53-CDs.

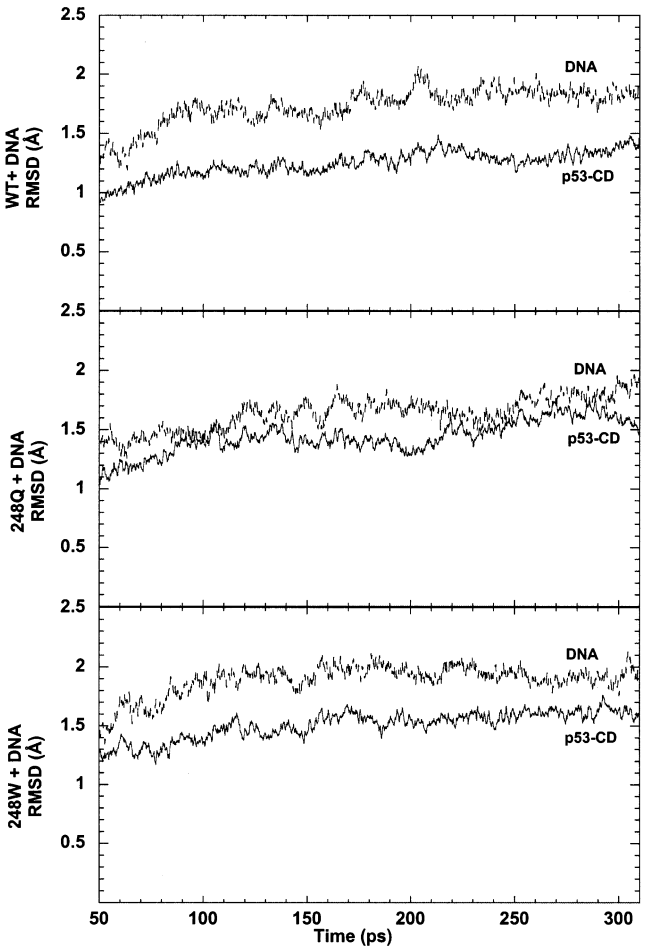

Figure 2. Plot of the RMSD of the protein and DNA backbone atoms from the initial structure as a function of simulation time in (a) wild-type p53-CD·DNA, (b) 248Q p53-CD·DNA, and (c) 248W p53-CD·DNA.

light donor−heavy acceptor distance ≤2.5 Å, and a deviation of less than ±60° from linearity. An average MD structure was obtained by averaging over the last 150 ps portion of the trajectory (see Results). Solvent accessible surface area (SASA)[32] calculations were performed on the MD structures using the GEPOL program[33] and a solvent probe radius of 1.4 Å for absolute values or using the MolMol program[34] for a percentage residue accessibility; i.e., the percentage ratio of the side chain X water-accessible surface area in the protein to that in the tripeptide −Gly−X−Gly−.

Free Energy Decomposition. *Relative Binding Free Energy*. The free energy difference between the wild-type and mutant p53-CD binding to DNA in solution, $\Delta\Delta G°$, is given by

$$\Delta\Delta G° = (\Delta G_{gas}^{mut} - \Delta G_{gas}^{wt}) + \{G_{solv}(p53^{mut}\cdot DNA) - G_{solv}(p53^{wt}\cdot DNA)\} - \{G_{solv}(p53^{mut}) - G_{solv}(p53^{wt})\} \quad (1)$$

ΔG_{gas} is the standard free energy change per mole for the noncovalent association of the p53-CD and DNA in the gas phase at 300 K, and $-G_{solv}$ corresponds to the work of transferring the molecule in its solution conformation to the same conformation in the gas phase at 300 K. In the following, a single Δ denotes the difference in energy/free energy between the DNA-bound and DNA-free state of wild-type or mutant p53-CD, and $\Delta\Delta$ represents the DNA-binding free energy of the mutant p53-CD relative to the wild-type protein. The $\Delta\Delta G°$ calculation was based on "snapshot" configurations taken every 5 ps from the last 100 ps of the MD trajectories.[35,36] All water molecules were removed from the MD trajectories to avoid boundary problems for the different sized systems, as the DNA-bound and free proteins had the same number of solute atoms but different numbers of solvent atoms. Previous works have shown that the free energy can be decomposed into physically meaningful components, even though the components are path dependent.[37−40] Here, ΔG_{gas} and G_{solv} are computed assuming that their energy/free energy components are additive.[41,42]

Gas-Phase Binding Free Energy Difference, $\Delta\Delta G_{gas}$. The free energy difference between wild-type and mutant p53-CD binding to DNA in the gas phase was estimated by

$$\Delta\Delta G_{gas} \sim \Delta\Delta E_{gas}^{vdW} + \Delta\Delta E_{gas}^{elec} = (\Delta E_{gas}^{vdW,mut} - \Delta E_{gas}^{vdW,wt}) + (\Delta E_{gas}^{elec,mut} - \Delta E_{gas}^{elec,wt}) \quad (2)$$

Equation 2 assumes similar conformational changes as well as similar vibrational entropy changes between wild-type and mutant p53-CDs binding to DNA. Based on the MD structures of wild type and mutant p53-CDs (see above), the changes in vdW and electrostatic energies upon binding were calculated using the CHARMM program[23] and an all-hydrogen force field[43] with unit dielectric constant and no cutoffs.

Solvation Free Energy, G_{solv}. The solvation free energy, G_{solv}, can be expressed as

$$G_{solv} = G_{solv}^{nonel} + G_{solv}^{elec} \quad (3)$$

The nonelectrostatic contribution to the solvation free energy, G_{solv}^{nonel}, was approximated by a linear function of the SASA; i.e.,

$$G_{solv}^{nonel} = \gamma \times \text{SASA} \quad (4)$$

where γ was set to 7.2 cal/mol/Å^2, as rationalized in previous works,[42,44] and the SASA of the free or bound molecule was computed using the GEPOL program (see Analyses above). Hence ΔG_{solv}^{nonel} is proportional to the loss in SASA at the protein−DNA interface.

The electrostatic contribution to the solvation free energy (G_{solv}^{elec}) was calculated using the finite-difference Poisson−Boltzmann (FDPB) method,[45] as implemented in CHARMM.[23] The protein was treated as a low dielectric medium ($\epsilon_{in} = 1$) surrounded by a high dielectric solvent ($\epsilon_{out} = 80$ for water). The ionic strength was set to 0.145 M. The low-dielectric region of the protein was defined as the region inaccessible to contact by a 1.4 Å sphere rolling over the molecular surface, defined by atomic coordinates of the MD structure and vdW radii taken from the all-hydrogen CHARMM27 force field. Due to the large dimension of the system and the long-range Coulombic forces, several focusing steps were performed to define proper boundary conditions. Initial electrostatic potential calculations for the wild-type and mutant structures employed a 200 × 200 × 200 grid with a spacing of 1.6 Å. Two focusing steps were performed to reach a final grid spacing of 0.4 Å.[46] The difference between the electrostatic potential calculated in solution and in the gas phase yielded the electrostatic contribution to the solvation free energy.

Results

Figures 1 and 2 show plots of the RMSDs of the backbone atoms from the initial coordinates of the p53-CD, free and bound to DNA, respectively, as a function of simulation time. For the free-state simulations (Figure 1), the protein backbone RMSDs of the wild-type, 248Q, and 248W p53-CDs plateaued after ∼160 ps and fluctuated around a mean of 1.55, 1.25, and 1.50 Å, respectively. For the bound-state simulations (Figure 2), the protein backbone RMSDs of the p53-CD·DNA complexes also plateaued after ∼160 ps, and fluctuated around a mean of 1.3 Å for the wild-type protein, and at a higher mean of 1.5 and 1.6 Å for the 248Q and 248W mutants, respectively. The protein backbone RMSDs of the MD structures from their respective initial ones are similar to those of NMR solution structures from their corresponding X-ray crystal structures[47] and also similar to analogous values reported for simulations of other proteins.[48−51] Relative to the protein backbone RMSD, the DNA backbone RMSD (not including the terminal bases) fluctuated around a higher mean of 1.8 to 2.0 Å, reflecting the greater mobility of the DNA compared to the protein. The last 150 ps of each trajectory were employed in analyses (see Methodology).

Changes in Mobility. Figure 3a shows that the mutant 248Q main-chain B-factors (averaged for the N, C^{α}, and C(=O) atoms) are generally higher than the corresponding wild-type values, indicating a relatively flexible structure. The mutant 248W main-chain B-factors are also generally higher than the corresponding wild-type values, except near the N-terminus and in the three loops connecting β-strands S3 and S4 (residues 147−155), S7 and S8 (residues 220−229), and S9 and S10 (residues 259−263). However, it is more rigid than the 248Q mutant, as its main-chain B-factors are generally lower than those of the 248Q mutant. When the protein is bound to DNA,

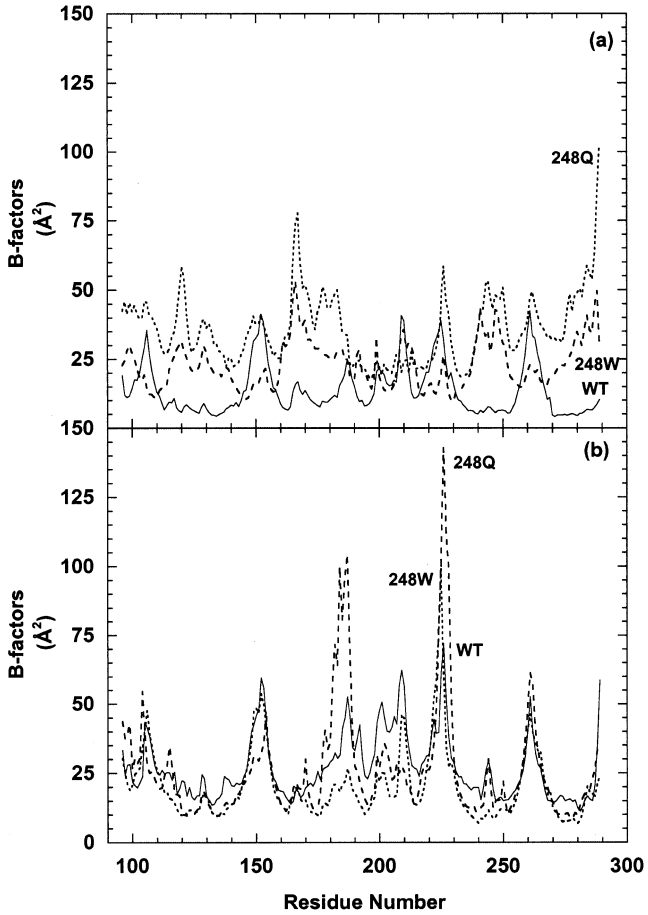

Figure 3. Plot of the B-factors averaged over the protein backbone atoms as a function of residue number in the simulations of (a) DNA-free and (b) DNA-bound p53-CDs. The solid, dotted, and dashed lines correspond to wild-type, 248Q, and 248W p53-CDs, respectively. The standard errors (not plotted) were 5.0, 6.4, and 3.7 Å^2 for the DNA-free wild-type, 248Q, and 248W p53-CDs, respectively, and 4.6, 2.5, and 3.4 Å^2 for the DNA-bound wild-type, 248Q, and 248W p53-CDs, respectively.

the two mutants and the wild-type protein exhibit similar trends in the main-chain B-factors, except that part of the L2 loop (residues 175−190) and the loop connecting β-strands S7 and S8 are disordered in the 248Q p53-CD·DNA complex (Figure 3b).

Figure 4 shows that in the 248Q mutant, the backbone RMSDs from the wild-type simulation structure for base-pairs 3−7, 9, and 10 are higher than the respective X-ray RMSDs, whereas those for base-pairs 11, 12, 14−17 are lower than the respective X-ray RMSDs. This indicates that the DNA major groove (base-pairs 6−10) in the 248Q p53-CD·DNA complex seems to be more flexible than that in the wild-type X-ray complex, whereas the DNA minor groove (especially base pairs 11 and 12) is more rigid. In the 248W mutant, the DNA major groove as well as minor groove are overall more flexible than those in the wild-type complex, except for base pairs 8, 11, and 12.

Changes in Component Free Energies Upon Mutation. The contributions to $\Delta\Delta G°$ (eq 1) from the protein−DNA interface residues were decomposed as follows:

$$\sum_i \Delta\Delta G_i = \sum_i \Delta\Delta G_{i,gas}^{vdW} + \sum_i \Delta\Delta G_{i,gas}^{elec} + \sum_i \Delta\Delta G_{i,solv}^{nonel} + \sum_i \Delta\Delta G_{i,solv}^{elec} \quad (5)$$

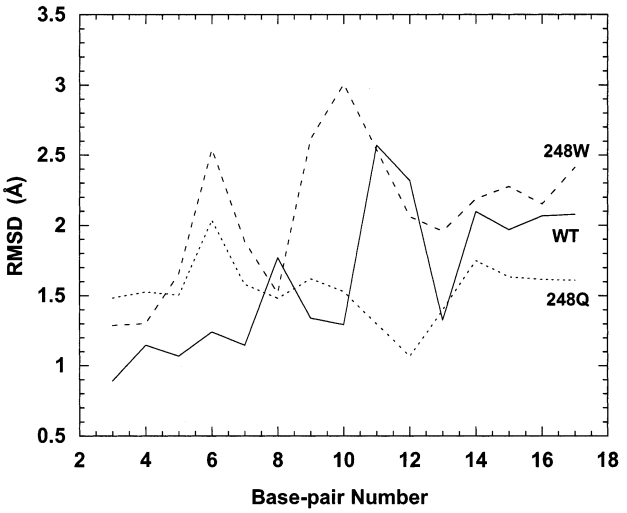

Figure 4. Plot of the RMSD of the DNA backbone atoms (averaged over the base pair) from the wild-type p53-CD·DNA simulation structure as a function of base-pair number for the crystal structure (solid line), 248Q p53 (dotted line) and 248W p53 (dashed line).

TABLE 1: Changes in the DNA-binding Free Energy Components of the Mutant 248Q p53-CD Relative to Wild-type[a]

interface residue i	$\Delta\Delta G_i$	$\Delta\Delta G_{i,\mathrm{gas}}^{\mathrm{elec}}$	$\Delta\Delta G_{i,\mathrm{solv}}^{\mathrm{elec}}$	$\Delta\Delta G_{i,\mathrm{gas}}^{\mathrm{vdw}}$	$\Delta\Delta G_{i,\mathrm{solv}}^{\mathrm{nonel}}$
L1: A119	5.2	4.2	−0.6	1.4	0.2
L1: **K120**	**18.6**	**30.6**	**−14.5**	2.2	0.4
L1: **S121**	**−11.7**	−6.9	−2.2	−2.0	−0.6
Q136	−0.2	1.9	−2.1	0.0	0.0
L137	−0.1	0.6	−0.6	−0.0	−0.0
L3: N239	2.0	**9.6**	**−7.4**	−0.2	−0.0
L3: S240	0.7	2.9	−2.2	−0.0	0.0
L3: S241	−1.1	−2.7	1.8	−0.2	−0.0
L3: C242	0.0	3.2	−3.2	0.0	0.0
L3: M243	0.7	−0.0	0.4	0.4	−0.1
L3: N247	4.3	5.2	−2.2	1.1	0.2
L3: 248Q	5.5	**143.4**	**−139.6**	1.4	0.3
L3: R249	−0.7	**−10.2**	**9.5**	−0.0	0.0
S10: V272	0.0	0.0	0.0	0.0	0.0
S10: R273	−5.4	−5.0	0.1	−0.5	0.0
S10: V274	0.6	0.2	0.5	−0.2	0.0
C275	−0.1	0.3	−0.4	−0.0	0.0
A276	−0.9	1.0	−2.6	0.6	0.1
C277	−0.2	0.9	−1.0	−0.0	−0.2
H2: R280	−2.8	−6.1	3.1	0.3	−0.1
H2: **D281**	**14.6**	6.3	**8.6**	−0.2	−0.1
H2: R283	6.8	−0.1	6.9	0.0	−0.0
$\sum_i$[b]	**35.8**	**179.2**	**−147.7**	4.1	0.2

[a] Values greater than 7 kcal/mol are highlighted in bold. [b] Free energy contributions from all interface residues.

where the summation is over all interface residues (defined as residues whose gas-phase interaction energy changes by >1 kcal/mol upon DNA binding). The relative free energies for 248Q (Table 1) and 248W (Table 2) p53-CDs are distinguished by $\Delta\Delta G$ and $\Delta\Delta' G$, respectively. The subscript "gas" in eq 5 denotes protein−protein or protein−DNA interactions (collectively known as solute−solute interactions), whereas the subscript "solv" denotes protein−solvent interactions. In what follows, we first identify the interface residues whose $|\Delta\Delta G_i| \geq (1/5)|\sum_i \Delta\Delta G_i|$. We then correlate the changes in the respective residue−protein/DNA/water hydrogen bonding interactions with the changes in the corresponding interaction energies upon mutation. In interpreting the results below, we emphasize the *trend* (sign) of the large free energy changes (given to ≥2 significant figures below) upon each mutation rather than their absolute values.

R248 → Q. When R248 is mutated to Gln, the p53-CD mutant is predicted not to bind to DNA as $\Delta\Delta G°$ is large and positive (37 kcal/mol), stemming mostly from a net unfavorable free energy contribution of the interface residues ($\sum_i \Delta\Delta G_i = 36$ kcal/mol, Table 1). The loss of binding free energy upon mutating R248 → Q is due mainly to *unfavorable solute−solute* electrostatic effects ($\sum_i \Delta\Delta G_{i,\mathrm{gas}}^{\mathrm{elec}} = 179$ kcal/mol), which outweigh favorable solute−solvent electrostatic changes ($\sum_i \Delta\Delta G_{i,\mathrm{solv}}^{\mathrm{elec}} = -148$ kcal/mol) by ∼31 kcal/mol. Solute−solute vdW energy ($\sum_i \Delta\Delta G_{i,\mathrm{gas}}^{\mathrm{vdW}} = 4.1$ kcal/mol) and nonelectrostatic solvation free energy ($\sum_i \Delta\Delta G_{i,\mathrm{solv}}^{\mathrm{nonel}} = 0.2$ kcal/mol) changes of the interface residues, however, do not make significant contributions to the observed loss of binding affinity.

Table 1 shows that as expected, the mutant 248Q contributes *less favorably* than the wild-type R248 to DNA binding ($\Delta\Delta G_{248\mathrm{Q}} = 5.5$ kcal/mol); however, it is not the major contributor to the observed loss of binding affinity. Instead, the main contributors to the loss of DNA binding upon mutating R248 → Q in the minor groove are two charged residues in the *major* groove, K120 ($\Delta\Delta G_{\mathrm{K120}} = 19$ kcal/mol) and D281 ($\Delta\Delta G_{\mathrm{D281}} = 15$ kcal/mol). The R248 → Q mutation resulted in a loss of two "close" contacts (defined in Introduction) to the Thy 12′ and Thy 11′ phosphate oxygens and a water-bridged hydrogen bond to the Thy 12′ sugar oxygen (compare Figure 6 with Figure 5), hence $\Delta\Delta G_{248\mathrm{Q,gas}}^{\mathrm{elec}}$ is large and positive (143 kcal/mol, Table 1). However, it is nearly compensated by a similarly large but negative $\Delta\Delta G_{248\mathrm{Q,solv}}^{\mathrm{elec}}$ (−140 kcal/mol) because in the *unbound* state, the neutral glutamine is poorly solvated compared to the positively charged arginine ($G_{\mathrm{solv}}^{\mathrm{elec}} =$ −6 kcal/mol for 248Q vs. −100 kcal/mol for R248, Table 3). Thus, the mutant 248Q incurred a smaller desolvation penalty upon binding DNA compared to the wild-type R248 ($\Delta G_{\mathrm{solv}}^{\mathrm{elec}} =$ 22 kcal/mol for 248Q vs 162 kcal/mol for R248, Table 3).

To compensate for the loss of charge−charge interactions with 248Q, the negatively charged Thy 11′ phosphate O^{2P} oxygen formed two new "close" contacts with the nearby positively charged R273 guanidinium nitrogens (the average distance between Thy 11′ O^{2P} to the two R273 amino nitrogens decreased from 4.92 ± 0.44 Å in the p53-CD·DNA simulation to 3.67 ± 0.42 Å in the respective 248Q mutant simulation). Furthermore, the Thy 11′ $O^{5'}\cdots$R273 N^{H2} distance in the wild-type complex simulation (3.95 ± 0.40 Å) decreased to 3.27 ± 0.24 Å in the mutant complex simulation, but the Thy 11′ $O^{5'}\cdots$R273 H^{NH2} distance (2.70 ± 0.47 Å) indicates that this is not a hydrogen bond. The favorable R273−Thy 11′ charge−charge interactions in the mutant complex likely result in a negative $\Delta\Delta G_{\mathrm{R273,gas}}^{\mathrm{elec}}$ (−5.0 kcal/mol, Table 1), which dominates $\Delta\Delta G_{\mathrm{R273}}$ (−5.4 kcal/mol). Thus, R273 contributes favorably to DNA binding upon substituting R248 for Gln, effectively canceling the unfavorable contribution from the mutant residue ($\Delta\Delta G_{248\mathrm{Q}} = 5.5$ kcal/mol).

The proximity of the R273 side chain to the Thy 11′ phosphate in turn drew the D281 side chain, to which it is hydrogen bonded, to the Thy 11′ phosphate oxygen in the mutant structure (D281 $O^{\delta 2}\cdots$Thy 11′ O^{1P} decreased from 5.79 ± 0.52 Å in the wild-type complex simulation to 3.85 ± 0.33 Å in the mutant complex simulation). The proximity of these two negatively charged groups resulted in a positive $\Delta\Delta G_{\mathrm{D281,gas}}^{\mathrm{elec}}$ (6.3 kcal/mol, Table 1). Although the $R273^{+}-D281^{-}-R280^{+}$ hydrogen-bonded network observed in the free and DNA-bound wild-type p53-CD simulations is preserved in the 248Q p53-CD·DNA simulation, D281 interacted with water molecules

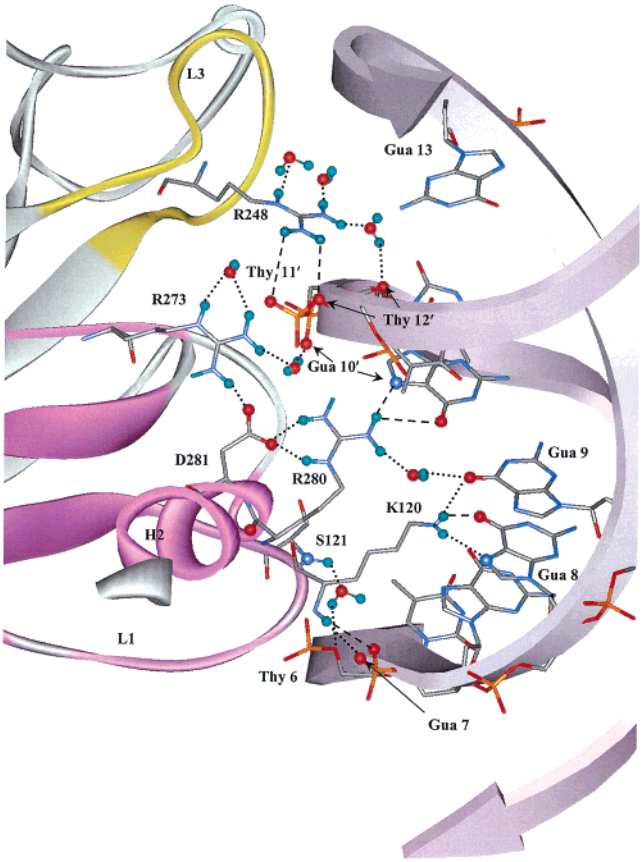

Figure 5. Hydrogen-bonding network in the average wild-type p53-CD·DNA structure in the vicinity of the major and minor grooves. Hydrogen bonds, as defined in Analyses, are depicted by dotted lines, whereas "close" contacts, as defined in Introduction, are depicted by dashed lines. Only certain key residues (discussed in text) and their hydrogen bonding partners are shown for the sake of clarity. The loop-sheet-helix motif is in magenta while the L3 loop is in yellow. Oxygen, phosphate, hydrogen, and nitrogen atoms are colored red, orange, green, and blue, respectively.

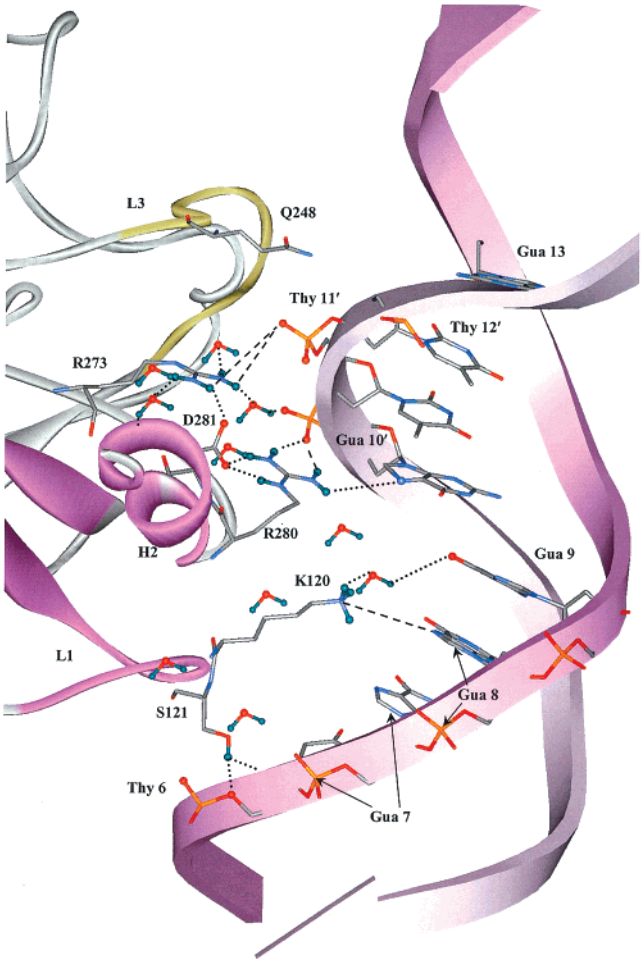

Figure 6. Hydrogen-bonding network in the average 248Q p53-CD·DNA structure in the vicinity of the major and minor grooves. S121 O^{γ} hydrogen bonds to Thy 6 O4′ (atom hidden from view) and Thy 6 O5′. See also legend to Figure 5.

rather than the R280 side chain but maintained its salt bridge with R273 in the free 248Q mutant simulation. Hence, the electrostatic solvation of the D281 side chain in the free mutant (G^{elec}_{solv} = 75 kcal/mol, Table 3) is not as unfavorable as that in the wild-type protein (G^{elec}_{solv} = 82 kcal/mol, Table 3), so $\Delta\Delta G^{elec}_{D281,solv}$ is positive (9 kcal/mol, Table 1). Consequently, both D281-DNA phosphate and D281-water interactions contribute unfavorably to DNA binding upon mutating R248 → Q in the wild-type p53-CD.

The proximity of the D281 side chain to the Thy 11′ phosphate in turn attracted the R280 side chain to hydrogen bond to a Gua 10′ phosphate oxygen (O^{1P}) in the mutant complex simulation, instead of a Gua 10′ base oxygen (O^6) seen in the wild-type complex simulation (R280 N^{H2}···Gua 10′ O^{1P} (3.80 ± 0.74 Å) decreased to 2.81 ± 0.16 Å in the 248Q mutant complex, whereas R280 N^{H1}···Gua 10′ O^6 (3.53 ± 0.50 Å) increased to 4.49 ± 0.51 Å; compare Figure 6 with Figure 5). The gain in the R280-phosphate salt bridge outweighed the loss of the R280-base hydrogen bond in the mutant complex, resulting in a negative $\Delta\Delta G^{elec}_{R280,gas}$ (−6.1 kcal/mol, Table 1) that almost canceled the unfavorable $\Delta\Delta G^{elec}_{D281,gas}$ contribution (6.3 kcal/mol). The observed changes in the R280-Gua 10′ interactions also caused R280 in the 248Q mutant to be less well solvated, especially when bound to the DNA. Consequently, upon binding DNA the R280 desolvation penalty in the 248Q mutant is slightly larger than that in the wild-type protein ($\Delta G^{elec}_{R280,solv}$ = 174 vs 171 kcal/mol, Table 3). The positive $\Delta\Delta G^{elec}_{R280,solv}$ (3.1 kcal/mol, Table 1) partially offsets the negative $\Delta\Delta G^{elec}_{R280,gas}$, hence R280 makes a small but favorable contribution to DNA binding upon mutating R248 → Q in the wild-type protein ($\Delta\Delta G_{R280}$ = −2.8 kcal/mol, Table 1).

The observed changes in the network of interactions among Thy 11′, R273, D281, R280, and Gua 10′ were accompanied by a distortion of the major groove conformation. This is suggested by the different behavior of the RMSDs of the major groove base pairs (6−10) in the 248Q mutant complex simulation and wild-type X-ray structure (compare dotted and solid lines in Figure 4). In particular, the RMSDs of base pairs 7−9 showed a V shape in the 248Q mutant complex simulation but an inverted V shape in the wild-type X-ray structure. This distortion in the Gua 7−9 conformation correlates with the loss of native hydrogen bonds from Gua 7−9 to K120: the Gua 7 O^{1P}···K120 N (3.21 ± 0.30 Å), Gua 8 N^7···K120 N^{ζ} (3.01 ± 0.21 Å), and Gua 9 O^6···K120 N^{ζ} (2.90 ± 0.21 Å) distances lengthen to 5.81 ± 0.42 Å, 3.81 ± 0.42 Å, and 5.34 ± 0.42 Å, respectively, in the mutant complex simulation. In addition, the close contact between Thy 6 $O^{3'}$ and K120 N (3.74 ± 0.39 Å) is lost (Thy 6 $O^{3'}$···K120 N = 7.46 ± 0.41 Å). Hence $\Delta\Delta G^{elec}_{K120,gas}$ is large and positive (31 kcal/mol, Table 1). The observed loss in K120-DNA electrostatic interactions is partially compensated by a gain in K120-solvent electrostatic interactions in the 248Q mutant complex relative to the wild-type complex; hence, K120 incurred a smaller desolvation penalty upon binding

TABLE 2: Changes in the DNA-Binding Free Energy Components of the Mutant 248W p53-CD Relative to Wild-type

interface residue i	$\Delta\Delta'G_i$	$\Delta\Delta'G_{i,\mathrm{gas}}^{\mathrm{elec}}$	$\Delta\Delta'G_{i,\mathrm{solv}}^{\mathrm{elec}}$	$\Delta\Delta'G_{i,\mathrm{gas}}^{\mathrm{vdw}}$	$\Delta\Delta'G_{i,\mathrm{solv}}^{\mathrm{nonel}}$
L1: A119	4.7	2.1	1.8	0.8	0.1
L1: K120	2.2	−3.5	4.2	1.5	0.1
L1: S121	−0.3	0.3	−0.2	−0.3	−0.1
Q136	−0.2	0.8	−1.1	0.0	0.0
L137	0.1	−0.6	0.7	0.0	0.0
L3: N239	0.0	1.4	−1.4	0.0	0.0
L3: S240	0.6	−0.6	1.2	0.1	−0.0
L3: S241	3.4	3.4	−0.5	0.5	0.0
L3: C242	−1.1	**−7.9**	**6.6**	0.1	0.1
L3: M243	2.9	2.0	−1.4	2.4	0.0
L3: N247	4.4	3.9	−1.1	1.4	0.2
L3: **248W**	**9.3**	**155.6**	**−147.7**	1.0	0.3
L3: R249	1.9	5.5	−3.8	0.0	0.1
S10: V272	0.0	0.1	−0.1	0.0	0.0
S10: R273	−4.3	−2.5	−1.6	−0.2	0.0
S10: V274	−0.4	−1.4	0.9	0.0	0.1
C275	−0.8	−1.4	0.7	−0.0	0.0
A276	0.9	1.8	−1.0	0.1	−0.0
C277	0.5	1.3	−0.8	−0.1	0.0
H2: **R280**	**−9.6**	**−11.1**	0.8	0.7	−0.0
H2: **D281**	**12.0**	**11.3**	1.0	−0.4	−0.0
H2: R283	4.7	−1.4	**6.0**	0.1	−0.1
$\sum_i$[b]	**30.9**	**159.2**	**−136.7**	**7.8**	0.7

[a] Values greater than 6 kcal/mol are highlighted in bold. [b] Free energy contributions from all interface residues.

DNA by the 248Q mutant compared to the wild-type protein ($\Delta G_{\mathrm{K120,solv}}^{\mathrm{elec}}$ = 177 vs 191 kcal/mol, Table 3).

The loss of the K120 N···Thy 6 "close" contact (see above) is compensated by the introduction of two hydrogen bonds from the serine side chain to Thy 6 in the mutant complex (the S121 O^{γ} to Thy 6 O4′ and Thy 6 O5′ distances decreased from 9.44 ± 0.48 and 7.42 ± 0.51 Å to 3.16 ± 0.31 and 3.04 ± 0.19 Å, respectively, Figure 6); hence, $\Delta\Delta G_{\mathrm{S121,gas}}^{\mathrm{elec}}$ is negative (−6.9 kcal/mol). Like R273 and R280, S121 contributes more favorably to DNA binding in the 248Q mutant than in the wild-type protein ($\Delta\Delta G_{\mathrm{S121}}$ = −12 kcal/mol, Table 1). However, the formation of two hydrogen bonds from S121 to a single base (Thy 6) in the DNA major groove is not sufficient to compensate for the loss of hydrogen bonds from K120 to three bases (Gua 7, Gua 8, and Gua 9).

The observed changes in the hydrogen bonding interactions and interaction energies indicate that mutating R248 → Q affects not only the interactions of the mutant side chain with DNA but also those of distant major groove-contact residues from the L1 loop (K120) and H2 helix (R280 and D281) as well as the phosphate-contact residue (R273). Thus, replacing R248 with Gln would significantly perturb binding of both DNA minor- and major-grooves.

R248 → W. As for the 248Q mutant, the 248W p53-CD is predicted not to bind to DNA as $\Delta\Delta G°$ = 30 kcal/mol, similar to the sum of the free energy contributions from the interface residues ($\sum_i\Delta\Delta'G_i$ = 31 kcal/mol, Table 2). The loss of binding free energy upon mutating R248 → W is due mainly to unfavorable solute−solute electrostatic interactions ($\sum_i\Delta\Delta'G_{i,\mathrm{gas}}^{\mathrm{elec}}$ = 159 kcal/mol) that are partially compensated by favorable solute−solvent electrostatic changes ($\sum_i\Delta\Delta'G_{i,\mathrm{solv}}^{\mathrm{elec}}$ = −137 kcal/mol). It is also due, to a much lesser extent, to unfavorable protein−protein/DNA vdW interactions ($\sum_i\Delta\Delta'G_{i,\mathrm{gas}}^{\mathrm{vdW}}$ = 7.8 kcal/mol). As for the 248Q mutant, the nonelectrostatic solvation free energy changes of the interface residues do not seem to contribute to the loss in binding affinity ($\sum_i\Delta\Delta'G_{i,\mathrm{solv}}^{\mathrm{nonel}}$ = 0.7 kcal/mol).

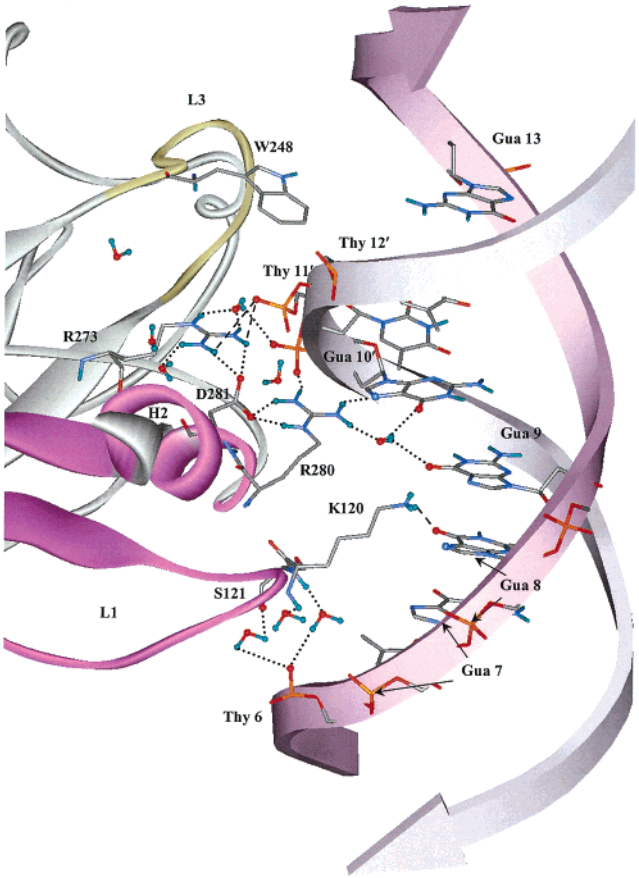

Figure 7. Hydrogen-bonding network in the average 248W p53-CD·DNA structure in the vicinity of the major and minor grooves. See also legend to Figure 5.

Table 2 shows that the key residues contributing to the loss of DNA binding upon mutating R248 → W are the mutant residue itself and D281 ($\Delta\Delta'G_i$ > 6 kcal/mol for i = 248W and D281, Table 2). Like the mutation of R248 → Q, the R248 → W substitution resulted in a loss of native protein−phosphate interactions at the mutation site (compare Figure 7 with Figure 5), hence $\Delta\Delta'G_{\mathrm{248W,gas}}^{\mathrm{elec}}$ is positive (156 kcal/mol, Table 2). The loss of 248W−DNA electrostatic interactions is partially compensated by a much smaller desolvation penalty upon binding DNA for the neutral 248W mutant residue ($\Delta G_{\mathrm{solv}}^{\mathrm{elec}}$ = 14 kcal/mol, Table 3) compared to the positively charged R248 wild-type residue, resulting in a negative $\Delta\Delta'G_{\mathrm{248W,solv}}^{\mathrm{elec}}$ (−148 kcal/mol, Table 2).

As observed in the 248Q mutant complex, to compensate for the loss of 248W−DNA electrostatic interactions Thy 11′ O^{2P} formed two new "close" contacts with the nearby R273 amino nitrogens in the 248W p53-CD·DNA simulation. Hence $\Delta\Delta'G_{\mathrm{R273,gas}}^{\mathrm{elec}}$ is negative (−2.5 kcal/mol, Table 2), but its magnitude is half of that found for the 248Q mutant in Table 1, probably because of the reduced charge−charge interactions between Thy 11′ $O^{5'}$ and R273 as the Thy 11′ $O^{5'}$···R273 N^{H2} distance (3.67 ± 0.40 Å) is longer than that in the 248Q complex simulation (3.27 ± 0.24 Å). The proximity of the R273 guanidinium side chain to Thy 11′ drew its hydrogen bonding partner, the D281 side chain, to within "close" contact of the Thy 11′ phosphate group in the 248W mutant complex (D281 $O^{\delta2}$···Thy 11′ O^{1P} = 3.91 ± 0.27 Å compared to 5.79 ± 0.52 Å in the wild-type complex). In addition, the D281 backbone amide proton and carbonyl oxygen lost two hydrogen bonds seen in the wild-type and 248Q complex

TABLE 3: Electrostatic Contributions to DNA-Binding Free Energies of Wild-type and Mutant p53-CDs

interface residue *i*	$\Delta G_{i,\mathrm{gas}}^{\mathrm{elec}}$			$\Delta G_{i,\mathrm{solv}}^{\mathrm{elec}}$ [a]			$\Delta G_{i,\mathrm{solv}}^{\mathrm{elec}}$		
	WT	248Q[b]	248W[c]	WT	248Q[b]	248W[c]	WT	248Q[b]	248W[c]
K120	−216.3	−186.1	−219.7	−100.2	−95.7	−98.7	191.0	176.5	195.1
				90.7	80.8	96.4			
S121	0.16	−6.7	0.4	−1.3	−0.15	−1.0	1.5	−0.7	1.3
				0.2	0.55	0.3			
N239	−10.48	−0.9	−9.1	−2.8	−1.6	−1.9	8.1	0.7	6.7
				5.3	−0.9	4.8			
248	−152.3	−10.7	3.3	−100.1	−5.9	−0.1	161.5	21.8	13.8
				61.4	15.9	13.7			
R249	−128.0	−138.1	−122.4	−79.1	−81.5	−72.8	145.9	155.4	142.1
				66.8	73.9	69.3			
C277	1.5	2.5	2.8	−0.4	0.5	0.1	8.1	7.1	7.3
				7.7	7.6	7.4			
R280	−172.9	−178.9	−184.0	−100.9	−97.2	−95.1	171.1	174.2	171.9
				70.2	77.0	76.8			
D281	129.9	136.1	141.2	81.7	74.8	72.5	−156.2	−147.6	−155.2
				−74.5	−72.8	−82.7			

[a] Numbers in top row are for the free state and those in the row below are for the DNA-bound state. [b] R248 → Q p53-CD. [c] R248 → W p53-CD.

simulation to the C277 carbonyl oxygen and the E285 amide proton, respectively. The additional loss of these backbone−backbone hydrogen bonds probably resulted in a more positive $\Delta\Delta' G_{\mathrm{D281,gas}}^{\mathrm{elec}}$ (11 kcal/mol, Table 2) than the $\Delta\Delta G_{\mathrm{D281,gas}}^{\mathrm{elec}}$ for the 248Q mutant (6.3 kcal/mol, Table 1).

As found in the 248Q mutant complex, the R280 amino nitrogen that is hydrogen bonded to a D281 carboxylate oxygen is also hydrogen bonded to a Gua 10′ phosphate oxygen (R280 $N^{H2}\cdots$Gua 10′ O^{1P} decreased from 3.80 ± 0.74 Å in the wild-type complex to 2.82 ± 0.18 Å), at the expense of the other R280 amino nitrogen losing a hydrogen bond with a Gua 10′ base oxygen (R280 $N^{H1}\cdots$Gua 10′ O^{6} increased from 3.53 ± 0.50 to 4.22 ± 0.58 Å, compare Figure 7 with Figure 5). The substitution of an R280-base charge-dipole interaction with an R280-phosphate charge−charge salt-bridge resulted in a negative $\Delta\Delta' G_{\mathrm{R280,gas}}^{\mathrm{elec}}$ (−11 kcal/mol, Table 2) that almost canceled the positive $\Delta\Delta' G_{\mathrm{D281,gas}}^{\mathrm{elec}}$ (11 kcal/mol), as found in the 248Q mutant (see above and Table 1).

In analogy to the 248Q mutant complex, the interaction changes in the Thy 11′···R273···D281···R280···Gua 10′ network correlated with a distortion in the major groove conformation, as evidenced by the different behavior of the RMSDs as a function of major groove base pairs (6−10) in the 248W mutant complex simulation and wild-type X-ray structure (compare dashed and solid lines in Figure 4). In contrast to the 248Q mutant complex, however, the Gua 9 O^6 and Gua 8 N^7 are ≤3.0 Å to the K120 side chain nitrogen, but their distances to the K120 ammonium protons (>2.5 Å) indicate an absence of hydrogen bonds. The Gua 7 $O^{1P}\cdots$K120 N hydrogen bond and the Thy 6 $O^{3'}\cdots$K120 N "close" contact found in the wild-type complex simulation are also lost in the 248W mutant complex simulation (compare Figure 7 with Figure 5). Although K120 contributes unfavorably to DNA binding upon mutating R248 → Trp in the wild-type protein, its relatively small $\Delta\Delta' G_{\mathrm{K120}}$ (2.2 kcal/mol) suggests that it does not play a major role in the observed DNA affinity loss as compared to its role in the 248Q mutant ($\Delta\Delta G_{\mathrm{K120}} = 19$ kcal/mol, Table 1).

The above results indicate that like the R248 → Q mutation, the R248 → W mutation affects not only the interactions of the mutant side chain with DNA but also those of distant major groove-contact residues. However, mutating R248 to Trp seems to have fewer long-range, extensive effects than mutating it to Gln (compare Figure 7 and Figure 6).

Discussion

Comparison with Available Experimental Data. *Mutant Structures.* The two mutant structures obtained from the simulations show an overall wild-type conformation despite structural changes at the protein−DNA interface. The predicted wild-type conformation is in accord with the findings that the 248Q mutant is 86% folded at 37 °C[21] and the R248 → W mutant yielded a stable wild-type-like core domain upon proteolytic digestion.[19]

Although 3d structures of the 248Q and 248W p53-CD mutants in either the DNA-bound or free state are unavailable, we can employ antibody binding data to verify the average mutant structures generated by the simulations. The wild-type p53 and the two mutants are recognized by the monoclonal antibody PAb1620 but not by the monoclonal antibody PAb240.[18,52] PAb1620 is specific for p53 native-like conformations and recognizes, in particular, R209 and N210.[52] These two residues are solvent exposed in the X-ray structure as well as in the free simulations reported here with SASA values greater than 44% (Table 4), indicating that they would be accessible to PAb1620. PAb240, however, recognizes a denatured conformation of the p53-CD that exposes its epitope, residues 212−217, located in the S7 β-strand.[53] In the wild-type p53 X-ray structure, this region is folded against the hydrophobic core of the β-sandwich and is not solvent exposed (average SASA of residues 212−217 is 14%). The latter is comparable to the average SASA values for these residues in the wild-type (20%), 248Q (17%), and 248W (22%) p53-CD simulations, indicating that they are relatively buried and would therefore be inaccessible to PAb240. Thus, in the mutant simulation structures R209 and N210 can bind to PAb1620, whereas residues 212−217 cannot bind to PAb240, in accord with the experimental antibody binding results.

NMR and equilibrium studies[21,54] of the *free* 248Q p53-CD showed that 248Q is an "intermediate mutant-type"[55] with an altered interface structure but an overall wild-type conformation. Chemical-shift changes of the 248Q p53-CD relative to the wild-type molecule were seen not only for residues in the vicinity of the mutation site but also for residues far from the mutation site, in helix H2 (containing R280 and D281), loop L1 (containing K120), loop L2, strand 2, and strand 2′.[54] This is consistent with the change in the hydrogen-bonding network involving these residues in the simulation of the DNA-free 248Q p53-CD relative to the wild-type protein. Upon mutating

TABLE 4: % SASA in Wild-type and Mutant p53-CD, Free and Bound to DNA[a]

interface residue *i*	WT[b]	248Q[c]	248W[d]	WT + DNA[b]	248Q[c] + DNA	248W[d] + DNA
R156	(21.0) 33.0	26.0	21.3	(18.4) 25.2	27.8	22.7
L206	(25.6) 18.2	27.3	22.7	(21.3) 49.7	23.3	22.9
R209	(60.8) 73.1	64.7	65.0	(61.3) 65.8	68.3	66.2
N210	(46.1) 50.5	44.1	50.5	(52.3) 55.9	51.2	51.4
212−217	(14.3) 20.4	16.6	21.6	(12.1) 18.8	19.1	19.2
K120	(55.3) 49.5	38.3	48.2	(5.9) 13.8	28.6	22.3
S121	(49.7) 57.7	40.8	34.3	(20.3) 23.3	45.6	38.6
V122	(11.6) 19.1	8.3	18.6	(13.1) 12.6	10.1	11.7
248	(52.6) 67.0	48.6	56.4	(21.9) 42.4	45.0	44.4
R249	(13.3) 21.1	21.6	23.8	(13.5) 22.0	22.4	23.3
R273	(23.1) 22.0	28.1	22.7	(13.2) 18.6	13.3	14.8
C277	(21.8) 32.2	33.1	33.5	(3.1) 6.6	23.3	6.3
R280	(41.7) 38.2	47.1	41.7	(10.9) 19.1	18.5	20.4
D281	(3.5) 11.4	19.3	7.5	(8.9) 8.3	9.6	7.8
R283	(39.6) 37.5	38.0	39.1	(31.4) 40.3	43.0	42.8
T284	(22.8) 31.6	42.8	34.6	(27.3) 42.1	44.4	48.0
E285	(19.1) 34.2	35.3	24.7	(21.9) 31.3	25.7	35.3

[a] Percentage ratio of the water-accessible surface area of the side chain X in the protein to the accessible surface area of X in the tripeptide −Gly−X−Gly− where <20% indicates buried, between 20% and 50% indicates partially buried, and >50% indicates solvent exposed.[61] [b] Numbers with and without brackets are based on the X-ray and average MD structures, respectively. [c] 248Q p53-CD. [d] 248W p53-CD.

R248 → Q in the free p53-CD, R249, located next to the mutation site, lost a backbone−backbone hydrogen bond with M246; D281 in helix H2 lost its salt-bridge with R280 and backbone−backbone hydrogen bonds with C277 O and E285 N; G117 in loop L1 lost its hydrogen bond with the T125 side chain. New hydrogen bonds linking helix H2 with β-strand S2 (via the R282 and T125 side chains) and with the L1 loop (via the G279 and K120 backbones) were formed during the 248Q p53-CD simulation.

DNA Binding Affinity. Sequence-specific DNA binding is essential for the p53 tumor suppression function.[3,56,57] The DNA-binding site for p53 contains a repeat of two consensus pentamer sequences: 5′-Pu-Pu-Pu-**C**-(A/T) (T/A)-**G**-Py-Py-Py-3′ (where Pu represents a purine (Gua/Ade), Py denotes a pyrimidine (Cyt/Thy), and invariant bases are in bold).[58] For the p53-CD and the DNA sequence studied here, DNA-binding specificity stems from the K120, C277, and R280 side chains hydrogen bonding to the Gua 8, Gua 9, Cyt 9′, and Gua 10′ base oxygen or nitrogen in the major groove, although K120 and C277 could also hydrogen bond to Ade 8 and Thy 9′, respectively. Nonspecific binding affinity is due to K120, S241, R273, A276, and R283 forming salt-bridges with the DNA phosphate.

The binding affinity of p53 to a DNA consensus site depends on the precise p53 and DNA sequences as well as on the temperature. For instance, p53 with the last 30 amino acids removed or masked by an antibody such as PAb421 has a much higher DNA binding affinity than full length p53,[56,59] and mutant p53-CDs often have higher DNA-binding affinity at 25 °C than at 37 °C.[3,55,56] Consequently, the predicted change in DNA binding affinity upon R248 → Q or R248 → W substitution was compared with experimental data corresponding to the p53 core-domain (if available) binding to a site similar to the consensus site studied here, ^{7}GGG**C**A^{11} ^{12}A**G**TCT^{16} at around 25°C. The 248Q p53-CD mutant failed to bind to the gadd45 oligonucelotide (GAA**C**A T**G**TCT, where bases that differ from the consensus DNA sequence studied here are underlined) at 20 °C.[21] Although data is not available for the binding affinity of the 248W p53-CD, the full-length mutant, in the absence or presence of monoclonal antibody PAb421, which binds to the p53 C-terminus, showed no detectable binding to the high-affinity DNA site, GGG**C**A TGTCC, at 25°C.[55,56,60] Furthermore, a peptide derived from the C-terminus of full-length p53 that included the binding site for monoclonal antibody PAb421 restored sequence-specific DNA binding to various degrees for a subset of p53 mutants, but not for the 248W mutant.[57] The nondetectable binding of the 248Q p53-CD or 248W full-length p53 mutant is consistent with the large positive free energy change computed for the 248Q and 248W mutants (≥30 kcal/mol).

Factors Affecting Loss of DNA Binding upon Mutating of R248 → Q in p53. In previous work it was believed that the DNA binding loss of 248 mutants was due to the loss of critical minor groove DNA contacts from the positively charged R248 side chain.[15] The results here show that indeed the 248Q mutant has lost two native close contacts with the DNA phosphate groups in the minor groove, but this is not the only effect the mutation has on DNA binding. The R248 → Q substitution also leads to significant structural changes in regions important for DNA binding distant from the mutation site, in accord with chemical shift observations (see above).[54] In particular, one of the structural changes involves a loss of critical major groove DNA interactions from K120, which is predicted here to contribute more to the observed loss of binding affinity than the mutant 248Q residue ($\Delta\Delta G_{K120}$ = 19 kcal/mol vs $\Delta\Delta G_{248Q}$ = 6 kcal/mol, Table 1).

The net DNA-binding free energy contribution from the three loop L1 interface residues (A119, K120 and S121) in the 248Q p53-CD is more unfavorable than that in the wild-type protein ($\Delta\Delta G_{A119} + \Delta\Delta G_{K120} + \Delta\Delta G_{S121}$ = 12 kcal/mol, Table 1). The motion of 248Q appears to be anticorrelated with that of the nearby R273: as 248Q moved away from the DNA phosphate, R273 moved closer to the Thy 11′ phosphate to compensate for the loss of 248Q-DNA phosphate contacts. Because R273 is salt-bridged to D281, which in turn is linked to R280, the proximity of R273 to the Thy 11′ phosphate drew the negatively charged D281 and the positively charged R280 side chains closer to the negatively charged DNA phosphate, resulting in a positive $\Delta\Delta' G_{D281}$ and a negative $\Delta\Delta' G_{R280}$, respectively. The net DNA-binding free energy contribution from the three hydrogen-bonded interface residues (R273, D281, and R280) in the 248Q p53 mutant is more unfavorable than that in the wild-type protein ($\Delta\Delta G_{R273} + \Delta\Delta G_{D281} + \Delta\Delta G_{R280}$ = 6 kcal/mol, Table 1).

Thus, the loss of DNA binding affinity upon mutating R248 → Q in the wild-type p53-CD is due mainly to the loss of major groove contacts from K120 in loop L1 as well as to

unfavorable interactions of D281 in helix H2. It is also due, to a lesser extent, to the loss of minor groove contacts from the mutant residue itself.

Factors Affecting Loss of DNA Binding upon Mutating of R248 → W in p53. As for the R248 → Q substitution, the mutation site has lost contact with the DNA minor groove, resulting in an unfavorable DNA-binding free energy contribution ($\Delta\Delta'G_{248W} = 9$ kcal/mol, Table 2). The net binding free energy contribution from the three loop L1 interface residues in the 248W p53 mutant is unfavorable compared to that in the wild-type protein ($\Delta\Delta'G_{A119} + \Delta\Delta'G_{K120} + \Delta\Delta'G_{S121} = 6.6$ kcal/mol), but favorable relative to that in the 248Q mutant (by 5.5 kcal/mol, see above). This is probably because, unlike the 248Q mutant, the K120 side chain nitrogen remains within hydrogen bonding distance of the Gua 8 and Gua 9 base oxygens, even though it does not form hydrogen bonds seen in the wild-type complex.

Again in common with the R248 → Q mutation, to compensate for the loss of 248W-DNA phosphate contacts, R273 moved closer to the Thy 11′ phosphate, dragging with it the salt-bridged D281, which in turn drew its hydrogen-bonding partner, R280, to interact with the Gua 10′ phosphate instead of the base oxygen. Consequently, the positively charged R273 and R280 contribute favorably, but the negatively charged D281 contributes unfavorably to binding the negatively charged DNA upon mutating R248 → W. The net binding free energy contribution from these three hydrogen-bonded residues, R273, D281, and R280, in the 248W p53 mutant is comparable to that in wild-type p53 ($\Delta\Delta'G_{R273} + \Delta\Delta'G_{D281} + \Delta\Delta'G_{R280} = -1.9$ kcal/mol) and more favorable than that in the 248Q mutant (by 8.3 kcal/mol, see above).

Thus, the loss of DNA binding affinity upon mutating R248 → W in the wild-type p53-CD is due mainly to the loss of minor groove contacts from the mutant residue itself. It is also due, to a lesser extent, to the loss of major groove contacts from loop L1 as well as to the poorer packing of the protein−DNA interface in the 248W mutant complex relative to the wild-type one ($\Delta\Delta'G_{i,gas}^{vdW} = 7.8$ kcal/mol).

Implications for Structure-Based Linker Rescue. The mutant structures generated in this study can be used with molecular docking methods to screen databases of small molecule compounds to identify potential ligands that may restore the binding affinity of mutant p53-CD toward DNA. The energetic and structural analyses suggest plausible sites that are not necessary at the mutation site to dock such ligands. They show that the major groove binding residue, K120, is a key contributor to the loss of DNA binding upon mutating R248 → Q in the wild-type p53-CD. Furthermore, experimental studies on the rescue of p53 mutants by second-site mutants found that the additional mutation of T284 → R in the DNA major groove enhanced binding of the 248Q full-length p53 mutant, with the last 30 residues deleted, to the high-affinity DNA site, GGG**C**A TGTCC, and rescued the tumor suppressor functions.[16] This together with the present results suggest that it may be possible to rescue the DNA binding affinity of the 248Q mutant using an appropriate small ligand in the vicinity of the 284 site, instead of the mutation site at position 248 to provide a bridge to the DNA. This would bring the DNA closer to the protein and force the removal of competing water molecules from the protein−DNA interface, which may allow K120 to reestablish its native major groove contacts. Such a rescue mechanism was found in our previous work, in which mutating T284 → R in the mutant 273H p53-CD did not play a role by acting simply in place of R280 to bind to the DNA major groove. Instead, it helped to restore a "native-like" conformation of the protein−DNA interface, thus enabling R280 and K120 to regain their major groove contacts.[22]

However, for the 248W p53-CD, the mutant residue itself is the major contributor to the loss of DNA binding upon mutating R248 → W in the wild-type protein. In this case it may be possible to rescue the DNA binding affinity of the 248W mutant using an appropriate small ligand in the vicinity of the mutation site itself to provide a bridge to the Thy 11′ phosphate. Apart from the hope of regaining favorable interactions with the Thy 11′ phosphate, this may inhibit Thy 11′ phosphate from contacting the nearby R273, which in turn would help to maintain the native network of interactions among Thy 11′, R273, D281, R280, and Gua 10′, and thus a "native-like" conformation of the protein−DNA interface. We are carrying out calculations to dock ligands at the proposed sites of the 248Q and 248W p53 mutants.

Acknowledgment. We thank Drs. C. S. Babu, and Y.-S. Lin for helpful discussions. We are grateful to Professor M. Karplus for the CHARMM program. S.N. is supported by a postdoctoral fellowship from Academia Sinica. This work is supported by grant # 91-2311-B-001-063 from the National Science Council, the Institute of Biomedical Sciences at Academia Sinica, and the National Center for High Performance Computing in the Republic of China.

References and Notes

(1) Harper, J. W.; Adami, G. R.; Wei, N.; Keyomarsi, K.; Elledge, S. J. *Cell* **1993**, *75*, 805−816.
(2) El-Deiry, W. S.; Tokino, T.; Velculescu, V. E.; Levy, D. B.; Parsons, R.; Trent, J. M.; Lin, D.; Mercer, W. E.; Kinzler, K. W.; Vogelstein, B. *Cell* **1993**, *75*, 817−825.
(3) Pietenpol, J. A.; Tokino, T.; Thiagalingam, S.; El-Deiry, W. S.; Kinzler, K. W.; Vogelstein, B. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1998−2002.
(4) May, P.; May, E. *Oncogene* **1999**, *18*, 7621−7636.
(5) Wu, W.; Levine, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3602−3606.
(6) Ko, L. J.; Prives, C. *Genes Dev.* **1996**, *10*, 1054−1072.
(7) Levine, A. *Cell* **1997**, *88*, 323−331.
(8) Hollstein, M.; Rice, K.; Soussi, T.; Fuchs, R.; Sorlie, T.; Hovig, E.; Smith-Sorensen, B.; Montesano, R.; Harris, C. C. *Nucl. Acids Res.* **1994**, *22*, 3551−3555.
(9) Bullock, A. N.; Henckel, J.; DeDecker, B. S.; Johnson, C. M.; Nikolova, P. V.; Proctor, M. R.; Lane, D. P.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14338−14342.
(10) Fisher, D. E. *Cell* **1994**, *78*, 539−542.
(11) Morgan, S. E.; Kim, R.; Wang, P. C.; Bhat, U. G.; Kusumoto, H.; Lu, T.; Beck, W. T. *Oncogene* **2000**, *19*, 5010−5019.
(12) Wang, Y.; Prives, C. *Nature* **1995**, *376*, 88−91.
(13) Hussain, S. P.; Harris, C. C. *Cancer Research* **1998**, *58*, 4023−4037.
(14) Hollstein, M.; Sidransky, D.; Vogelstein, B.; Harris, C. C. *Science* **1991**, *253*, 49−53.
(15) Cho, Y.; Gorina, S.; Jeffrey, P. D.; Pavletich, N. P. *Science* **1994**, *265*, 346−355.
(16) Wieczorek, A. M.; Waterman, J. L.; Waterman, M. J.; Halazonetis, T. D. *Nat. Med.* **1996**, *2*, 1143−1146.
(17) Selivanova, G.; Kawasaki, T.; Ryabchenko, L.; Wiman, K. G. *Seminars Cancer Biol.* **1998**, *8*, 369−378.
(18) Ory, K.; Legros, Y.; Auguin, C.; Soussi, T. *EMBO J.* **1994**, *13*, 3496−3504.
(19) Bargonetti, J.; Manfredi, J.; Chen, X.; Marshak, D. R.; Prives, C. *Genes Dev.* **1993**, *7*, 2565−2574.
(20) Park, J. D.; Nakamura, H.; Chumakov, A. M.; Said, J. W.; Miller, C. W.; Chen, D. L.; Koeffler, H. P. *Oncogene* **1994**, *9*, 1899−1906.
(21) Bullock, A. N.; Henckel, J.; Fersht, A. R. *Oncogene* **2000**, *19*, 1245−1256.
(22) Wright, J. D.; Noskov, S. Y.; Lim, C. *Nucl. Acids Res.* **2002**, *30*, 1563−1574.
(23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(24) Dunbrack, R. L. J.; Karplus, M. *Nat. Struct. Biol.* **1994**, *1*, 334−340.

(25) Bower, M. J.; Cohen, F. E.; Dunbrack, R. L. *J. Mol. Biol.* **1997**, *267*, 1268−1282.
(26) Wright, J. D.; Lim, C. *Prot. Eng.* **2001**, *14*, 479−486.
(27) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(28) MacKerell, J. A.; Wiorkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.
(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−923.
(30) Berkowitz, M.; McCammon, J. A. *Chem. Phys. Lett.* **1982**, *90*, 215−217.
(31) Eriksson, M. A. L.; Hard, T.; Nilsson, L. *Biophys. J.* **1995**, *68*, 402−426.
(32) Lee, B.; Richards, F. M. *J. Mol. Biol* **1971**, *55*, 379−400.
(33) Pascual-Ahuir, J. L.; Silla, E. *J. Comput. Chem.* **1990**, *11*, 1047−1060.
(34) Koradi, R.; Billeter, M.; Wuthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.
(35) Nifosi, R.; Reyes, C. M.; Kollman, P. A. *Nucl. Acids Res.* **2000**, *28*, 4944−4955.
(36) Wang, W.; Lim, W. A.; Jakalian, A.; Wang, J.; Wang, J.; Luo, R.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **2001**, *123*, 3986−3994.
(37) Boresch, S.; Archontis, G.; Karplus, M. *Proteins: Struct. Funct. Genet.* **1994**, *20*, 25−33.
(38) Boresch, S.; Karplus, M. *J. Mol. Biol.* **1995**, *254*, 801−807.
(39) Brady, G. P.; Sharp, K. A. *J. Mol. Biol.* **1995**, *254*, 77−85.
(40) Brady, G. P.; Szabo, A.; Sharp, K. A. *J. Mol. Biol.* **1996**, *263*, 123−125.
(41) Gilson, M. K.; Given, J. A.; Bush, B.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.
(42) Noskov, S.; Lim, C. *Biophys. J.* **2001**, *81*, 737−750.
(43) MacKerell, J. A. D.; Bashford, D.; Bellott, M.; Dunbrack, R.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E. I.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(44) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem.* **1997**, *101*, 3005−3014.
(45) Gilson, M. K.; Honig, B. *Proteins: Struct. Func. Genet.* **1988**, *4*, 7−18.
(46) Sharp, K. *Biophys. Chem.* **1996**, *61*, 37−49.
(47) Billeter, M. *Quart. Rev. Biophys.* **1992**, *25*, 325−377.
(48) Chandrasekhar, I.; Clore, G. M.; Szabo, A.; Gronenborn, A. M.; Brooks, B. R. *J. Mol. Biol.* **1992**, *226*, 239−250.
(49) Daggett, V.; Levitt, M. *J. Mol. Biol.* **1993**, *232*, 600−619.
(50) Smith, P. E.; van Schaik, R. C.; Szyperski, T.; Wuthrich, K.; van Gunsteren, W. F. *J. Mol. Biol.* **1995**, *246*, 356−365.
(51) Philippopoulos, M.; Lim, C. *J. Mol. Biol.* **1995**, *254*, 771−792.
(52) Wang, P. L.; Sait, F.; Winter, G. *Oncogene* **2001**, *20*, 2318−2324.
(53) Gannon, J. V.; Greaves, R.; Lane, D. P. *EMBO J.* **1990**, *9*, 1595−1602.
(54) Wong, K.-B.; DeDecker, B. S.; Freund, M. V.; Proctor, M. R.; Bycroft, M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8438−8442.
(55) Rolley, N.; Butcher, S.; Milner, J. *Oncogene* **1995**, *11*, 763−770.
(56) Friedlander, P.; Legros, Y.; Soussi, T.; Prives, C. *J. Biol. Chem.* **1996**, *271*, 25468−25478.
(57) Abarzua, P.; LoSardo, J. E.; Gubler, M. L.; Lu, Y.-A.; Felix, A.; Neri, A. *Oncogene* **1996**, *13*, 2477−2482.
(58) El-Deiry, A. A.; Kern, S. E.; Pietenpol, J. A.; Kinzler, K. W.; Vogelstein, B. *Nature Genet.* **1992**, *1*, 45−49.
(59) Hupp, T. R.; Meek, D. W.; Midgley, C. A.; Lane, D. P. *Cell* **1992**, *71*, 875−886.
(60) Zhang, W.; Funk, W. D.; Wright, W. E.; Shay, J. W.; Deisseroth, A. B. *Oncogene* **1993**, *8*, 2555−2559.
(61) Gromiha, M. M.; Oobatake, M.; Kono, H.; Uedaira, H.; Sarai, A. *Prot. Eng.* **1999**, *12*, 549−555.