Mechanism of DNA–binding loss upon single-point mutation in p53

Jon D Wright[1] and Carmay Lim[1,2,*]

[1]*Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan, ROC*
[2]*Department of Chemistry, National Tsing Hua University, Hsinchu 300, Taiwan, ROC*

**Corresponding author (Fax, +886 2 2788 7641; Email, carmay@gate.sinica.edu.tw)*

Over 50% of all human cancers involve p53 mutations, which occur mostly in the sequence−specific DNA−binding central domain (p53c), yielding little/non–detectable affinity to the DNA consensus site. Despite our current understanding of protein−DNA recognition, the mechanism(s) underlying the loss in protein−DNA binding affinity/specificity upon single−point mutation are not well understood. Our goal is to identify the common factors governing the DNA−binding loss of p53c upon substitution of Arg 273 to His or Cys, which are abundant in human tumours. By computing the free energies of wild–type and mutant p53c binding to DNA and decomposing them into contributions from individual residues, the DNA−binding loss upon charge/noncharge–conserving mutation of Arg 273 was attributed not only to the loss of DNA phosphate contacts, but also to longer–range structural changes caused by the loss of the Asp 281 salt–bridge. The results herein and in previous works suggest that Asp 281 plays a critical role in the sequence−specific DNA−binding function of p53c by (i) orienting Arg 273 and Arg 280 in an optimal position to interact with the phosphate and base groups of the consensus DNA, respectively, and (ii) helping to maintain the proper DNA–binding protein conformation.

[Wright J D and Lim C 2007 Mechanism of DNA-binding loss upon single-point mutation in p53; *J. Biosci.* **32** 827–839]

1. Introduction

Statistical analyses of the protein−DNA complexes in the Nucleic Acid Database (Berman *et al* 1992) have elucidated the key interactions governing protein−DNA binding affinity and specificity (Jayaram and Jain 2004). Charge–charge interactions between positively charged amino acid (aa) side chains and the negatively charged DNA phosphate backbone govern binding affinity, whereas hydrogen–bonding interactions between polar aa side chains and nucleic acid bases dictate binding specificity. Amino acids containing both donor and acceptor atoms tend to contact DNA using their donor atoms (Mandel-Gutfreund and Margalit 1998). Despite our current understanding of protein−DNA recognition, the mechanism(s) underlying the loss in protein−DNA binding affinity and specificity upon single point mutation are not well understood.

The tumour suppressor protein, p53, which binds DNA and activates transcription of genes that mediate DNA damage repair and cell cycle arrest (El-Deiry *et al* 1993; Harper *et al* 1993; Pietenpol *et al* 1994; May and May 1999), provides an excellent prototype system to elucidate the different mechanisms underlying the loss in DNA binding affinity and specificity upon single point mutation because over 50% of all human cancers involve p53 mutations, which occur mostly in the sequence–specific DNA–binding central domain (residues 102–292, referred to as p53c) (Hollstein *et al* 1991). Many commonly occurring single–point mutations in the core DNA–binding domain produce a full–length p53 molecule with little or non–detectable binding affinity to the DNA consensus site. Approximately 40% of these mutations occur at just six hot spots (Arg 175, Gly 245, Arg 248, Arg 249, Arg 273 and Arg 282) (Hollstein *et al* 1994). Two of the hot–spot residues (Arg 248 and Arg 273) directly contact DNA, whereas the other four hot–spot residues (Arg 175, Gly 245, Arg 249, and Arg 282) do not contact DNA directly but stabilize the p53 DNA–binding surface (Cho *et al* 1994). Based on the p53c•DNA X–ray structure and biochemical

Keywords. p53 mutants; protein-DNA interactions; molecular dynamics simulations; free energy decomposition

Abbreviations used: aa, amino acid; MD, molecular dynamic; RMSD, root-mean-square-deviations; vdW, van der Walls

data (Cho *et al* 1994), failure of p53 contact mutants, which retain the wild–type p53c conformation (Ory *et al* 1994), to act as a sequence–specific transactivator has been attributed to loss of critical DNA contacts, while failure of p53 conformational mutants to bind DNA in a sequence–specific manner has been attributed to structural defects, ranging from small structural shifts to local/global unfolding of the core domain.

Previous studies on the hot–spot 273 site have elucidated the mechanisms of p53c inactivation in the 273H single mutant and reactivation in the 273H+284R double mutant. In the wild–type p53c•DNA X–ray structure (1TSR, chain B), Arg 273 interacts with the Thy 11′ phosphate oxygen and one of the Asp 281 carboxylate oxygen atoms, while the other Asp 281 carboxylate oxygen is hydrogen bonded to Arg 280, which in turn is hydrogen bonded to an invariant base of the consensus DNA sequence (Cho *et al* 1994). On the basis of the X–ray structures of the free superstable quadruple M133L/V203A/N239Y/N268D mutant p53c (referred to as t-p53c) and the t-p53c-273H, the observed reduction in DNA binding has been ascribed exclusively to the loss of interactions between the Arg 273 guanidinium side chain and DNA (Joerger *et al* 2005). On the other hand, our previous calculations suggest that the DNA–binding loss upon the charge–conserving mutation of Arg 273 to histidine in the p53c is not due solely to the loss of charge–charge interactions with the Thy 11′ phosphate, but also to the loss of the salt–bridge with Asp 281, resulting in the loss of DNA major groove contacts from Lys 120 and Arg 280 (Wright *et al* 2002). Remarkably, the calculations show that the wild–type conformation and the DNA–binding ability of the 273H p53c mutant can be rescued by a second site mutation of a neutral Thr 284 to a positively charged arginine, as found experimentally (Wieczorek *et al* 1996). The second site mutation of Thr 284 to arginine reactivates p53 mainly by stabilizing the local DNA–binding interface and promoting the restoration of the Lys 120 and Arg 280 interactions with the DNA major groove, rather than simply adding favorable charge–charge interactions between the positively charged 284R mutant residue and the negatively charged DNA phosphate backbone.

To extend our understanding of the effect of mutation on the binding affinity of proteins to DNA we have here studied the *non*charge–conserving mutation of Arg 273 to cysteine because it is the second–most commonly seen mutation at the 273 position after the mutation to histidine (Soussi and May 1996). Furthermore, the 273C mutant p53c retains the wild–type p53c conformation: at 30°C, it binds to the PAb1620 antibody, which recognizes wild–type p53, but not to the PAb240 antibody, which is specific for denatured p53. However, the 273C mutant p53c does not exhibit specific DNA binding at 30°C (Rolley *et al* 1995). As for the 273H p53 mutant, a second site mutation of Thr 284 to arginine has been observed to restore the sequence–specific binding affinity of the 273C single mutant; i.e. both the 273H+284R and 273C+284R p53 double mutants could bind to a specific DNA sequence (Wieczorek *et al* 1996). Also, in common with the t-p53c-273H pentamutant, the observed reduction in DNA binding in t-p53c-273C pentamutant has also been attributed simply to the removal of DNA contact made by the wild–type Arg 273 on the basis that the X–ray structures of the free t-p53c, t-p53c-273H, and t-p53c-273C proteins are virtually identical (Joerger *et al* 2006). However, unlike the 273H mutant, the mutation of the positively charged Arg 273 to a neutral cysteine eliminates a positive charge, therefore, it is not clear if the cause for the loss of DNA sequence–specific binding would be similar to that found for the 273H mutant. In other words, is the observed DNA–binding loss upon the noncharge–conserving mutation of Arg 273 to cysteine in the p53c due to (i) the loss of Arg 273–phosphate protein–DNA interactions, and/or (ii) the loss of Arg 273–Asp 281 protein–protein interactions, resulting in the loss of major groove contacts, and/or (iii) other factors?

The goal of this work is to answer the aforementioned question, and to identify the common factors governing the loss of DNA binding upon charge–conserving or noncharge–conserving mutations of Arg 273 in p53c. Because the p53c domain contains a tightly bound Zn^{2+} ion that is critical for its DNA–binding function, it is important to accurately model the interactions of the Zn^{2+} dication. However, current popular force fields such as the CHARMM force field, which employs Zn^{2+} van der Waals (vdW) parameters that reproduce the hydration properties of only Zn^{2+} (Stote and Karplus 1995) and the conventional 12–6 vdW and Coulombic potential energy function, could not preserve the native tetrahedral Zn–Cys_3His binding site geometry in simulations of the DNA–free/bound wild–type/273C p53c (J D Wright and C Lim, unpublished results). In contrast, the CTPOL force field, which incorporates charge transfer effects and local electrostatic polarization energy (Sakharov and Lim 2005), and employs Zn^{2+} vdW parameters that reproduce the observed hydration structure and relative hydration free energies of Zn^{2+} (Babu and Lim 2006), could reproduce the observed tetrahedral Zn–Cys_3His binding site geometry in simulations of the DNA–free/bound wild–type/mutant p53c (J D Wright and C Lim, unpublished results). Furthermore, a change in the Zn^{2+} CN from four to six in the wild–type/273C mutant p53c was found to affect the conformation and the relative mobility of secondary structural elements that are involved in binding Zn^{2+} and/or the DNA (J D Wright and C Lim, unpublished results). Hence, the CTPOL force field was employed in molecular dynamics (MD) simulations of the wild–type and 273C mutant p53c, free and DNA–bound, in the presence of explicit water molecules. Based on the solution MD structures, the free energies of wild–type and 273C mutant p53c binding to DNA were computed and decomposed into component energies and contributions

from individual residues. The theoretical results obtained were verified by their consistency with available experimental data. They help to elucidate the mechanism underlying the loss in protein−DNA binding affinity and specificity upon single point mutation at the 273 hot spot, which in turn might aid the search for suitable ligands to restore the lost interactions with DNA.

2. Methodology

2.1 *MD simulations*

The wild–type and 273C p53c domains (residues 96–289), free and bound to a 15–basepair DNA duplex, 5'–^{3}AATTGGGCAAGTCTA^{17}–3', were subjected to MD simulation at a mean temperature of 300 K and at pH 7 using the CHARMM29 (Brooks *et al* 1983) program. A full description of the simulation starting structures, force fields, protonation states, nonbonded interaction treatment, and the simulation protocol is given in our previous work (J D Wright and C Lim, unpublished results). In summary, the initial wild–type p53c^{wt}•DNA complex was taken from PDB entry 1TSR, chain B (Cho *et al* 1994), while the initial mutant p53c^{mut}•DNA complex was generated using a backbone–dependent side chain rotamer library (Dunbrack and Karplus 1993; Bower *et al* 1997). Starting structures for the DNA–free simulations were obtained by removing the DNA from the respective wild–type 1TSR, chain B structure and initial mutant complex structure. Each simulation was carried out using the CTPOL force field (*see* Sakharov and Lim 2005 and Babu and Lim 2006) for the metal–binding site, the all–hydrogen CHARMM27 parameter set (MacKerell *et al* 1998) for the protein and DNA atoms, and the TIP3P model for the water molecules (Jorgensen *et al.* 1983). The DNA–free and DNA–bound structures were immersed in the center of a previously equilibrated TIP3P water sphere of radius 42Å and 50Å, respectively. For each DNA–bound structure, twenty sodium ions were added to the system by replacing water molecules with the highest electrostatic energies on the oxygen atom and located ≥7Å apart from each other and ≥4Å apart from the protein or DNA atoms. The resulting numbers of atoms in the simulations of wild–type p53c (p53c^{wt}), mutant 273C (p53c^{mut}), p53c^{wt}•DNA, and p53c^{mut}•DNA are 32,782, 32,889, 58,389, and 58,394, respectively. Each simulation was continued until the protein backbone rmsd from the initial structure had stabilized (i.e. fluctuated less than 0.1Å) for at least 100 ps, yielding simulation periods of 550−650 ps.

The last 100 ps of each trajectory were used in generating an average structure, in analysing hydrogen bonds and vdW contacts, and in computing the B–factor and relative percentage solvent accessibility of each residue. A hydrogen bond is defined as a heavy–heavy donor–acceptor distance ≤ 3.5 Å and a heavy–light distance ≤ 2.5 Å. vdW contacts between the protein and DNA are defined as having a heavy–heavy donor–acceptor distance between 3.5 and 4.0 Å. The percentage solvent accessibility of each residue is defined as the percentage ratio of the side chain X water–accessible surface area in the protein to that in the tripeptide –Gly–X–Gly–. It was computed using the MolMol (Koradi *et al* 1996) program with a probe radius of 1.4 Å.

2.2 *Free energy calculations*

Based on scheme 1, where p = wild–type or mutant p53c,

$$\begin{array}{ccccc} & & G_{gas}(p) & & \\ [p]_{gas} & + & [DNA]_{gas} & \rightarrow & [p\bullet DNA]_{gas} \\ \uparrow -G_{solv}(p) & & \uparrow -G_{solv}(DNA) & & \downarrow G_{solv}(p\bullet DNA) \\ [p]_{sln} & + & [DNA]_{sln} & \rightarrow & [p\bullet DNA]_{sln} \\ & & G^{o}(p) & & \end{array} \quad \text{(scheme 1)}$$

the DNA–binding free energy of wild–type or mutant p53c (denoted by a single Δ) in solution is given by:

$$\Delta G^{o}(p) = \Delta G_{gas}(p) + G_{solv}(p\bullet DNA) - G_{solv}(p) - G_{solv}(DNA), \quad (1)$$

where ΔG_{gas} is the standard free energy change per mole for the noncovalent association of the p53c 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. The free energy difference between the wild–type and mutant p53c binding to DNA in solution (denoted by double $\Delta\Delta$) is given by:

$$\Delta\Delta G^{o} = \Delta G_{gas}(p53c^{mut}) - \Delta G_{gas}(p53c^{wt}) + G_{solv}(p53c^{mut} \bullet DNA) - G_{solv}(p53c^{wt} \bullet DNA) - G_{solv}(p53c^{mut}) + G_{solv}(p53c^{wt}). \quad (2)$$

The ΔG^{o} calculations were based on the 20 coordinate sets saved every 5 ps during the final 100 ps of each simulation (complexed with and without the DNA), omitting the first and last residues of each coordinate set. As the DNA–bound and free proteins had the same number of protein and DNA atoms but different numbers of water atoms, all water molecules were removed from the MD trajectories to avoid boundary problems for the different sized systems. Although the counterions from each coordinate set could be included in the ΔG^{o} calculations, differences in their locations between the wild-type and mutant complex simulations would affect the ΔG^{o} calculations. Therefore, the counterions from each coordinate set of the wild-type and mutant complex simulations were replaced by a common set of counterions, whose positions correspond to their average positions in the wild–type simulation over the 100 ps under study. To relieve any close vdW contacts among these pseudo counterions,

they were minimized for 10 steps using conjugate gradients with the protein/DNA/Zn^{2+} fixed to preserve the simulation structures. Using this common set eliminates any artifacts in the free energy decomposition that may be caused from slight differences in the counterion distribution in the simulations.

The ΔG_{gas} in eq. 1 was approximated by:

$$\Delta G_{gas} \sim \Delta E_{gas}^{vdW} + \Delta E_{gas}^{elec}, \quad (3)$$

where the changes in the gas–phase vdW and electrostatic energies upon DNA binding were calculated using the CHARMM program (Brooks *et al* 1983) and force field (MacKerell *et al.* 1998) with a dielectric constant of 1 and a nonbonded cut–off of 999 Å.

The nonelectrostatic contribution to the solvation free energy (G_{solv}^{nonel}) was approximated by a linear function of the solvent accessible surface area (SASA) (Lee and Richards 1971):

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

where γ^{cav} = 46.0 cal/mol/Å^2 (Sharp *et al.* 1991) and γ^{vdW} = –38.8 cal/mol/Å^2 (Jayaram *et al* 1999), and the absolute SASA of the free or DNA–bound protein was computed using the CHARMM program and a solvent probe radius of 1.4Å.

The electrostatic contribution to the solvation free energy (G_{solv}^{elec}) was estimated by finite–difference solution to the Poisson–Boltzmann equation (Gilson and Honig 1988), implemented in the CHARMM program, using the optimized radius parameters in reference (Nina *et al* 1997), with a protein dielectric equal to 1 and a solvent dielectric equal to 80. The interior and exterior grid points were assigned either a solvent dielectric or solute dielectric using the reentrant surface from the 1.4 Å solvent probe radius. The grid was set to 200 points per side with an initial grid spacing of 1.6 Å, followed by an intermediate grid spacing of 0.8 Å and a final spacing of 0.4 Å. The sum of G_{solv}^{nonel} and G_{solv}^{elec} yielded the solvation free energy (G_{solv}).

3. Results

3.1 *Validation of the model solution p53c structures*

To evaluate if the simulation structures conform to the observed X–ray conformations of the DNA–free (1TSR, chain C) and DNA–bound (1TSR, chain B) wild–type p53c, we computed the root–mean–square–deviations (RMSDs) of each residue's backbone heavy atoms in the average $p53c^{wt}$ or $p53c^{wt}$•DNA structure from the respective X–ray structure and evaluated if the main chain hydrogen bonds found in the X–ray structures were preserved during the simulations (J D Wright and C Lim, unpublished results). As the overall backbone RMSD in the "best" representative conformer of the free p53c NMR ensemble (2FEJ, model 1) (Canadillas *et al* 2006) from the corresponding X–ray structure (1TSR, chain C) is 1.81 Å, residues with <2 Å backbone RMSDs from the respective X–ray structure were assumed to adopt native–like solution conformations. The backbone RMSD and hydrogen bond analyses show that the free and DNA–bound $p53c^{wt}$ simulations have preserved the key protein–protein and protein–DNA contacts. For the free $p53c^{wt}$, the percentage of residues with <2 Å backbone RMSDs from the 1TSR, chain C structure in the average MD structure (86%) is greater than that in the "best" representative NMR conformer (73%). However, the percentage of all backbone–backbone hydrogen bonds found in the free X–ray crystal structure that are preserved in the simulation (~71%) is less than that preserved in the free NMR solution structure (77% for model 1 of 2FEJ). Nevertheless, it is close to the percentage of all backbone–backbone hydrogen bonds found in 60 X–ray structures that are preserved in the corresponding NMR structures (~69%) (Garbuzynskiy *et al* 2005). For the DNA–bound $p53c^{wt}$, 87% of all residues in the average simulation structure have <2 Å backbone RMSDs from the X–ray complex structure (1TSR chain B), while ~77% of the backbone–backbone hydrogen bonds found in the X–ray structure are preserved during the simulation.

The majority of the backbone–backbone hydrogen bonds that were not preserved in the free and DNA–bound $p53c^{wt}$ simulations do *not* involve residues in the loop–sheet–helix motif (aa 113–140 and 270–287), L2 loop (aa 171–181), and L3 loop (aa 237–250) that are critical for DNA and/or Zn^{2+} binding, which include the hot–spot "conformational" sites (Arg 175, Gly 245, Arg 249, and Arg 282). However, the distance between Arg 282 N and Pro 278 O (3.13 Å) in the $p53c^{wt}$•DNA X–ray structure has increased to 4.08 Å in the respective simulation. For the free wild–type p53c backbone, a large deviation from the crystal structure backbone occurs for residues 181–188, 208–213, 224–230, and 260–264, which all pertain to loops that are not involved in Zn^{2+} and/or DNA binding. Interestingly, the backbone–backbone hydrogen bonds in the free simulation structure are very similar to those in the free NMR solution structure (2FEJ, model 1) (Canadillas *et al* 2006).

3.2 *Validation of the computed B–factors*

To evaluate if the simulation structures conform to the experimentally observed protein mobility, we compared the average main chain B–factors per residue derived from the simulations of the DNA–free/bound wild–type p53c with the respective X–ray values. Comparison between the computed and X–ray B–factors in our previous work (J D Wright and C Lim, unpublished results) show that the simulations could reproduce qualitative features of the protein mobility.

Because the computed B–factors do not reflect protein motions occurring on a timescale slower than 10^{-9} s, which are inherent in the X–ray B–factors (Petsko and Ringe 1984; Philippopoulos and Lim 1995), they generally underestimate the experimental numbers. The exceptions correspond to residues 96, 105, 167, 168, 171, 243, and 244 in the free $p53c^{wt}$ and residues 225 and 226 in the DNA–bound $p53c^{wt}$, whose computed B–factors exceed the respective X–ray values. However, this may be because residues 96, 105, 167, 170, 243, and 244 in the free X–ray structure (1TSR, chain C) and residues 225 and 226 in the complex X–ray structure (1TSR, chain B) are stabilized by crystal contacts, which would artificially reduce the B–factors of these residues.

3.3 *Effect of the Arg 273 ⟶ Cys mutation on the protein/DNA conformation*

To determine the changes in the protein conformation upon mutating Arg 273 to cysteine in p53c, we computed the protein backbone RMSD in the average wild–type/mutant p53c structure from the 1TSR, chain C crystal structure for the free simulations and from the 1TSR, chain B structure for the DNA–bound simulations. The RMSD difference between the wild–type and mutant proteins in the absence and presence of DNA (figure 1) show that the Arg 273→ Cys mutation has caused conformational changes in the loop–sheet–helix motif, which is involved in DNA binding.

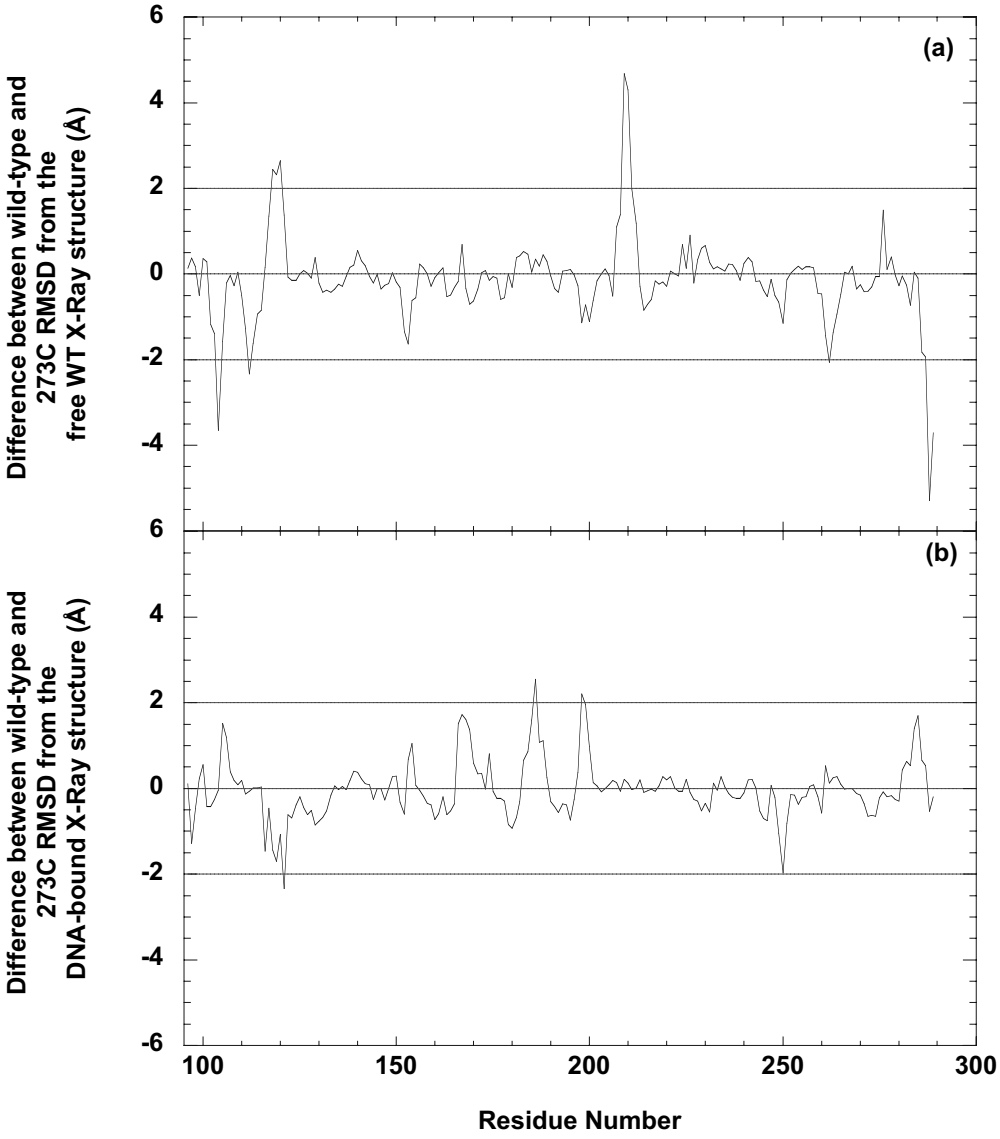

Figure 1. Plot of the protein backbone RMSD difference between the wild–type and mutant p53c proteins as a function of residue number in the (**a**) absence of DNA and (**b**) presence of DNA.

The loop–sheet–helix motif contains the L1 loop (aa 113–123), a 3–stranded β sheet consisting of the S2–S2′ hairpin (aa 124–140) and the four COOH–terminal residues of the S10 β strand (aa 271–274), and the H2 α helix (aa 278–287). Residues 117–121 and 283–287 in the free p53c as well as Lys 120 in the DNA–bound protein exhibit significant (>2 Å) backbone RMSD differences between wild–type and 273C p53c (*see* figure 1). In addition, Q106 and G112 at the N–terminus and aa 208–213 in the loop connecting β–strands S6 and S7 in the free p53c as well as D186 and E198 in the DNA–bound protein also exhibit significant (>2Å) backbone RMSD differences upon mutating Arg 273→Cys in p53c (*see* figure 1). As Lys 120 from the L1 loop hydrogen bonds to the O6 and N7 atoms of the Gua 8 base in the major groove, the significant backbone RMSD differences between the wild–type and 273C free/DNA–bound p53c indicate that the Arg 273→Cys mutation may have affected the binding of Lys 120 in the DNA major groove.

To determine the changes in the DNA conformation upon mutating Arg 273 to cysteine in p53c, we computed the DNA backbone RMSD in the average wild–type/mutant p53c structure from the corresponding 1TSR, chain B structure. Comparison of the DNA backbone RMSDs derived from the average wild–type and mutant p53c complexes (figure 2) show that the Arg 273→Cys mutation has resulted in conformational changes in part of the DNA that binds to the p53c protein. In the 1TSR, chain B crystal structure, the p53c protein binds to the consensus DNA sequence, ^{7}GGGCAAG^{13}. Compared to the DNA backbone in the X–ray structure, the RMSDs of base pairs 7–13 in the average wild–type complex is <2.5Å, which seems reasonable considering that the average B–factors of these base pairs in the X–ray structure seems relatively high (~56Å^2). However, the RMSDs of base pairs 8–12 in the average mutant complex exceed 2.5Å, and are significantly greater than the corresponding RMSDs in the average wild–type complex. These high RMSD values suggest that the backbone conformation of base pairs 8–12 in the average mutant complex differ significantly from that in the average wild–type complex.

In summary, the Arg 273→Cys mutation has caused structural changes in the protein–DNA interface, which in turn may be expected to affect the DNA binding ability of the 273C p53c mutant.

3.4 *Effect of the Arg 273→Cys mutation on the protein mobility*

To determine the changes in the protein mobility upon mutating Arg 273 to cysteine in p53c, we compared the average main chain B–factors of each aa residue derived

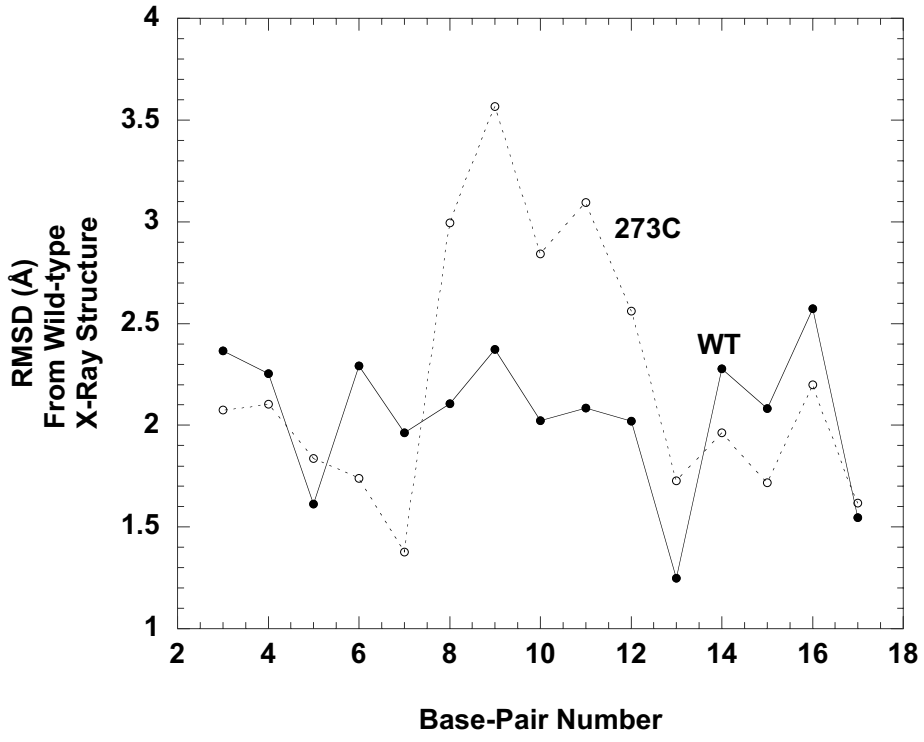

Figure 2. Plot of the RMSD of the DNA backbone atoms (averaged over the base pair) from the wild–type X–ray complex structure (1TSR, chain B) as a function of base pair number for the wild–type (solid line) and 273C mutant (dotted line) MD complex structures.

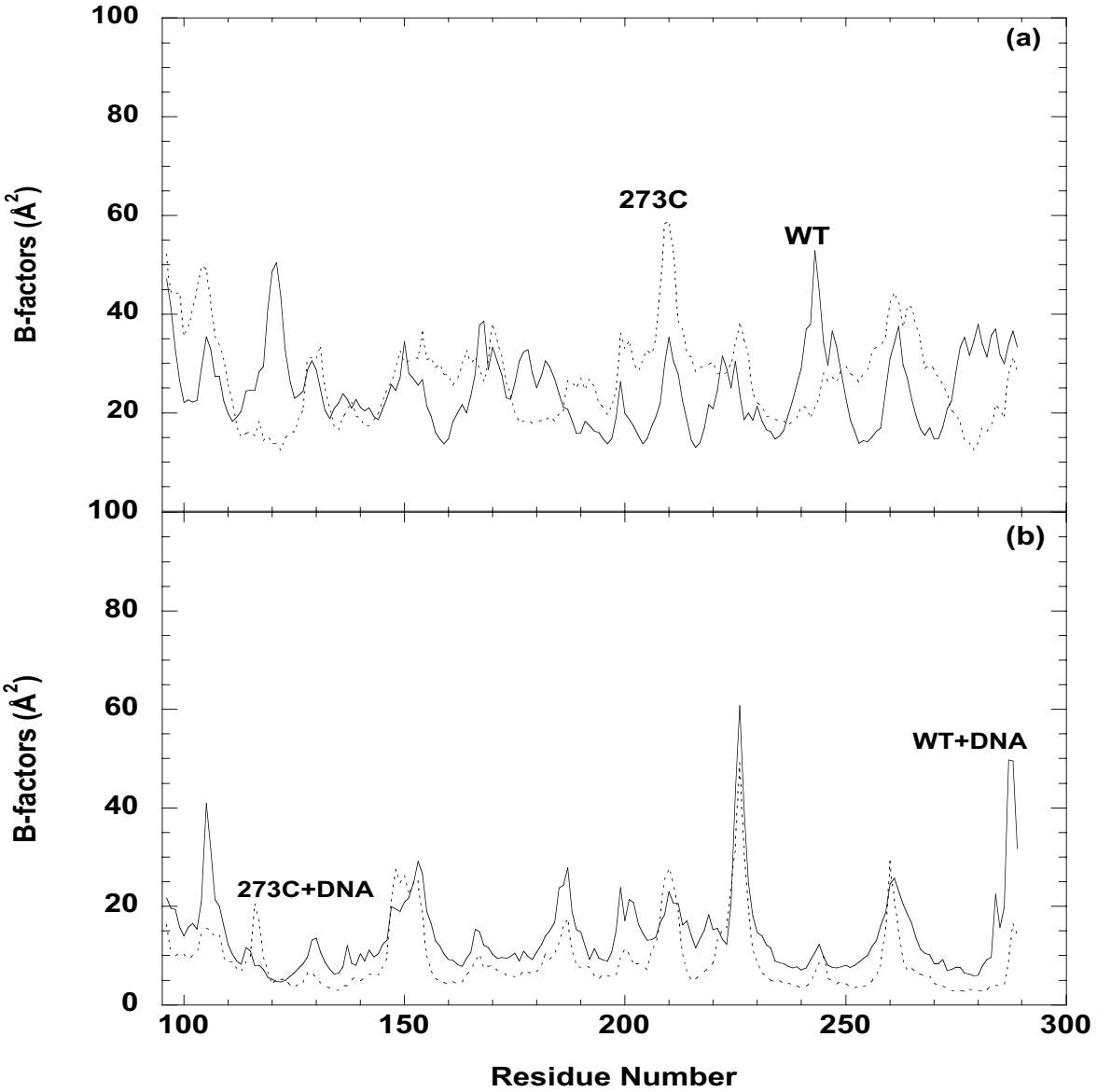

Figure 3. Plot of the B–factors averaged over the protein backbone atoms as a function of residue number in the simulations of DNA–free **(a)** and DNA–bound **(b)** p53c. The solid and dotted curves correspond to wild–type and 273C mutant p53c, respectively.

from the DNA–free/bound wild–type p53c simulation (figure 3, solid curves) with those derived from the respective mutant simulation (figure 3, dashed curves). The average main chain B–factors as a function of the residue number in figure 3 show that the Arg 273→Cys mutation seems to rigidify some of the residues that are involved in binding Zn^{2+} and/or the DNA especially in the free protein. In the absence of DNA, the main chain B–factors of the residues in the mutant p53c are generally higher than those in the wild–type protein, except for aa 113–123 in the L1 loop, aa 175–185 in the L2 loop, 240–248 in the L3 loop, and aa 276–286 in the H2 helix (figure 3a). Notably, Lys 120 from the L1 loop as well as Cys 277 and Arg 280 from the H2 helix interact with the DNA base atoms in the major groove, while Arg 248 from the L3 loop, which has extensive contacts with the L2 loop, interacts with phosphate and sugar groups in the minor groove. In the presence of DNA, the 273C p53c mutant demonstrates a similar trend in the main chain *B*–factors to the wild–type p53c, although the average main chain B–factors are generally slightly lower than the respective wild–type values (figure 3b).

3.5 *Effect of the Arg 273→Cys mutation on DNA binding*

For normal p53 tumour suppressor function, it is essential for the p53c to be able to bind to the correct specific DNA sequence in order to initiate the anti–oncogenic processes (Pietenpol *et al.* 1994; Abarzua *et al* 1996; Friedlander *et al* 1996). The DNA–binding site for p53c contains a repeat of the 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) and Py represents a pyrimidine (Cyt/Thy) (El-Deiry *et al* 1992). The consensus DNA sequence studied here corresponds to ^{7}GGGCA^{11} ^{12}AGTCT^{16}. DNA–binding specificity stems from the Lys 120, Cys 277 and Arg 280 side chains hydrogen bonding to the base oxygen and/or nitrogen atoms of bases Gua 8, Cyt 9', and Gua 10' in the major groove, while non–specific DNA–binding affinity is due to salt bridges between Lys 120, Ser 241, Arg 273, Ala 276 and Arg 283 with the DNA phosphate backbone. Experimental binding data for the full–length 273C mutant p53 at 30°C shows that it fails to bind the consensus DNA sequence, 5′–GGGCA TGTCC–3′ (where the underlined bases are different to the sequence studied here) (Rolley *et al.* 1995). To determine if the calculations could predict the observed loss of the DNA–binding ability of the 273C mutant, we computed the DNA–binding free energy of the 273C p53c relative to that of the wild–type p53c according to eq 2 (*see* § 2.2). The free energy difference between the mutant and wild–type p53c binding to DNA in solution is large and positive ($\Delta\Delta G^{o}$ = 24 kcal/mol). This indicates a loss of the DNA–binding ability of the 273C mutant, consistent with the experimental binding results.

3.6 *Key residues contributing to the DNA–binding loss of 273C p53c*

To determine the key residue(s) responsible for the loss of the DNA–binding ability of the 273C mutant p53c, the $\Delta\Delta G^{o}$ was decomposed into contributions from each residue i ($\Delta\Delta G_i$), where the $\Delta\Delta$ denotes the DNA–binding free energy of the mutant p53c relative to that of the wild–type protein. The free energy components of the interface residues, defined as residues whose atoms are within 4.0Å of any DNA atom in the X-ray structure, are listed in table 1. As expected, the mutant residue contributes less favorably to DNA binding than the wild–type Arg 273 residue ($\Delta\Delta G_{273}$ = 11 kcal/mol, table 1). This is mainly because of the loss of favorable charge–charge interactions that the wild–type Arg 273 side

Table 1. DNA-binding free energy components of interface residues in the 273C mutant p53c relative to the wild-type protein[a]

Interface residue i	$\Delta\Delta G_i$	$\Delta\Delta G^{elec}_{i,gas}$	$\Delta\Delta G^{elec}_{i,solv}$	$\Delta\Delta G^{vdw}_{i,gas}$	$\Delta\Delta G^{nonel}_{i,solv}$
L1: G117	5.90	8.3	-3.1	0.7	0.0
L1: A119	0.47	-2.1	2.8	-0.2	-0.1
L1: K120	**10.18**	**1.4**	**6.9**	**1.8**	**0.0**
L1: S121	0.87	-2.9	2.2	1.7	-0.1
L1: V122	4.08	2.0	0.1	2.1	-0.1
N239	-2.16	-0.6	-1.7	0.1	0.0
L3: S241	-3.54	-2.7	-1.6	0.7	0.1
L3: M243	-2.00	-3.0	1.3	-0.1	-0.2
L3: N247	-5.89	-6.1	1.3	-1.0	-0.1
L3: R248	**-11.58**	**-19.8**	**6.9**	**1.4**	**-0.1**
S10: 273C	**10.75**	**122.9**	**-113.8**	**1.8**	**-0.2**
C275	-1.13	1.5	-3.3	0.6	0.0
A276	-3.96	-5.6	0.6	1.1	0.0
C277	0.46	-0.7	1.9	-0.8	0.0
H2: R280	0.22	-18.7	17.7	1.1	0.0
H2: D281	**23.72**	**-12.0**	**37.7**	**-2.1**	**0.1**
H2: R282	**8.41**	**21.7**	**-13.5**	**0.2**	**0.0**
H2: R283	3.96	2.9	1.8	-0.7	0.0
H2: T284	1.28	1.2	1.0	-1.3	0.4
$\sum_i^b$	**40.0**	**87.7**	**-54.6**	**7.1**	**-0.1**

[a]Bold rows correspond to residues with $\Delta\Delta G_i$ values >8 kcal/mol. [b]Free energy contributions from all interface residues.

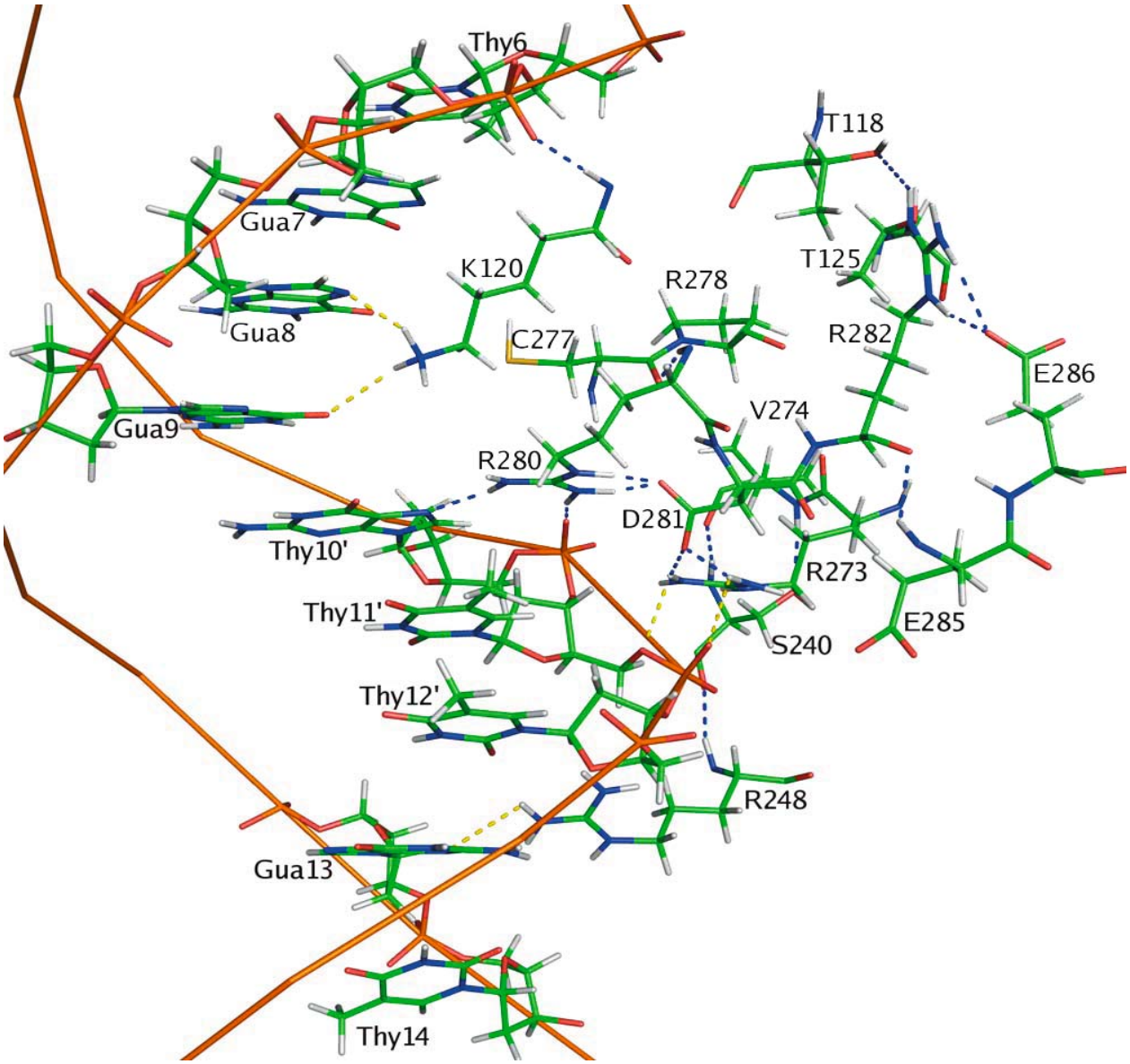

Figure 4. Hydrogen–bonding network in the average wild–type p53c·DNA simulation structure in the vicinity of the DNA major and minor grooves. Hydrogen bonds and "vdW" contacts, as defined in § 2, are depicted by blue and yellow dashed lines, respectively. Only certain key residues (discussed in § 3) and their hydrogen bonding partners are shown for the sake of clarity. Oxygen, phosphate, carbon, nitrogen, and hydrogen atoms are colored red, orange, green, blue, and white, respectively. Figures were prepared using the program PyMOL (DeLano 2004).

chain would have had with the DNA phosphate backbone ($\Delta\Delta G_{273,gas}^{elec}$ = 123 kcal/mol, compare figures 4 and 5). The unfavorable $\Delta\Delta G_{273,gas}^{elec}$, however, is compensated partly by the negligible desolvation cost of the neutral mutant residue upon DNA binding ($\Delta G_{273C,solv}^{elec}$ = ~1 kcal/mol), as compared to the large desolvation penalty of the positively charged wild–type arginine ($\Delta G_{R273,solv}^{elec}$ = 115 kcal/mol).

The unfavourable contribution of the mutant cysteine towards DNA binding ($\Delta\Delta G_{273}$ = 11 kcal/mol) is compensated by the minor groove binding residue, Arg 248 ($\Delta\Delta G_{R248}$ = –12 kcal/mol), which is linked to the mutation site by Ser 240. In the free and DNA–bound wild–type p53c, Ser 240 makes backbone–backbone hydrogen bonds to both Arg 248 and Val 274, neighbouring the mutation site (*see* figure 4). Thus, any conformational changes due to mutation at position 273 are propagated to Arg 248. The backbone–backbone hydrogen bond between the Arg 248 amide hydrogen and the Ser 240 carbonyl oxygen is retained in simulations of the free and DNA–bound 273C mutant proteins. However, the Arg 248 side chain in the 273C

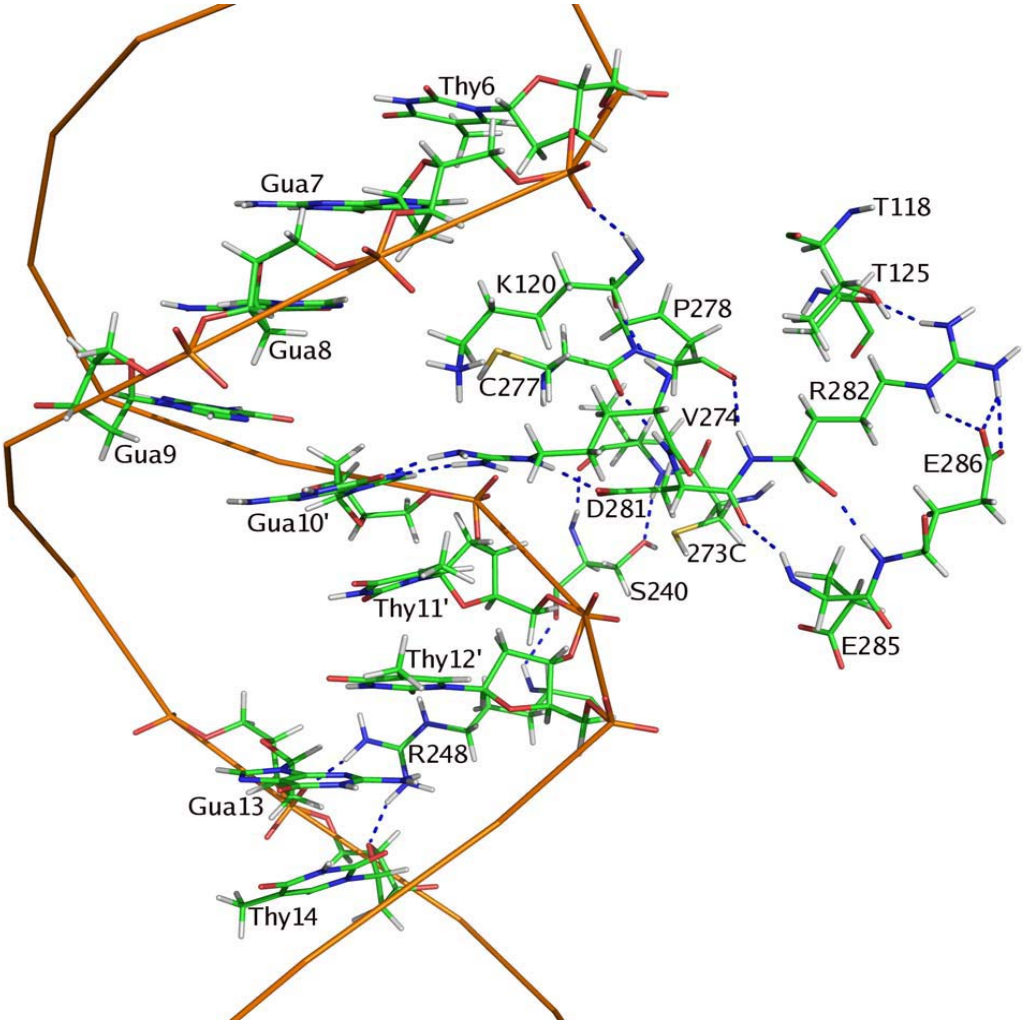

Figure 5. Hydrogen–bonding network in the average 273C p53c·DNA structure in the vicinity of the major and minor grooves. See also footnotes to figure 1.

mutant gains two hydrogen bonds with the Gua 13 sugar $O^{3'}$ and Thy 14 sugar $O^{4'}$ (*see* figure 5), which are absent in the wild–type p53c complex, where the Arg 248 side chain makes vdW contacts (rather than hydrogen bonds) with the DNA in the minor groove (*see* figure 4). This results in a gain in the Arg 248 electrostatic interactions in the 273C mutant upon DNA binding ($\Delta\Delta G_{R248,gas}^{elec}$ = –20 kcal/mol, table 1), which is partly offset by a more unfavourable solvation free energy change of the mutant p53c upon DNA binding, as compared to that of the wild–type protein ($\Delta\Delta G_{R248,solv}^{elec}$ = 7 kcal/mol). Thus, Arg 248 makes a favourable free energy contribution to the net $\Delta\Delta G^{o}$.

As the favourable contribution of Arg 248 cancels the unfavorable contribution of Cys 273, the mutant residue is neither the sole nor the key determinant of the observed DNA–binding loss of the 273C p53c: Asp 281 makes the largest unfavorable free energy contribution ($\Delta\Delta G_{D281}$ = 24 kcal/mol, table 1) to the net $\Delta\Delta G^{o}$, which is roughly twice that of the mutant residue. Surprisingly, the unfavourable free energy contribution is not because of the loss of hydrogen bonding interactions between the Asp 281 carboxylate oxygen and the mutant cysteine, as evidenced by the favorable gas–phase electrostatic energy ($\Delta\Delta G_{D281,gas}^{elec}$ = –12 kcal/mol, table 1). In the free wild–type p53c, the Asp 281 amide hydrogen and carbonyl oxygen hydrogen bond to the Cys 277 carbonyl oxygen and Glu 285 amide hydrogen, respectively, while its carboxylate side chain forms salt bridges with both the Arg 273 and Arg 280 guanidinium groups. These side chain–side chain interactions are lost, whereas the two backbone–backbone hydrogen bonds are preserved in the free 273C mutant, consistent with the X–ray structure of the free t-p53c-273C

pentamutant. Upon binding DNA, Asp 281 gains a side chain–side chain hydrogen bond with the Arg 280 $N^{\varepsilon}H$ in the 273C mutant p53c (*see* figure 5), which may account for the favorable $\Delta\Delta G_{D281,gas}^{elec}$. Rather, the highly unfavorable $\Delta\Delta G_{D281}$ contribution is due to an unfavourable solvation free energy difference between the mutant and wild–type p53c binding to DNA in solution ($\Delta\Delta G_{D281,solv}^{elec}$ = 38 kcal/mol, table 1). This may be because upon DNA binding, Asp 281 loses 3 hydrogen bonds to water molecules in the 273C mutant, but retains the same number of hydrogen bonds to water in the free and DNA– bound wild–type p53c.

Arg 282, which is next to Asp 281, also contributes less favourably to DNA binding in the 273C mutant p53c, as compared to its contribution in the wild–type protein ($\Delta\Delta G_{R282}$ = 8.4 kcal/mol, table 1). Arg 282 is linked to Asp 281 via Glu 285, whose amide nitrogen is within hydrogen bonding distance (≤3.5 Å) to both the Asp 281 and Arg 282 carbonyl oxygen atoms. In the free wild–type p53c, the Arg 282 backbone amide and carbonyl oxygen hydrogen bond to Pro 278 and Glu 286, whose neighbours, Cys 277 and Glu 285, respectively, form backbone–backbone hydrogen bonds with Asp 281 (*see* above), while the Arg 282 $N^{\varepsilon}H$ and NH1 atoms hydrogen bond to Glu 286 $O^{\varepsilon 1}$ and Thr 125 $O^{\gamma 1}$, respectively. These hydrogen–bonding interactions are preserved in the free 273C mutant, where the Arg 282 side chain additionally interacts with both Glu 286 carboxylate oxygen atoms. Upon binding DNA, Arg 282 gains two side chain–side chain hydrogen bonds with Thr 118 and Glu 286 in the wild–type p53c, whereas its hydrogen bonding interactions do not change in the 273C mutant. The gain in hydrogen bonding interactions upon binding DNA for the wild–type p53c but not for the 273C mutant probably accounts for the unfavorable $\Delta\Delta G_{R282,gas}^{elec}$ (22 kcal/mol, table 1) and an overall unfavorable free energy contribution to the net $\Delta\Delta G^{o}$.

In contrast to Arg 282, Arg 280, which forms a salt bridge with Asp 281, does not appear to contribute to the DNA–binding loss of the 273C mutant, and interacts with the invariant Gua 10' in both the wild–type and 273C mutant p53c (*see* figures 4 and 5). On the other hand, the neighbouring K120, which also interacts with the DNA major groove in the wild–type protein, seems to contribute less favourably contribution to DNA binding in the 273C mutant p53c, as compared to its contribution in the wild–type protein ($\Delta\Delta G_{K120}$ = ~10 kcal/mol, table 1). The unfavorable free energy contribution is mainly because upon DNA binding, the solvation free energy change of Lys 120 in the 273C mutant is more unfavorable than that in the wild–type protein, resulting in an unfavorable $\Delta\Delta G_{D281}$ contribution (of ~7 kcal/mol) to the net $\Delta\Delta G^{o}$. It is also partly due to the loss of electrostatic interactions and vdW contacts with the DNA bases in the major groove (compare figures 4 and 5).

4. Discussion

In previous work, it was thought that mutations of Arg 273 in p53c caused DNA–binding loss due solely to the loss of direct charge–charge interactions between the wild–type arginine side chain and the DNA phosphate backbone (Cho *et al* 1994; Joerger *et al* 2005, 2006). Notably, on the basis of the virtually identical X–ray structures of the free t-p53c, t-p53c-273H, and t-p53c-273C, the observed reduction in DNA binding by the t-p53c-273H and t-p53c-273C pentamutants has been attributed to "the loss of one particular DNA contact and not the result of secondary effects caused by structural distortions in neighbouring regions" (Joerger *et al* 2005, 2006). Contrary to this conclusion, our previous work on the charge–conserving mutation of Arg 273 to histidine suggested that the loss of specific DNA binding is mainly because of the loss of DNA major groove contacts from Lys 120 and Arg 280 resulting from the loss of the salt–bridge between the wild–type Arg 273 and Asp 281 (Wright *et al* 2002). To determine if these factors are also responsible for the observed DNA–binding loss upon the noncharge–conserving mutation of Arg 273 to cysteine, we have herein decomposed the free energies of wild–type and 273C p53c binding to DNA into the contributions from individual residues based on the structures from MD simulations of the DNA–free/bound molecules.

The results herein imply that the noncharge–conserving mutation of Arg 273 to cysteine in the p53c causes a loss in DNA binding not, as previously thought, due only to the loss of charge–charge interactions between the wild–type Arg 273 side chain and the DNA phosphate backbone. Instead, the loss of the DNA–binding ability of the 273C p53c mutant is also due to a combination of local conformational changes in the vicinity of the mutation site caused by the loss of the salt bridge between the wild–type Arg 273 and Asp 281 side chains. The loss of this protein–protein hydrogen bond is consistent with the recently solved 1.8-Å X–ray structure of the free t-p53c-273C pentamutant, which shows that Cys 273 can no longer hydrogen bond with Asp 281 (Joerger *et al* 2006). The loss of the Arg 273–Asp 281 salt bridge is accompanied by conformational changes that appear to rigidify certain residues that are involved in binding Zn^{2+} and/or the DNA in the free protein, especially Lys 120 from the L1 loop and Arg 280 from the H2 helix, which make significant major groove DNA contacts (*see* figure 3). Notably, the NMR solution structure of the free wild–type p53c was found to be far more mobile than the respective X–ray structure, and the L1 loop was found to be especially flexible (Canadillas *et al.* 2006). The increased rigidity of these important DNA–binding residues in the 273C mutant p53c may inhibit them from adopting the correct spatial orientation for DNA binding.

Thus, the present results and previous results on the charge–conserving mutation of Arg 273 to histidine as well as biochemical data (Rolley *et al* 1995) suggest that Asp 281 is a crucial residue in the binding of the p53c to DNA even though it does not directly contact DNA. It hydrogen bonds to both the wild–type Arg 273 and Arg 280 residues, which interact with the phosphate backbone and invariant base of the DNA consensus sequence, respectively. The Arg 273^{+}…Asp 281^{-}…Arg 280^{+} hydrogen–bonded network appear critical in holding the local structure in a favourable DNA–binding conformation. Indeed, Asp 281 has been observed to be conserved across 27 species that p53 has been sequenced for (El-Deiry *et al.* 1992). Mutation of Asp 281 to either Asn or Glu in p53 has been implied to lead to a loss of specific DNA–binding, while mutations to Tyr and Gly show complete loss of DNA binding (Zhang *et al* 1993; Rolley *et al* 1995).

Acknowledgements

We thank Dr S Noskov for helpful discussions, we are grateful to Prof. M Karplus for the CHARMM program. JW is supported by a research fellowship from the National Science Council. This work is supported by grant number NSC 93–2811–B–001–057 from the National Science Council and the Institute of Biomedical Sciences, Academia Sinica, Taiwan, ROC.

References

Abarzua, P, LoSardo J E, Gubler M L, Lu Y-A, Felix A and Neri A 1996 Restoration of the transcription activation function to mutant p53 in human cancer cells; *Oncogene* **13** 2477–2482

Babu C S and Lim C 2006 Empirical force fields for biologically active divalent metal cation in water; *J. Phys. Chem. A* **110** 691–699

Berman H M, Olson W K, Beveridge D L, Westbrook J, Gelbin A, Demeny T, Hsieh S H, Srinivasan A R and Schneider B 1992 The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids; *Biophys. J.* **63** 751–759

Bower M J, Cohen F E and Dunbrack R L 1997 Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: A new homology modeling tool; *J. Mol. Biol.* **267** 1268–1282

Brooks B R, Bruccoleri R E, Olafson B D, States D J, Swaminathan S and Karplus M 1983 CHARMm: A program for macromolecular energy, minimization, and dynamics calculations; *J. Comput. Chem.* **4** 187–217

Canadillas J M P, Tidow H, Freund,S M V, Rutherford T J, Ang H C and Fersht A R 2006 Solution structure of p53 core domain: Structural basis for its instability; *Proc. Natl. Acad. Sci. USA* **103** 2109–2114

Cho Y, Gorina S, Jeffrey P D and Pavletich N P 1994 Crystal structure of a p53 tumor suppressor-DNA complex: Understanding tumorigenic mutations.; *Science* **265** 346–355

DeLano W L 2004 *The PyMOL molecular graphics system* (DeLano Scientific)

Dunbrack R L J and Karplus M 1993 Backbone-dependent rotamer library for proteins. Application to side-chain prediction; *J. Mol. Biol.* **230** 543–574

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

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 and Vogelstein B 1993 WAF1, a potential mediator of p53 tumor suppression; *Cell* **75** 817–825

Friedlander P, Legros Y, Soussi T and Prives C 1996 Regulation of mutant p53 temperature-sensitive DNA binding; *J. Biol. Chem.* **271** 25468–25478

Garbuzynskiy S O, Melnik B S, Lobanov M Y, Finkelstein A V and Galzitskaya O V 2005 Comparison of X-ray and NMR structures: Is there a systematic difference in residue contacts between X-ray and NMR-resolved protein structures?; *Proteins Struct. Function Bioinform.* **60** 139–147

Gilson M K and Honig B H 1988 Calculation of the electrostatic potential in solution: Method and error assessment; *J. Comp. Chem.* **9** 327–335

Harper J W, Adami G R, Wei N, Keyomarsi K and Elledge S J 1993 The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases; *Cell* **75** 805–816

Hollstein, M., Rice, K., Soussi, T., Fuchs, R., Sorlie, T., Hovig, E., Smith-Sorensen, B., Montesano R and Harris C C 1994 Database of p53 gene somatic mutations in human tumours and cell lines; *Nucleic Acids Res.* **22** 3551–3555

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

Jayaram B and Jain T 2004 The role of water in protein-DNA recognition; *Annu. Rev. Biophys. Biomol. Struct.* **33** 343–361

Jayaram B, McConnell K J, Surjit B D and Beveridge D L 1999 Free energy analysis of protein-DNA binding: The EcoRI Endonuclease-DNA complex; *J. Comp. Phys.* **151** 333–357

Joerger A C, Ang H C and Fersht A R 2006 Structural basis for understanding oncogenic p53 mutations and designing rescue drugs; *Proc. Natl. Acad. Sci. USA* **103** 15056–15061

Joerger A C, Ang H C, Veprintsev D B, Blair C M and Fersht A R 2005 Structures of p53 Cancer Mutants and Mechanism of Rescue by Second-site Suppressor Mutations; *J. Biol. Chem.* **280** 16030–16037

Jorgensen W L, Chandrasekhar J, Madura J D, Impey R W and Klein M L 1983 Comparison of simple potentials for simulating liquid water; *J. Chem. Phys.* **79** 926–923

Koradi R, Billeter M and Wuthrich K 1996 MOLMOL: a program for display and analysis of macromolecular structures; *J. Mol. Graph.* **14** 51–55

Lee B and Richards F M 1971 The interpretation of protein structures: Estimation of static accessibility; *J. Mol. Biol.* **55** 379–400

MacKerell J A D, Bashford D, Bellott M, Dunbrack R, Evanseck J D, Field M J, Fischer S, Gao J *et al* 1998 All-hydrogen empirical potential for molecular modelling and dynamics studies of

proteins using the CHARMM22 force field; *J. Phys. Chem. B.* **102** 3586–3616

Mandel-Gutfreund Y and Margalit H 1998 Quantitative parameters for amino acid-base interaction: implications for prediction of protein-DNA binding sites; *Nucleic Acids Res.* **26** 2306–2312

May P and May E 1999 Twenty years of p53 research: structural and functional aspects of the p53 protein; *Oncogene* **18** 7621–7636

Nina M, Beglov D and Roux B 1997 Atomic radii for continuum electrostatic calculations based on molecular dynamics free energy calculations; *J. Phys. Chem. B* **101** 5239–5248

Ory K, Legros Y, Auguin C and Soussi T 1994 Analysis of the most representative tumour-derived p53 mutants reveals that changes in the protein conformation are not correlated with loss of transactivation or inhibition of cell prolifereation; *EMBO J.* **13** 3496–3504

Petsko G and Ringe D 1984 Fluctuations in protein structure from X-ray diffraction; *Annu. Rev. Biophys. Bioeng.* **13** 331–371

Philippopoulos M and Lim C 1995 Molecular dynamics simulation of *E. coli* ribonuclease H1 in solution: Correlation with NMR and X-ray data and insights into biological function; *J. Mol. Biol.* **254** 771–792

Pietenpol J A, Tokino T, Thiagalingam S, El-Deiry W S, Kinzler K W and Vogelstein B 1994 Sequence-specific transcriptional activation is essential for growth suppression by p53; *Proc. Natl. Acad. Sci. USA* **91** 1998–2002

Rolley N, Butcher S and Milner J 1995 Specific DNA binding by different classes of human p53 mutants; *Oncogene* **11** 763-770

Sakharov D and Lim C 2005 Zn protein simulations including charge transfer and local polarization effects; *J. Am. Chem. Soc.* **127** 4921–4929

Sharp K A, Nicholls A, Fine R F and Honig B 1991 Reconciling the magnitude of the macroscopic and microscopic hydrophobic effects; *Science* **252** 106–109

Soussi T and May P 1996 Structural aspects of the p53 protein in relation to gene evolution: a second look; *J. Mol. Biol.* **260** 623–637

Stote R and Karplus M 1995 Zinc binding in proteins and solution: A simple but accurate nonbonded representation; *Proteins Struct. Func. Genet.* **23** 12–31

Wieczorek A M, Waterman J L, Waterman M J and Halazonetis T D 1996 Structure-based rescue of common tumor-derived p53 mutants; *Nat. Med.* **2** 1143–1146

Wright J D, Noskov S Y and Lim C 2002 Factors governing loss and rescue of DNA binding upon single and double mutations in the p53 Core Domain; *Nucleic Acids Res.* **30** 1563–1574

Zhang W, Funk W D, Wright W E, Shaym J W and Deisseroth A B 1993 Novel binding of p53 mutants and their role in transcription activation; *Oncogene* **8** 2555–2559

*e*Publication: 28 June 2007