Tumor Suppressor INK4: Determination of the Solution Structure of $p18^{INK4C}$ and Demonstration of the Functional Significance of Loops in $p18^{INK4C}$ and $p16^{INK4A}$ †,‡

Junan Li, In-Ja L. Byeon, Karen Ericson, Ming-Jye Poi, Paul O'Maille, Thomas Selby, and Ming-Daw Tsai*

Departments of Chemistry and Biochemistry and Campus Chemical Instrument Center, 100 West 18th Avenue, The Ohio State University, Columbus, Ohio 43210

Received September 23, 1998; Revised Manuscript Received December 7, 1998

ABSTRACT: Since the structures of several ankyrin-repeat proteins including the INK4 (inhibitor of cyclin-dependent kinase 4) family have been reported recently, the detailed structures and the functional roles of the loops have drawn considerable interest. This paper addresses the potential importance of the loops of ankyrin-repeat proteins in three aspects. First, the solution structure of $p18^{INK4C}$ was determined by NMR, and the loop structures were analyzed in detail. The loops adapt nascent antiparallel β-sheet structures, but the positions are slightly different from those in the crystal structure. A detailed comparison between the solution structures of p16 and p18 has also been presented. The determination of the p18 solution structure made such detailed comparisons possible for the first time. Second, the $\{^1H,^{15}N\}$HSQC NMR experiment was used to probe the interactions between $p18^{INK4C}$ and other proteins. The results suggest that $p18^{INK4C}$ interacts very weakly with dna K and glutathione *S*-transferase via the loops. The third aspect employed site-specific mutagenesis and functional assays. Three mutants of p18 and 11 mutants of p16 were constructed to test functional importance of loops and helices. The results suggest that loop 2 is likely to be part of the recognition surface of $p18^{INK4C}$ or $p16^{INK4A}$ for CDK4, and they provide quantitative functional contributions of specific residues. Overall, our results enhance understanding of the structural and functional roles of the loops in INK4 tumor suppressors in particular and in ankyrin-repeat proteins in general.

The structure and function of the INK4 family of tumor suppressors are active subjects of current research in many laboratories. There are four known members in this family: $p15^{INK4B}$, $p16^{INK4A}$, $p18^{INK4C}$, and $p19^{INK4D}$ (which are abbreviated as p15, p16, p18, and p19, respectively, in this paper) (*1*−*5*). The tumor suppressor function has been well established for p16 (*5*) and to a lesser extent for p15 (*6*). Only one natural mutant from tumor cells has been identified for p18 (*7*) and p19 (*8*).

The INK4 proteins are only one type of the very large family of ankyrin-repeat proteins. Ankyrin repeats exist in a large number of proteins found in a variety of organisms ranging from plants, poxviruses, and prokaryotes to yeast and animals. They are found in sets of at least four repeats, a repeat being approximately 33 amino acids long (*9*). The ankyrin repeats have been suggested to be involved in protein−protein interactions (*9*−*12*).

The structure of ankyrin-repeat proteins was unknown until recently. Upon the basis of NMR analyses of the secondary structures of tumor suppressor p16, we first reported that ankyrin-repeats consist of helix-turn-helix (H-T-H) motifs (*13*). Subsequently, the tertiary structures of several proteins containing ankyrin repeats have been reported: the crystal structures of 53BP2 (*14*), p18 (*15*), and transcriptional regulator GABPα/β bound to DNA (*16*) and the solution structures of p19 (*17*), p16 (*18*), and myotrophin (*19*). In all of these structures, ankyrin repeats exist as H-T-H motifs linked by loops.

The loops of the ankyrin-repeat proteins have drawn considerable interest for three reasons. (a) It has been suggested that the loops (*18*) or the tips of the loops (*16*) are the possible sites for protein−protein interactions, including the interaction between p16 and CDK4, which is of enormous significance in cancer research. (b) In the crystal structure of the transcriptional regulator GABPα/β-DNA complex, the β-subunit consists of 5 ankyrin repeats linked by 4 loops. The interface between α- and β-subunits involves mainly the tips of the loops of the β-subunit (*16*). The tip of one of the loops in 53BP2 is also involved in the binding to p53 (*14*). (c) The structural details of the loops seem to differ between proteins. They are reported to form a continuous

† This work was supported by NIH Grant CA69472, and by a grant from the Comprehensive Cancer Center of OSU. The DMX-600 NMR spectrometer used was funded in part by NIH Grant RR08299 and NSF Grant BIR-9221639. The DRX-800 NMR spectrometer was funded by Ohio Board of Regents. Support from the Ohio Supercomputer Center is acknowledged.

‡ The coordinates of p18 have been deposited in the Brookhaven Protein Data Bank (file name 1BU9). This is paper 3 in the series Tumor Suppressor INK4; for paper 2, see ref *18*.

* Author to whom correspondence should be addressed. Phone: (614) 292-3080. Fax: (614) 292−1532. E-mail: Tsai.7@osu.edu.

[1] Abbreviations: AEBSF, 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride; BSA, bovine serum albumin; CDK 4, cyclin-dependent kinase 4; CDK6, cyclin-dependent kinase 6; DTT, dithiothreitol; EDTA, ethylenediamine tetraacetic acid; dna K, an *E. coli* chaperonin, heat shock protein 70 homologue; GST, glutathione *S*-transferase; HEPES, *N*-[2-hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid]; INK 4, inhibitor of CDK4; pRb, retinoblastoma gene product; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

10.1021/bi982286e CCC: $18.00
Published on Web 02/17/1999

β-sheet structure in some of the structures solved to date, but not in others. Among the proteins reported to form β-sheet structure, the extent of the β-sheet also varies.

This paper consists of three aspects related to the structure−function relationship of INK4 proteins in particular and ankyrin-repeat proteins in general. The first aspect is the determination of the solution structure of human p18. The improved stability and reduced propensity for aggregation of p18 and the use of 800 MHz NMR allowed determination of the solution structure at substantially higher resolution (relative to p16), which in turn allowed detailed examination of the structures of the loops. Comparisons between the solution structures of p18 and p16 and between the crystal structure and the solution structure of p18 are also presented. A nascent β-sheet structure was observed, but the positions differ slightly from those defined in the crystal structure.

In the second aspect, NMR was used to test whether the loops of p18 are involved in binding with other proteins, in the hope that the information learned will be applicable to p16 and other types of ankyrin-repeat proteins. Although p16 is best known to inhibit CDK4 and CDK6, there is sufficient evidence that it may also interact with other proteins, such as the Tax oncoprotein of human T-cell leukemia virus type 1 (*20*). It is likely that there are other yet unidentified proteins that can interact with p16 or p18, either specifically or nonspecifically. Our results suggest that p18 interacts with dna K and GST proteins weakly, and the sites of interactions are localized at the loop regions.

In the third aspect of this paper, we use site-specific mutagenesis to show that loop 2 is important for the function of p16 and p18. The interaction of p16 or p18 with CDK4 is likely to be of high affinity and to involve more than the loops. The size of the complex is also larger than the current limit of NMR technology. Thus, we constructed 11 mutants of p16 designed to test loop 2 and loop 3 residues, as well as the helical regions that have been suggested to be functionally important by others. The results indicate that the three loop 2 mutants show the largest increases in IC_{50} values. We then also constructed three loop 2 mutants of p18 and showed that their IC_{50} values also increase. The results provide experimental support to the importance of loop 2 in p16−CDK4 interactions. Furthermore, the functional analysis of site-specific mutants has provided the quantitative contribution of specific residues to the function of p16 or p18, which is independent of the structural information of the complexes.

MATERIALS AND METHODS

Expression and Purification of p18 and Mutants. Human p18 was expressed in a soluble form as a GST fusion protein in *Escherichia coli* BL 21 (DE 3) pLys S (Novagen). The cell lysate was purified on a glutathione-agarose column (Sigma); p18 was cleaved from the column with thrombin (Sigma) and further purified on an S-100 column equilibrated with 4 mM HEPES buffer containing 1 mM DTT and 5 mM EDTA (pH 7.5). All p18 mutants were constructed in Quickchange method (Stratagene), and their expression and purification were done in the same way as wild-type p18. Note that p18 was engineered to have a 15 residue tag at the N-terminus to facilitate cleavage by thrombin. Since the native p18 already consists of a flexible (conformationally random) N-terminus (residues 1−7), this additional tag does not interfere with structural or functional analyses of p18.

Construction of p16 Mutants. Human p16 cDNA was cloned into a GST-fusion protein expression vector pTG-p16 and was expressed in *E. coli* BL 21 (DE 3) pLys S (Novagen). All p16 mutants except W110A were constructed by the Quickchange method (Stratagene) using pTG-p16 as template, and their expression and purification were the same as p18 mutants. For W110A, the template DNA was pTN-p16, in which a Δ1−8 p16 cDNA was cloned into pTN vector (Stratagene) and expressed in *E. coli* BL 21 (DE 3) as an inclusion-body protein, which was refolded as described previously (*13*).

Assay for CDK4 Inhibition. The in vitro CDK4 inhibition assay (*1*, *3*, *18*) involved 3−10 units of the CDK4/cyclin D2 complex and varying concentrations of p16, p18, or their mutants in kinase buffer (50 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 2.5 mM EGTA, 0.1 mM Na_3VO_4, 1 mM NaF, 10 mM β-glycerophosphate, and 1 mM DTT) supplemented with 0.2 mM AEBSF, 2.5 mg/mL leupeptin, and 2.5 mg/mL aprotinin in a total volume of 15 μL. After a 30 min preincubation at 30 °C, 50 ng of "GST-Rb" and 2 μCi of [γ-^{32}P]ATP were added into the reaction mixture and the samples were incubated at 30 °C for an additional 15 min, then 3.75 μL of regular 5× SDS−PAGE loading buffer was added to each sample to stop the reaction. The reaction mixture was separated by SDS−PAGE and the phosphorylation was analyzed by Phosphorimager.

The substrate of CDK4, GST-Rb, was prepared as follows: human "large pocket" Rb cDNA corresponding residues 379−928 of wild-type Rb was cloned into pGEX-2T plasmid, and expressed as a GST-fusion protein in *E. coli* BL 21 (DE 3). The cell lysate was purified on a glutathione-agarose column. After washing with 20 mM Tris-HCl (pH 7.5)−5 mM EDTA−0.2 M NaCl−5 μM β-mercaptoethanol, the bound proteins were eluted out with the above washing buffer containing 50 mM reduced glutathione, then dialyzed against the kinase buffer.

NMR Experiments and Structural Determination. All NMR experiments were performed on a Bruker DMX-600 or DRX-800 spectrometer at 20 °C. Unlabeled, uniform ^{15}N-labeled, and uniform ^{15}N,^{13}C-double labeled proteins were used. The protein concentration of the NMR samples was 0.2−0.4 mM. The p18 samples contained 4 mM HEPES, 1 mM DTT, and 5 μM EDTA in 95% H_2O/5% ^2H_2O or 100% ^2H_2O at pH 7.5. Most of the triple resonance experiments for the total assignments were conducted on DMX-600 while all of the 2D to 4D NOESY experiments (*21*−*24*) for NOE assignments were conducted on DRX-800 using 100 ms mixing time. All data were processed using the XWIN-NMR (Bruker) or the Felix 95.0 (Molecular Simulations) program. NOE intensities were calibrated against those between backbone protons within the identified α-helical regions (*25*, *26*), then classified into strong, medium, weak, and very weak categories corresponding to the interproton distance restraints of 1.8−2.7 Å (1.8−2.9 Å for NOEs involving NH protons), 1.8−3.5 Å (1.8−3.7 Å for NOEs involving NH protons), 1.8−5.0 Å, and 1.8−6.0 Å, respectively. Upper limits for distances involving nonstereospecifically assigned protons were corrected for center averaging (*27*), and an additional 0.5 Å was added to the upper distance limits for NOEs

involving methyl protons (*28, 29*). Hydrogen bond restraints ($r_{NH-O} = 1.5-2.5$ Å, $r_{N-O} = 2.4-3.5$ Å) were deduced on the basis of the secondary structure and the slowly exchanging amide protons. The slowly exchanging amide protons were identified in the ^{15}N-HSQC spectra of p18 recorded after exchanging with ^2H_2O at 10 °C for 1 h. The HNHA (*30*) data of p18 were not very useful in giving any ϕ torsion angle restraints because only a few cross-peaks were detected. Side-chain χ1 torsion angles were derived from the coupling constants and NOE data. The ^{15}N-$^1H_\beta$ coupling ($J_{N\beta}$) was estimated qualitatively from the intensity of the correlation in the 3D HNHB spectrum (*31*). The $J_{N\beta}$ value was further verified from the 3D ^{15}N-edited NOESY in H_2O such that the H_β proton which showed strong (weak) $J_{N\beta}$ correlation should exhibit weak (strong) NOE to its backbone H_N proton. The $H_\alpha-H_\beta$ *J* coupling ($J_{\alpha\beta}$) was estimated qualitatively from the 2D 30 ms clean TOCSY (*32*) and further verified from the 3D ^{13}C-edited NOESY in ^2H_2O (*33*). When the $J_{\alpha\beta}$ information was combined to that of the $J_{N\beta}$, 18 χ1 torsion angle restraints were unambiguously obtained. The minimum range of $\pm30°$ was employed for the χ1 restraints.

Structures were calculated on Silicon Graphics O2 workstations or a Cray T90 Supercomputer using a simulated annealing method (*34*) with the X-PLOR program (*35*). The simulated annealing protocol for p18 consists of two stages: in the first stage, the simulated annealing structures were determined based on the experimental distance and dihedral restraints. The resulting structures were then used as initial structures for the second stage of simulated annealing calculations where, in addition to the distance and dihedral restraints, the structures were refined against secondary $^{13}C_\alpha$/$^{13}C_\beta$ chemical shift restraints. A total of 80 structures were generated using the protocol. The structures were inspected by MOLMOL (*36*) and Insight II (Molecular Simulations Inc.) and analyzed by PROCHECK (*37*). An ensemble of 24 lowest-energy structures was selected from a total of 60 converged structures.

Expression and Purification of dna K and GST. Both dna K and GST were purified as byproducts when we expressed CDK 4 gene in a glutathione *S*-transferase fusion protein in *E. coli* BL21 (DE 3) pLys S. Because dna K can be co-purified with GST-CDK 4 (*38*), we loaded the cell lysate on a glutathione-agarose column, then eluted the bound proteins with 50 mM reduced glutathione (pH 7.5). After dialysis to remove reduced glutathione, the eluent was loaded on a Q Fastflow Sepharose column (Pharmacia) equilibrated with 50 mM Tris-HCl (pH 8.4), then eluted with a NaCl gradient from 0 to 600 mM.

RESULTS AND DISCUSSION

Determination of the Solution Structure of p18. The total assignments of p18 were performed under a low protein concentration (0.2−0.4 mM). Although p18 is less aggregating than p16, aggregation was still a problem at higher protein concentration. P18 exhibited a number of slowly exchanging backbone amide peaks in ^2H_2O, indicating it is more stable and more structured than p16 under the same condition. The improved stability is likely caused by the additional ankyrin repeat in p18 (relative to p16). These intrinsic properties of p18, combined with the use of 800 MHz NMR, resulted in excellent spectral quality.

Nearly complete assignments of p18 were obtained using an assignment protocol similar to that described previously (*13, 39*). The assignments were based on the triple resonance correlation from the HNCA, HNCACB, HNHA, HNHB, HN(CO)CA, and CBCA(CO)NH experiments combined with the NOE patterns from the 3D ^{15}N-edited NOESY experiment. The side chain assignments were completed using HCCH-TOCSY and 3D ^{13}C-edited NOESY experiments. Aromatic side-chain assignments were obtained from the intraresidue correlation between the H_α and H_β protons and the ring protons using the 3D ^{13}C-edited NOESY spectrum. The amide side-chain assignments of asparagine and glutamine were obtained from the intraresidue correlation to the H_β or H_γ protons using the 3D ^{15}N-edited NOESY data and from those to the C_α, C_β, or C_γ carbons using the CBCA(CO)NH data. More than 97% of backbone and 90% of side-chain resonances were assigned.

The secondary structure of p18 was identified from a qualitative analysis of sequential and medium-range NOEs and the $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shift index (*40*). The final 24 simulated annealing structures were calculated on the basis of 3405 experimental restraints consisting of 2993 inter proton NOE distance restraints (see Figure 1 for the distribution of NOEs along the p18 sequence), 92 hydrogen bond restraints, 18 χ1 torsion angle restraints, and 159 C_α and 143 C_β chemical shift restraints. The ensemble of 24 structures of p18 is shown in Figure 2. Like p16, p18 also contains a random segment at the N-terminus (residues 1−7) and at the C-terminus (residues 163−168), which are not shown in Figure 2. The structures in Figure 2 converge very well with the average root-mean-square deviation (rmsd) from the mean structure of 0.50 Å for the backbone (N, C_α, and C′) atoms and 0.97 Å for all heavy atoms, and show good covalent geometry and good agreement with the restraints. None of the structures shows NOE violations of >0.6 Å and dihedral angle violations of $>6°$. The structural statistics are summarized in Table 1. The local structural precision in the ensemble of p18 structures is shown in Figure 3. Note that the local structural precision correlates well with the number and the quality of NOEs for each residue (Figure 1).

Comparison with p16 Solution Structure. The structure of p18 (five ankyrin-repeat units) shows the same structural topology as that of p16 (four ankyrin-repeat units), that is, a linear array of a repeating H-T-H structure from the ankyrin repeats. The H-T-H structures form a helix bundle, which is mainly stabilized by hydrophobic interactions. The residues of the hydrophobic cores in the helix bundles of p18 are shown in Figure 4. Each H-T-H motif is connected by a loop. The orientation of the loop is perpendicular to helical axes.

It is difficult to compare the structures of p16 and p18 in detail because of the large difference in the structural quality. As many as 3405 restraints were extracted from p18 assignments whereas less than a half (1437) restraints were obtained for p16 (*18*). However, a few differences are noticeable, especially in the loop regions. First, all four loops of p18 interact with the neighboring loops in an orderly manner, adapting continuous antiparallel β-sheet structures. The hairpin-like β-sheet structures consist of the last three residues of the preceding repeat and the first four residues of the next. The NOE pattern that clearly characterizes this structure is shown in Figure 5. In contrast, p16 (see Figure

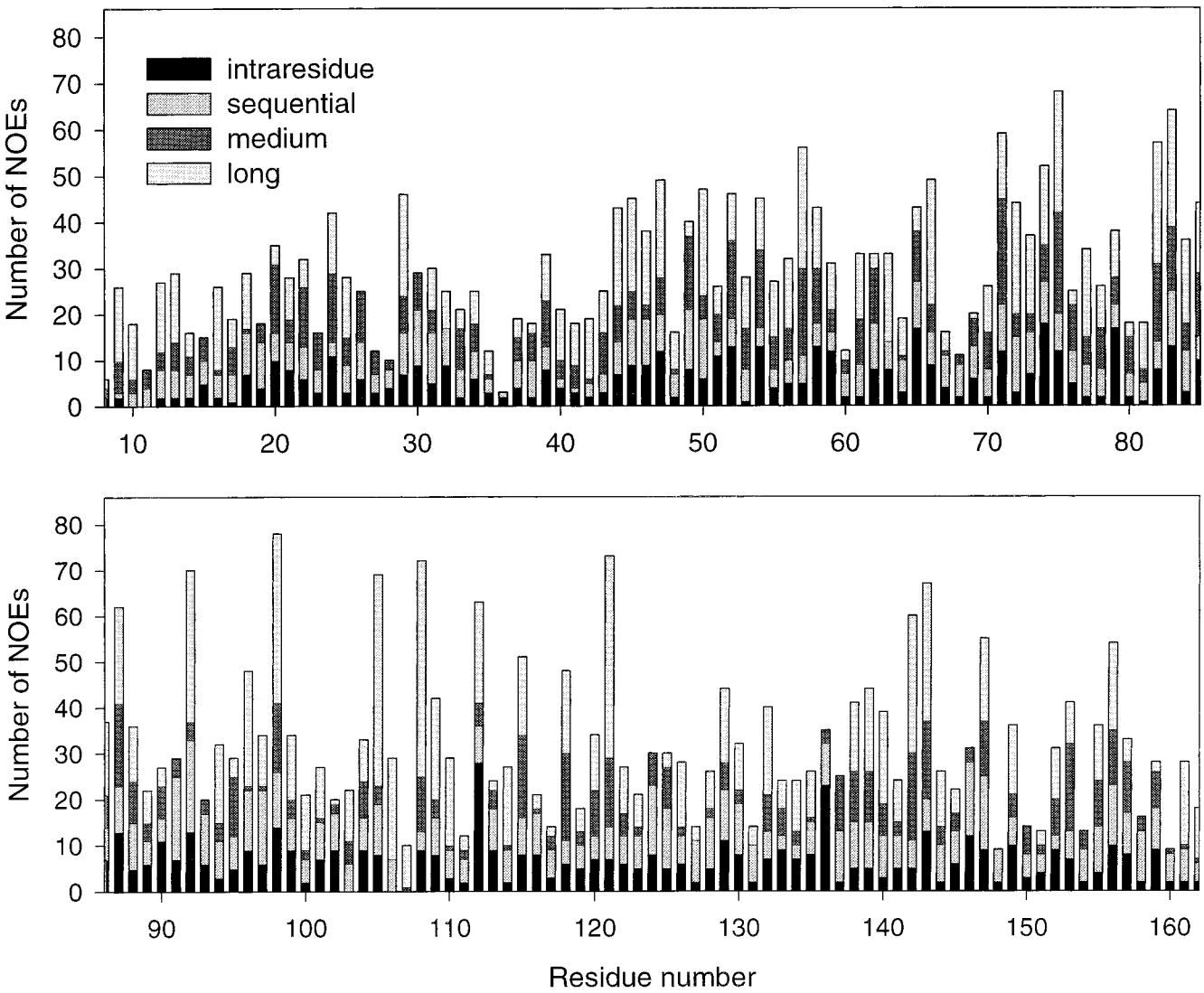

FIGURE 1: Distribution of NOEs along the p18 sequence. NOEs are grouped into intraresidue ($|i - j| = 0$), sequential ($|i - j| = 1$), medium range ($1 < |i - j| < 5$) and long range ($|i - j| = 5$ or greater).

9) shows close contact only between loops 2 and 3, not between loops 1 and 2. Further refinement of p16 structure is required to confirm this difference. Note that the solution structure of p19 (*17*) also does not show close contacts between the first and the second loops similarly to that of p16. Although the p16 structure does not show a continuous β-sheet structure (*18*), it is possible that antiparallel β hairpin structure could form between loops 2 and 3. This is supported by the observation that some of the NOEs that define the β-sheet structure of the second and the third loops in p18 were also observed in p16 in the corresponding positions.

Second, in p18 both the loops and the helices are well-defined (rmsd 0.56 and 0.43 Å, respectively; see Figure 2 and Table 1), whereas in p16 the loops (rmsd 1.32 Å) are much more poorly defined compared to the helix segments (rmsd 0.58 Å) (*18*). This difference in the structural quality for the loops may indicate that the p18 loops are rigid and well structured whereas the p16 loops are rather flexible and less structured. This is further supported by the observation that as many as 43 amide protons were still present in p18 after exchanging with ^2H_2O at 20 °C for 1 h, while none of the p16 amide protons was observed under the same condition. All but one (Leu-57) of these slow-exchanging amide protons of p18 are located on the third to the fifth ankyrin repeats (in loops and helices). The slow amide exchange for the loop residues should be mainly due to hydrogen bonds from the β-sheet since most of these residues are from the β-sheet region. A few amide protons (from residues Val-119, Glu-120, Leu-122, and Val-123 at the second helix of the fourth ankyrin repeat) were still observable even after 6 days in ^2H_2O.

Comparison with p18 Crystal Structure. The p18 solution structure is very similar to its crystal structure (*15*). All of the interactions described in the p18 crystal structure paper are also observed in the solution structure. Also, the loops of the p18 solution structure adapt nascent antiparallel β-sheet structures, similar to those in the crystal structure. However, our NMR data such as the NOE patterns (Figure 5) and the chemical shift index for the C_α and C_β resonances of p18 show that the antiparallel β hairpin structure in solution consists of seven residues instead of six suggested in the crystal structure (*15*). This makes the position of H-bonding residues due to the β-sheet structure in solution somewhat different from that in the crystal structure: the former consists of residues 34−35, 39−40, 66−67, 71−72, 99−100, 104−105, 133−134, and 138−139 whereas the latter consists of

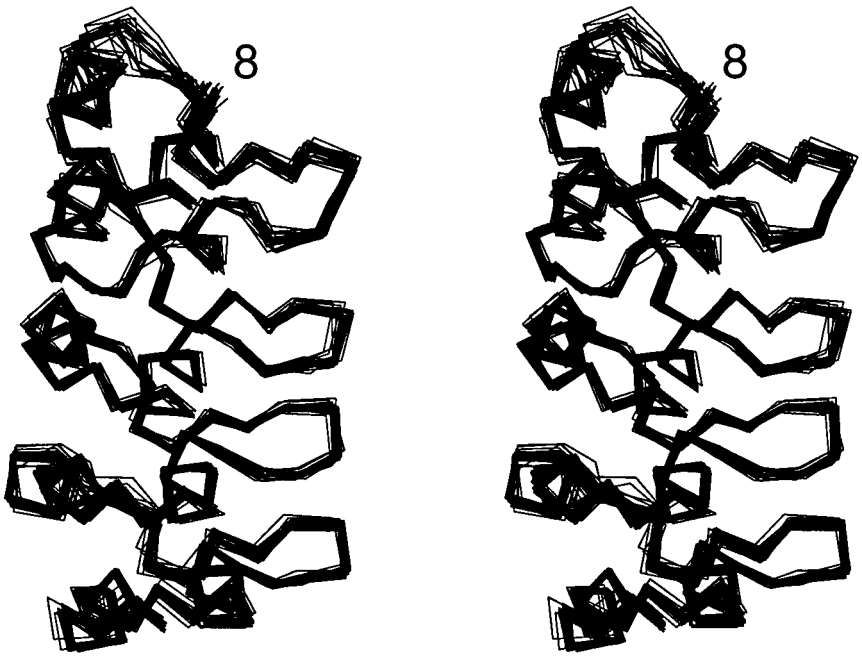

FIGURE 2: Stereoview showing the best fit superposition of the α-carbon traces of residues 8−162 of the ensemble of 24 structures. The random segments at the N-terminus (residues 1−7) and the C-terminus (residues 163−168) are not shown.

Table 1: Structural Statistics for p18[a]

	⟨*SA*⟩	$\overline{(SA)}$r
rmsd from experimental distance restraints (Å)[b]		
all (3085)	0.052 ± 0.001	0.056
sequential ($\|i-j\|=1$) (593)	0.034 ± 0.005	0.035
medium ($1<\|i-j\|<5$) (519)	0.055 ± 0.002	0.064
long range ($\|i-j\|=5$ or greater) (874)	0.041 ± 0.003	0.047
intraresidue (1007)	0.066 ± 0.002	0.068
hydrogen bonds (92)	0.049 ± 0.007	0.047
rmsd from experimental dihedral restraints (deg) (18)	1.65 ± 0.29	1.73
rmsd from experimental $^{13}C_\alpha$ restraints (ppm)(159)	1.28 ± 0.05	1.29
rmsd from experimental $^{13}C_\beta$ restraints (ppm)(143)	0.97 ± 0.05	0.99
rmsd from idealized covalent geometry		
bonds (Å)	0.005 ± 0.001	0.006
angles (deg)	0.78 ± 0.02	0.83
impropers (deg)	0.61 ± 0.02	0.67
PROCHECK (Ramachandran plot)		
most favored regions (%)	65.2 ± 1.9	68.5
additionally allowed region (%)	29.2 ± 2.6	26.0
generously allowed region (%)	4.3 ± 1.5	3.4
disallowed region (%)	1.4 ± 0.5	2.1
atomic rmsd between ⟨SA⟩ and $\overline{SA}$		
backbone heavy atoms for residues 8−162	0.50 ± 0.04	
all heavy atoms for residues 8−162	0.97 ± 0.04	
backbone heavy atoms for helical residues[c]	0.43 ± 0.05	
all heavy atoms for helical residues[c]	0.86 ± 0.07	
backbone heavy atoms for nonhelical residues[d]	0.56 ± 0.05	
all heavy atoms for nonhelical residues[d] (Å)	1.08 ± 0.05	

[a] ⟨*SA*⟩ are the final 24 simulated annealing structures; $\overline{SA}$ is the mean structure obtained by averaging the coordinates for the individual SA structures best fit to each other; and $\overline{(SA)}$r is the restrained minimized mean structure (residues 8−162) obtained by restrained minimization of the mean structure $\overline{SA}$. The number of various restraints is given in parentheses. [b] None of the structures exhibited distance violations greater than 0.6 Å or dihedral angle violations greater than 6°. [c] Residues 8−15, 18−26, 41−44, 50−58, 73−79, 83−91, 106−112, 116−123, 140−147, and 150−158 are identified as helical residues. [d] The residues from 8-162 which are not assigned as helical residue.

residues 34−35, 38−39, 66−67, 70−71, 99−100, 103−104, 133−134, and 137−138, where residue numbers are underlined if there is a difference between the crystal and solution structures in the β-sheet positions. It is interesting to note that the fine structure of the loops also varies in other ankyrin-repeat proteins: the crystal structure of 53BP2 (*14*) and the solution structure of p19 (*17*) exhibit four-residue β-hairpin structures, while myotrophin (*19*) does not show β-sheet structure.

Figure 6 shows the overlay between our high-resolution solution structure and the 1.95 Å resolution crystal structure (*15*), along with a plot of the backbone rmsd as a function of residue. Clearly, the overall folds of the two structures are similar. The rmsd for the backbone atoms of residues 8−160 is 1.48 Å. The helical region (rmsd 1.26 Å) shows better agreement than that of the β-sheet (rmsd 1.82 Å). The largest difference appears in the first β-sheet (rmsd for residues 36−39 3.14 Å) and in the last helix (rmsd for

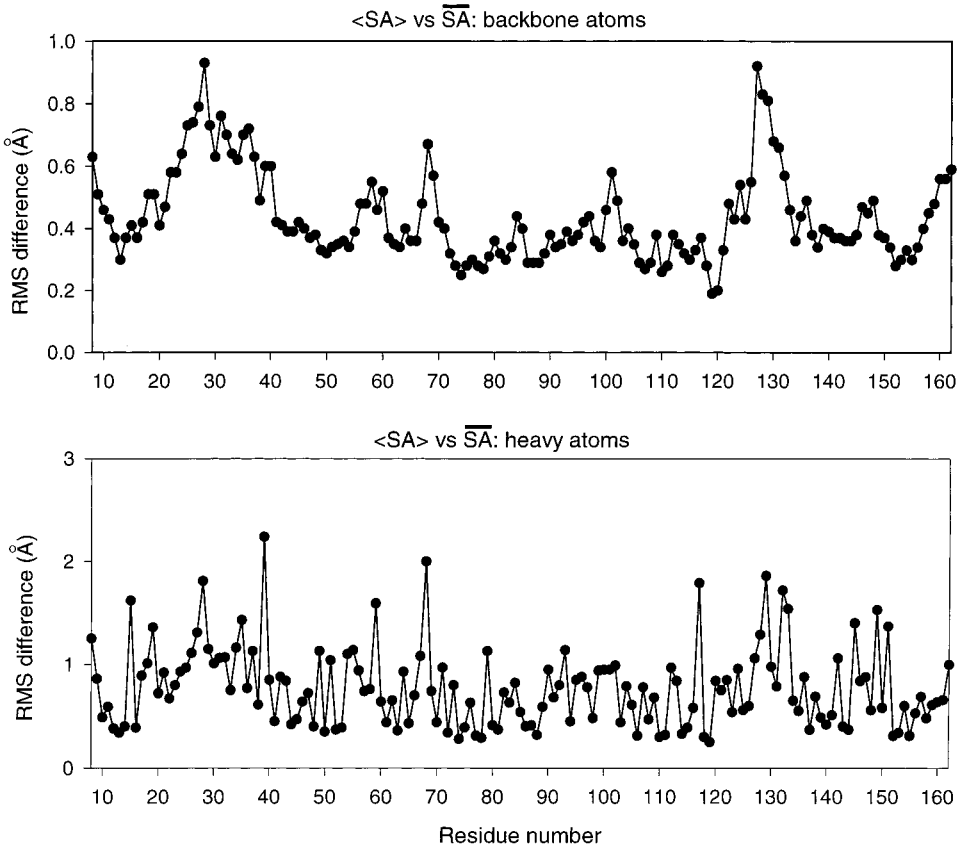

FIGURE 3: Atomic rmsd of the 24 individual simulated annealing structures relative to the mean structure, for the backbone (N, C_α, and C) and all heavy atoms as a function of residue number.

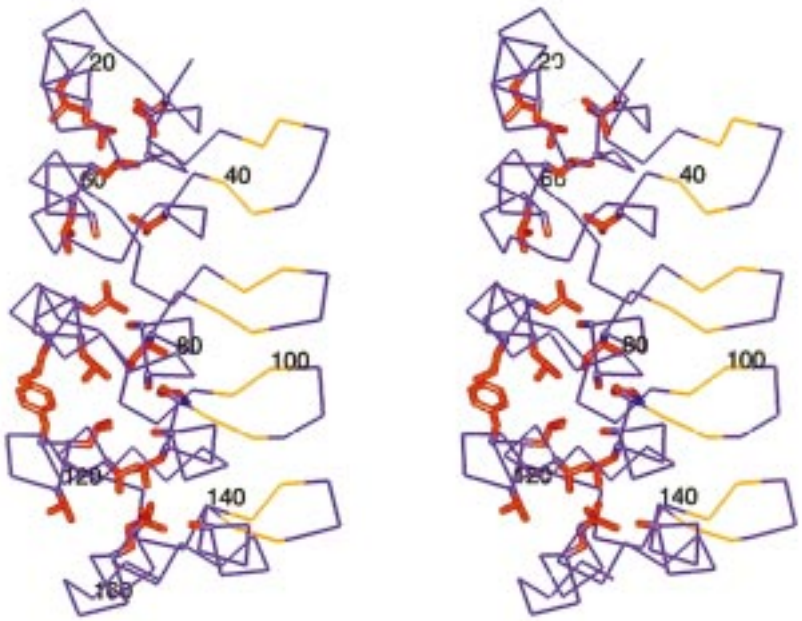

FIGURE 4: The side chains of the hydrophobic core residues in the helix bundle structures of p18 are shown in red. The β-sheet residues are indicated in orange. The restrained minimized mean structure of p18 is used to generate this figure.

residues 153−158 = 2.76 Å), possibly because loop 1 and the C-terminal helix are involved in crystal packing as noted in (*15*).

Overall, our results suggest that the differences between p18 and p16 solution structures, and between solution and crystal structures of p18, are noticeable. The functional significance of such differences remains to be established. However, the determination of the p18 solution structure has made such detailed comparisons possible for the first time.

Detection of p18 Interactions with Other Proteins via Loops. The {^{1}H,^{15}N}HSQC experiment is becoming a useful method to study protein−protein interactions. The method is based on the detection of the chemical shift changes in the spectra of ^{15}N-labeled protein upon addition of another

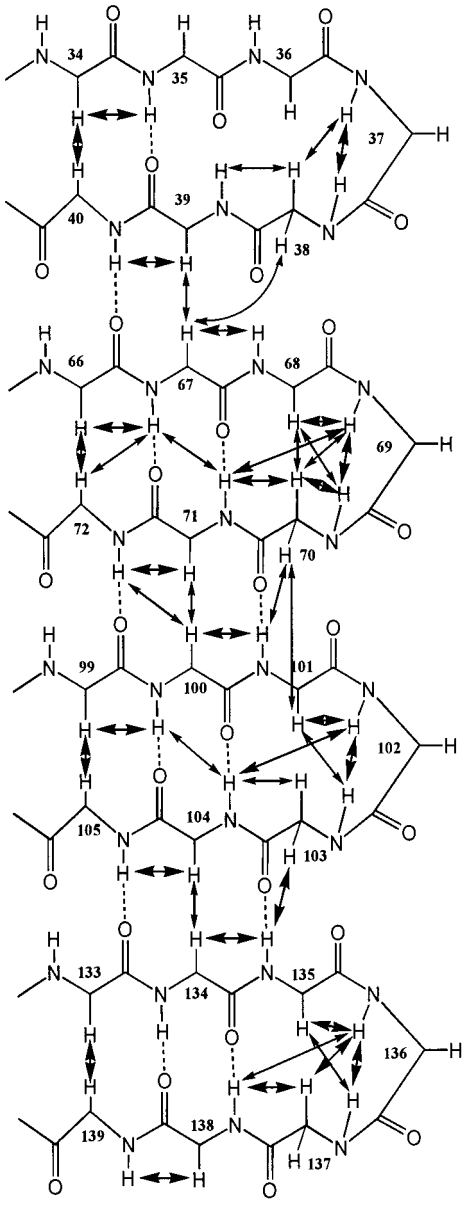

FIGURE 5: Proposed β-sheet structures in p18 from NOE data. NOEs used to define the β-sheet structures are indicated by double arrows. Heavy double arrows indicate medium to strong NOEs, whereas light double arrows indicate weak NOEs. Possible hydrogen bonds are indicated by dashed lines.

protein. The method is particularly useful if the binding is of relatively low affinity and the free and the bound forms are in fast exchange on the NMR time scale. Under this condition, the changes in the ^{15}N and ^{1}H chemical shifts of amide NH can be monitored. If the changes are confined to a limited number of residues, then these residues are most likely to be involved in binding. It has been demonstrated that the chemical shift changes usually correspond to the sites of interactions (*41*). On the other hand, this method is less useful if the binding is tight (the signals of the complex could be too broad to be detected in an intermediate exchange condition) or if binding leads to global structural changes (the signals change too much to be recognizable). In the latter two cases changes in the $\{^{1}H,^{15}N\}$HSQC spectrum will be difficult to detect and/or interpret, and it will require total assignments or structural determination of the complex to obtain the binding information.

The effect of CDK4 binding to p18 probably falls into the situation of tight binding and/or global changes and no useful information was obtained from our $\{^{1}H,^{15}N\}$HSQC experiments. However, it was found unexpectedly that both dna K (the heat shock protein 70 homologue in *E. coli*) and GST can also interact weakly with p18. Figure 7A shows the overlay of the spectra of p18 in the absence and presence of dna K. There are notable and reproducible, albeit small, changes in several cross-peaks. The small and localized changes suggest that the interaction is very weak and localized. Since all of the cross-peaks have already been assigned for free p18, the shifted cross-peaks can be readily assigned. As shown in Figure 8, the residues involved in binding are localized in the loops. Similar results were obtained for GST (spectra not shown). The results suggest that ankyrin-repeat proteins are likely to interact with other proteins via the loops.

The biological significance (if any) of the weak interactions between p18 and dna K or GST remains to be established. However, they may not be totally nonspecific since a control protein, bovine serum albumin (BSA), did not show detectable interaction with p18 as shown in Figure 7B. While some very minor changes induced by BSA cannot be ruled out, they are clearly different from that induced by dna K.

In our view, ankyrin-repeat proteins are likely to interact with a number of other proteins, with different degrees of specificity. The NMR approach presented here can be used to identify and characterize protein−protein interactions that fall between the two extreme cases (the very weak p18/dna K or p18/GST interactions and the very strong p18/CDK4 interactions). p18 is particularly useful for this purpose since it is a lot more stable and less aggregating than p16. It serves as a model for the important tumor suppressor p16, with implication for the broad family of ankyrin-repeat proteins as well.

Loop 2 Is Important for p16 Function in CDK4 Inhibition. Analysis of the interaction between p18 and CDK4 by the above NMR method was not successful possibly because it is a tight binding, and pure CDK4 was available only in very small quantity. We therefore used site-specific mutagenesis to probe the interactions. While p18 is a good structural homologue for p16, the interaction between p16 and CDK4 is most cancer-relevant. Thus, we have focused our mutagenesis studies on p16.

The interaction between p16 and CDK4 should not be confined to the loops, since residues in other regions have also been shown to be important by our lab and others previously. Other than the mutants that have already been reported previously (*13*, *18*), we here report structural and functional characterization of 11 additional mutants. The mutated residues are located on loop 2 (E69A, N71S, and L78A), loop 3 (W110A and R112A), the third helix-turn-helix motif (F90A, V95A, V96A, and R99A), and helix IVA (E119Q and L121A). Please note that the numbering of residues is different between p18 and p16. The numbering of p16 can be found in Byeon et al. (*18*). The selection of these residues was based on several factors: (a) the potential significance of the loops as described in the introductory portion of this paper and the p18 results; (b) the potential significance of loop 2 from our previous modeling result of p16−CDK4 (18); (c) the report that a 20 residue synthetic

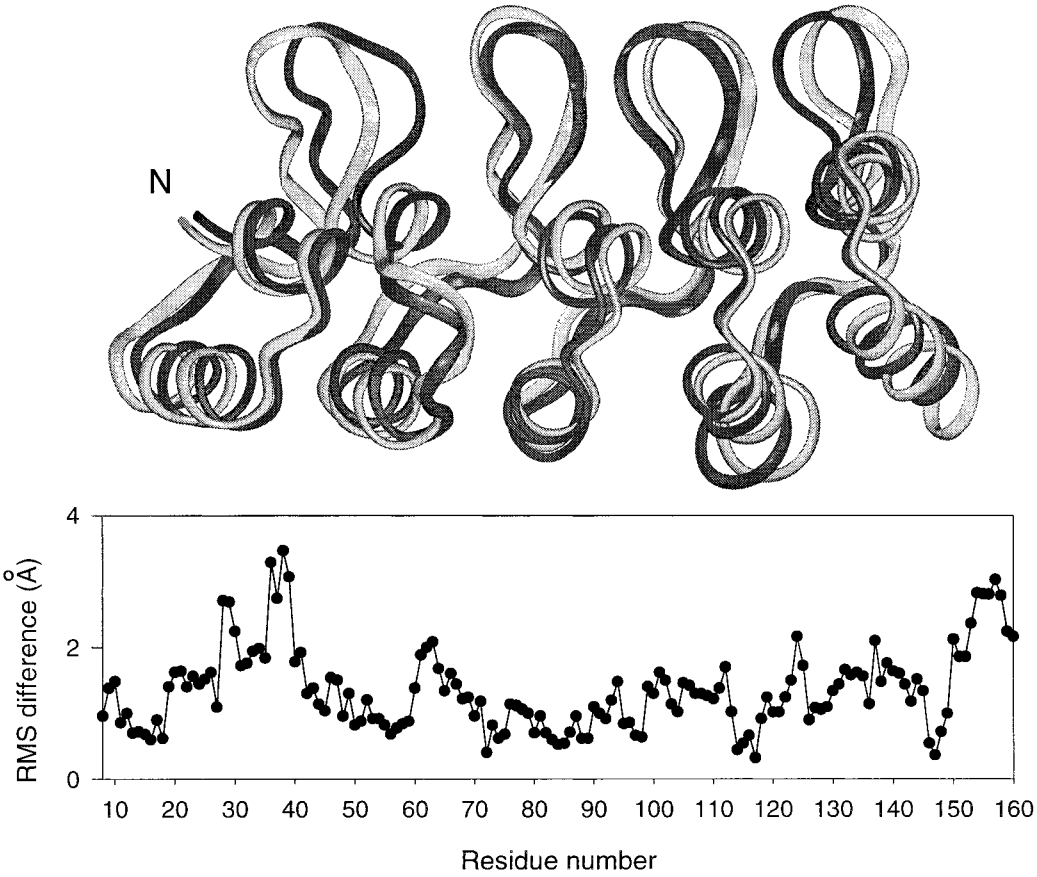

FIGURE 6: Top, best fit superposition of the backbone (N, C_α, and C) atoms of the restrained minimized mean NMR structure (black) and the X-ray structure (gray) of p18. The coordinates of the crystal structure were kindly provided by R. Marmorstein (*15*). Bottom, backbone atomic rmsd between the restrained minimized mean NMR structure and the X-ray structure.

peptide of residues 84−103 of p16 is able to bind CDK4 and CDK6 (*42*); and (d) the p19−CDK4 binding model proposed by Luh et al. (*17*) which emphasized the importance of the third helix-turn-helix motif.

The structural properties of these mutants were analyzed by 1D and/or NOESY NMR and compared with the spectra of WT p16. The activities of p16 mutants in the inhibition of CDK4 activity were analyzed as described previously (*18*). The IC_{50} values and the structural properties are listed in Table 2. As shown in Table 2, most of these mutants show unperturbed structures, except that N71S shows significantly enhanced aggregation. Thus, the mutated residues (except N71) should not play key structural roles and the functional changes of the mutants can be attributed to the functional roles of the mutated residues. Interestingly, the only two mutants showing significant increases in the IC_{50} values are E69A and L78A (5- and 33-fold, respectively), which are both loop 2 mutants. The other loop 2 mutant, N71S, shows a 3.5-fold increase. The IC_{50} values of all other mutants increase by less than 3-fold. These results suggest that loop 2 should be an important part of p16-CDK4 interactions.

Loop 2 Mutants of p18. Only one natural mutant of p18, A72P, has been identified from cancer cells (*7*), which happens to be located at loop 2 also. Therefore, we constructed and characterized this mutant, along with two other mutants on the loop 2 region: R68A and T69A. The 1D and NOESY NMR spectra of R68A and T69A are similar to those of WT p18, suggesting that the mutations did not cause global structural perturbation. The spectrum of A72P is somewhat perturbed. However, detailed analysis indicated that the perturbation is localized.

The activities of p18 mutants in the inhibition of CDK4 activity were analyzed as described previously for the assay of p16 activities (*18*). The IC_{50} values of p18 and its mutants are listed in Table 2. The results indicate that the IC_{50} values increase by 3−5-fold. Since the structures of the mutants have not been globally perturbed, the changes in the IC_{50} values can be attributed to the contributions of these residues to the function of p18. Although the effects are only modest, the results do suggest that the loop 2 of p18 is also likely to be part of the contacts between p18 and CDK4.

Summary of Functional Contributions of p16 Residues. It is commonly perceived that the purpose of functional analyses of site-specific mutants is to map out the binding site in the absence of the tertiary structure of the p16−CDK4 complex. This is not necessarily accurate, since the structure of p16 could be significantly altered upon binding to CDK4. It is also possible that there are many residues in contact between p16 and CDK4, but only some of them are important functionally. The most important information from the data in Table 2 and the mutant data published previously (*13*, *18*) is the identification of the quantitative contribution of each residue to the function of p16. Such functional information is independent of the structure of the complex, even after such a structure becomes available. The functional contributions of the residues characterized in this paper and from our previous studies are summarized in Figure 9. The color of the highlighted residues reflects the relative contri-

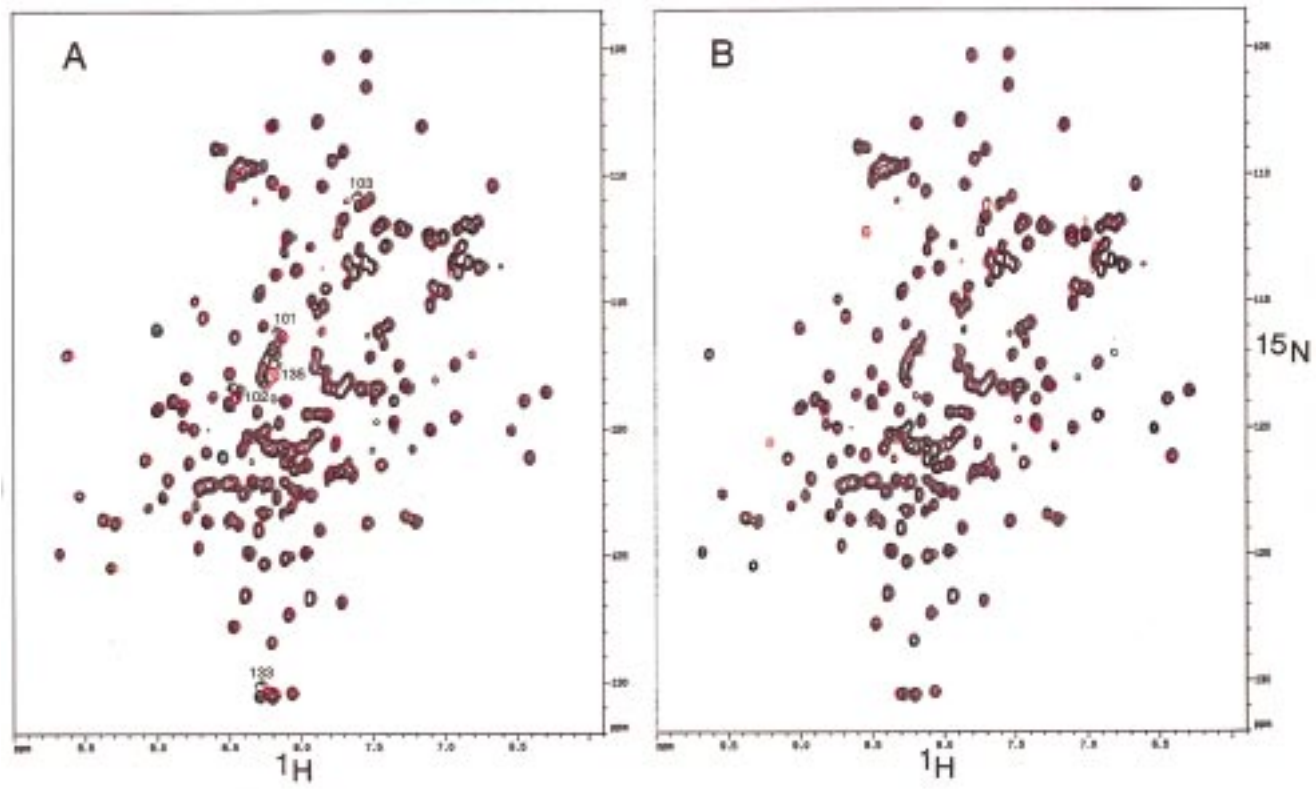

FIGURE 7: Binding of p18 with dna K. (A) Superposition of the {^{1}H, ^{15}N}HSQC spectra of free p18 (black) and p18 in the presence of 2 equivalents of dna K (red). Essentially the same results were obtained for GST. The residues with larger chemical shift changes are labeled. (B) Control experiments showing that BSA caused little changes in the {^{1}H, ^{15}N}HSQC spectrum of p18 (free p18 is black; p18 + BSA is red). The two extra red peaks in B can also be detected in the free form (black) if it is printed at a higher noise level.

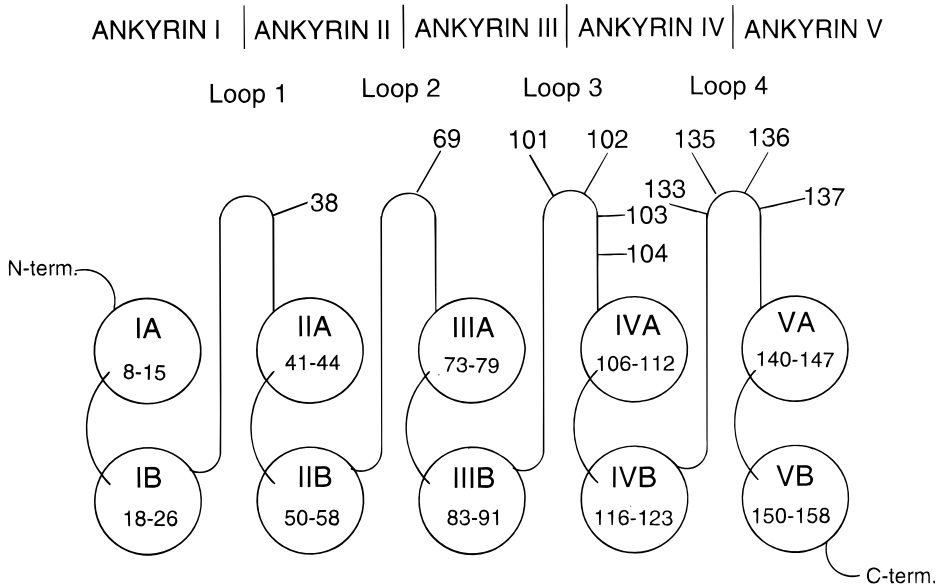

FIGURE 8: Residues involved in the interaction of p18 with dna K. The residues whose chemical shifts are perturbed upon binding of p18 with dna K are indicated. Essentially the same results were obtained for GST. The residues include those labeled in Figure 7A plus others with smaller changes.

butions to the function. Only the mutants from our own lab, which went through systematic structural and functional analyses, are used in constructing Figure 9.

Unlike enzyme catalysis where the key catalytic function is often contributed by a limited number of residues, protein−protein interactions often involve a greater number of residues with smaller contributions from each. Thus, we consider an increase of even only 3−5-fold in the IC_{50} value to be modestly significant (purple color). Larger increases are given green, orange, and red colors. An increase of <2-fold is considered insignificant and not colored.

On the basis of the literature information and the results presented in this paper, we extend our previous hypothesis (*18*) to suggest that the loops are an essential element in protein−protein interactions involving ankyrin-repeat proteins and that additional interactions involving residues at

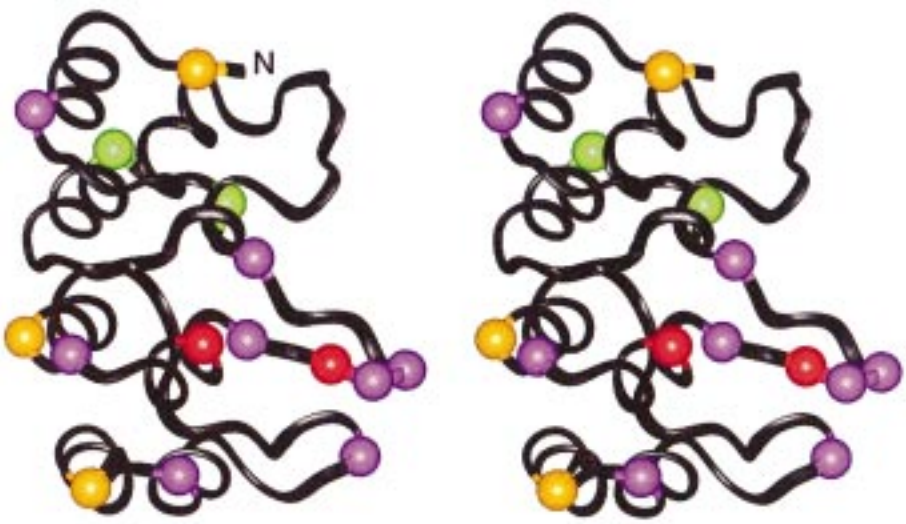

FIGURE 9: Contributions of functionally important residues of p16. The quantitative contributions of specific residues to the inhibition of CDK4 are indicated by color using the NMR structure of p16 (*18*). If mutation of this residue causes the IC_{50} to increase by >20-fold, it is indicated by red (residues 78 and 84); 10−20-fold, orange (residues 15, 92 and 124); 5−10-fold, green (residues 66 and 69); 3−5-fold, purple (residues 26, 71, 76, 77, 80, 83, 90, 110, and 121). Only the results from our lab are included. The results of p18 are incorporated in the p16 structure (R68, T69, and A72 of p18 correspond to P76, A77, and T80 of p16, respectively). The mutants with severe structural perturbations are not included. In interpreting the results, one should take into consideration that they are based mainly on one mutant, and that some of the mutants are not alanine mutants.

Table 2: IC_{50} Values and Structural Properties of Loop 2 Mutants of p18 and p16

protein	location	IC_{50} (nM)	structural properties	natural mutation
p16 WT		70 ± 25		
p16 E69A	loop 2	350 ± 45	retained	E69K and V
p16 N71S	loop 2	250 ± 87	serious aggr.	yes
p16 L78A	loop 2	2300 ± 700	retained	frameshift
p16 F90A	turn III	200 ± 60	retained	frameshift
p16 V95A	helix IIIB	110 ± 15	retained	yes
p16 V96A	helix IIIB	85 ± 15	retained	no
p16 R99A	helix IIIB	95 ± 8	retained	R99Q
p16 W110A	loop 3	160 ± 65	retained	stop
p16 R112A	loop 3	70 ± 14	retained	R112G
p16 E119Q	helix IVA	80 ± 11	retained	yes
p16 L121A	helix IVA	150 ± 25	retained	frameshift
p18 WT		80 ± 20		
p18 R68A	loop 2	250 ± 35	retained	no
p18 T69A	loop 2	360 ± 50	retained	no
p18 A72P	loop 2	400 ± 30	slightly perturbed	yes

other regions can provide additional specificity and affinity in complicated systems such as the complexes between p16 or p18 and CDK4 or CDK6.

Comparison with Crystal Structures of Complexes. After submission of our paper, the crystal structures of p16−CDK6 complex (*43*) and p19−CDK6 complex (*44*) have been published. While our work has been performed on CDK4 (which has not yet been overexpressed in large quantity), CDK6 is a close homologue of CDK4. Thus, it makes sense to evaluate our functional results on the basis of these crystal structures. Overall, our results are consistent with the crystal structures in the following ways: (a) It has been clearly demonstrated in both crystal structures that the loops, particularly loop 2, are involved in binding with CDK6. This agrees with our view that loops are important for p16−CDK4 interactions. (b) In the crystal structures, most of the interactions come from the most conserved second and third ankyrin repeats of p16, which are connected by loop 2. As shown in Figure 9, this is also the region where the majority of binding residues have been identified from our functional analyses. (c) Specifically, a hydrogen-bonding interaction between the side chain of a tip residue at loop 2 (D71 in p19 and D74 in p16) and the side chain of R31 in CDK6 has been identified in the complexes. This residue corresponds to Asp67 in p18. This result agrees with our finding that mutations of surrounding residues in p18 (R68A, T69A, and A72P mutants) led to modest increases in IC_{50} values.

The availability of crystal structures of the p16−CDK6 and p19−CDK6 complexes represents a major advancement in the field. It will allow us to further evaluate p16−CDK4 interactions from functional analyses and to evaluate quantitative contributions of each specific residue to the function of p16. Such work is actively in progress. It is clear from the crystal structures and from our functional results that the interactions are not limited to the loops, and it is also anticipated that details of the interactions between ankyrin repeats and other proteins could vary for different proteins.

As an interesting note, we see a remote similarity between ankyrin repeat proteins and antibodies. It is well established that the complementation determining region of antibodies lies in the variable loops on the surface of the Fab fragments. The loops are held together by β-sheets. Thus, while ankyrin-repeat proteins and antibodies have very different structures and functions, they both employ loops in protein−protein interactions, only that the loops are held together by α-helices in one case and by β-sheets in the other.

ACKNOWLEDGMENT

We thank Dr. Ronen Marmorstein, University of Pennsylvania, for generously providing us the coordinates of the crystal structure of p18. Human p18 expression vector, pGEX-p18, was a generous gift from Dr. Hengming Ke at the University of North Carolina. L. Ciobanu is acknowledged for assistance in purifying a p18 mutant, and J. Cloud for purifying some p16 mutants. Chemical shift tables will be available upon request.

REFERENCES

1. Serrano, M., Hannon, G. J., and Beach, D. (1993) *Nature 366*, 704−707.
2. Kamb, A., Gruis, N., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavigian, S., Stockert, E., Day, R., III, Johnson, B., and Skolnick, M. (1994) *Science 264*, 436−440.
3. Hannon, G. J., and Beach, D. (1994) *Nature 371*, 257−261.
4. Guan, K. L., Jenkins, C. W., Li, Y., O'Keefe, S. N., Wu, X. Y., Zariwala, M., Matera, A. G., and Xiong, Y. (1996) *Mol. Biol. Cell 7*, 57−70.
5. Serrano, M. (1997) *Exp. Cell Res. 237*, 7−13.
6. Okamoto, A., Hussain, S. P., Hagiwara, K., Spillare, E. A., Rusin, M. R., Demetrick, D. J., Serrano, M., Hannon, G. J., Shiseki, M., Zariwala, M., Xiong, Y., Beach, D. H., Yokota, J. and Harris, C. C. (1995) *Cancer Res. 55*, 1448−1451.
7. Lapointe, J., Lachance, Y., Labrie, Y., and Labrie, C. (1996) *Cancer Res. 56*, 4586−4589.
8. Zariwala, M., and Xiong, Y. (1996) *Oncogene 13*, 2033−2038.
9. Bork, P. (1993) *Proteins 17*, 363−374.
10. Lux, S. E., John, K. M., and Bennett, V. (1990) *Nature 344*, 36−42.
11. Greenwald, I., and Rubin, G. M. (1992) *Cell 68*, 271−281.
12. Helps, N. R., Barker, H. M., Elledge, S. J., and Cohen, P. T. (1995) *FEBS Lett. 377*, 295−300.
13. Tevelev, A., Byeon, I.-J. L., Selby, T., Ericson, K., Kim, H.-J., Kraynov, V., and Tsai, M.-D. (1996) *Biochemistry 35*, 9475−9487.
14. Gorina, S., and Pavletich, N. P. (1996) *Science 274*, 1001−1005.
15. Venkataramani, R., Swaminathan, K. and Marmorstein, R. (1998) *Nat. Struct. Biol. 5*, 74−81.
16. Batchelor, A. H., Piper, D. E., de la Brousse, C., McKnight, S. L., and Wolberger, C. (1998) *Science 279*, 1037−1041.
17. Luh, F. Y., Archer, S. J., Domaille, P. J., Smith, B. O., Owen, D., Brotherton, D. H., Raine, A. R. C., Xu, X., Brizuela, L., Brenner, S. L., and Laue, E. D. (1997) *Nature 389*, 999−1003.
18. Byeon, I. L., Li, J., Ericson, K., Selby, T. L., Tevelev, A., Kim, H. K., O'Maille, P. and Tsai, M.-D. (1998) *Mol. Cell 1*, 421−431.
19. Yang, Y., Rao, N. S., Walker, E., Sen, S. and Qin, J. (1997) *Protein Sci. 6*, 1347−1351.
20. Suzuki, T., Kitao, S., Matsushime, H., and Yoshida, M. (1996) *EMBO J. 15*, 1607−1614.
21. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry 28*, 6150−6156.
22. Sklenar, V., Piotto, M., Leppik, R., Saudek, V. (1993) *J. Magn. Reson. A 102*, 241−245
23. Fesik, A. W., and Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 78*, 588−593.
24. Vuister, G. W., Clore, G. M., Gronenborn, A. M., Powers, R., Garrett, D. S., Tshudin, R., and Bax, A. (1993) *J. Magn. Reson. B 101*, 210−213.
25. Williamson, M. P., Havel, T. F., and Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295−315.
26. Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., and Gronenborn, A. M. (1986) *EMBO J. 5*, 2729−35.
27. Wüthrich, K., Billeter, M., and Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
28. Clore, G. M., Gronenborn, A. M., Nilges, M., and Ryan, C. A. (1987) *Biochemistry 26*, 8012−8023.
29. Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., and Wüthrich, K (1987) *J. Mol. Biol. 196*, 611−639.
30. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR 4*, 871−878.
31. Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) *J. Magn. Reson. 95*, 636−41.
32. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870.
33. Driscoll, P. C., Gronenborn, A. M., and Clore, G. M. (1989) *FEBS Lett. 243*, 223−33
34. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 239*, 129−136.
35. Brünger, A. T. (1992) X-PLOR, Version 3.1: A system for X-ray 36. Crystallography and NMR (Yale University Press, New Haven and London).
36. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
37. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−91.
38. Sherman, M. Y., and Goldberg, A. L. (1992). *EMBO J. 11*, 71−77.
39. Byeon, I.-J. L., Yan, H., Edison, A. S., Mooberry, E. S., Abildgaard, F., Markley, J. L., and Tsai, M.-D. (1993) *Biochemistry 32*, 12508−12521.
40. Spera, S., and Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490−5492.
41. Rajagopal, P., Waygood, E. B., Reizer, J., Saier, M. H., Jr. and Klevit, R. E. (1997) *Protein Sci. 6*, 2624−2627.
42. Fåhraeus, R., Paramio, J. M., Ball, K. L., Lain, S., and Lane, D. P. (1996) *Curr. Biol. 6*, 84−91.
43. Russo, A. A., Tong, L., Lee, J.-O., Jeffrey, P. D., and Pavletich, N. P. (1998) *Nature 395*, 237−243.
44. Brotherton, D. H., Dhanaraj, V., Wick, S., Brizuela, L., Domaille, P. J., Volyanik, E., Xu, X., Parisini, E., Smith, B. O., Archer, S. J., Serrano, M., Brenner, S. L., Blundell, T. L., and Laue, E. D. (1998) *Nature 395*, 244−250.

BI982286E