Protein Science (1996), *5*:1776–1784. Cambridge University Press. Printed in the USA.

Structural characterization of the tumor suppressor p16, an ankyrin-like repeat protein

JUDITH A. BOICE AND ROBERT FAIRMAN

Division of Macromolecular Structure, Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, New Jersey 08543-4000

(RECEIVED May 10, 1996; ACCEPTED June 18, 1996)

Abstract

The p16 protein has been identified as a tumor suppressor that functions by inhibiting the cyclin-dependent kinases CDK4 and CDK6. Deletions or point mutations in the p16 gene have been found in a number of human cancers, emphasizing its importance in regulating cell cycle progression. Inhibition by p16 occurs through protein–protein interactions with its targets. This is not surprising, since p16 is thought to contain ankyrin-like repeats, motifs implicated in protein–protein interactions. Our goal was to identify structural characteristics of p16 not only as an important step towards understanding CDK4 inhibition but also to explore the role of ankyrin repeats in the p16 structure, as no detailed structure of any protein containing these motifs has been reported. We have expressed, refolded, and purified p16 from *E. coli* and have shown it to be functionally active by specific binding to CDK4. Analytical ultracentrifugation has shown that p16 weakly self-associates to form dimers with a K_d = 270 μM. The CD spectrum indicates that the protein is composed of 33% α-helix, 22% β-sheet, 19% β-turn, and 27% other (which includes aromatic and random coil contributions). Further CD experiments suggest that p16 exhibits low structural stability with a ΔG of -2.3 kcal/mol. This weak stability is a consequence of a highly dynamic structure as measured by ANS-binding, NMR hydrogen-deuterium exchange, and fluorescence. It is possible that a well-defined tertiary structure is imparted upon the binding of p16 to CDK4.

Keywords: analytical ultracentrifugation; ankyrin-like repeat; cyclin; circular dichroism; cyclin dependent kinase inhibitor

Control of the eukaryotic cell cycle is a highly regulated process that involves the interaction of a large number of macromolecules. Critical checkpoints in this process are found during the entry of cells into both the G1/S and G2/M phases of the cell cycle (for reviews, see Draetta, 1994; Sherr, 1995). The cyclin-dependent kinases (CDKs) are enzymes whose activities govern these checkpoints. CDK holoenzymes, which consist of a CDK catalytic subunit and a cyclin regulatory subunit, are regulated, in part, by their temporal expression, degradation, phosphorylation, and dephosphorylation. This regulation throughout the cell cycle serves to drive the cell through the successive phases.

Recent studies have identified CDK inhibitors, a new class of proteins, adding yet another level of control to CDK regulation (for review, see Elledge & Harper, 1994). These CDK inhibitors are thought to function by blocking and/or disrupting CDK-cyclin association. p16, one of the first of these inhibitors to be discovered, was originally cloned from a yeast-two hybrid interaction screen using CDK4 as the bait protein (Serrano et al., 1993). Subsequent studies support its role as a tumor-suppressor protein (Kamb et al., 1994; Nobori et al., 1994; Moulton et al., 1995; Okamoto et al., 1995; Ranade et al., 1995; Shapiro et al., 1995; Washimi et al., 1995).

The p16 protein binds specifically to CDK4 and CDK6. These CDKs, in association with the D-type cyclins, are involved in the progression of cells from the G1 to S-phase (Weinberg, 1995). Binding of p16 inhibits the ability of either CDK4 or CDK6 to phosphorylate, and thus, inactivate the retinoblastoma protein (RB). RB, a well-established tumor-suppressor protein, functions by sequestering, and thereby inactivating, the E2F family of transcription factors (Hiebert et al., 1992; Nevins, 1992). These transcription factors are involved in the activation of genes that are necessary for cell cycle progression and DNA synthesis. When RB is phosphorylated by CDK4 or CDK6, it is no longer able to bind E2F family members (Chellappan et al., 1991; Hunter & Pines, 1994; Sherr, 1994), resulting in the activation of E2F target genes. By inhibiting CDK4 and CDK6, p16 keeps the cell from progressing through the cycle by ensuring that RB remains in the hypophosphorylated state, able to sequester E2F.

We purified active p16 and carried out structural characterization to understand how this protein interacts with CDK4. The

Reprint requests to: Robert Fairman, Division of Macromolecular Structure, Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, New Jersey 08543-4000; e-mail: fairman@bms.com.

structural work described here will serve as an important basis for understanding its role in inhibiting CDK4 through protein–protein interactions. The structure of p16 has been proposed to be composed almost exclusively of ankyrin repeats, which are thought to be critical determinants in many protein–protein interactions. Although ankyrin repeats are found in a large number of proteins, there is surprisingly limited structural information available. Therefore, p16 should serve as a model system for understanding the role of ankyrin repeats in protein–protein interactions.

Results

We subcloned human p16 from a HeLa cell cDNA library clone using standard PCR methods. Sequencing of our p16 clone revealed a valine to glycine change at amino acid position 27 as compared to the originally published sequence (Serrano et al., 1993) However, this change was noted upon deposition of the sequence to GenBank (accession # l27211). The GenBank sequence also includes eight additional amino acids upstream of the original published sequence. These two versions of p16 have been shown to be identical in activity when compared in other laboratories (Serrano et al., 1993; Yang et al., 1995). The eight additional amino acids are not included in our protein, as the cloning was started based on the original publication. At this time, it is not clear which methionine is the actual starting amino acid.

Overexpression of p16 in *E. coli* at 150 mg/L of media, as a fusion with the S-peptide portion of ribonuclease (Richards & Vithayathil, 1955), resulted in the partitioning of the fusion protein almost exclusively into inclusion bodies (Fig. 1). When compared to Luria Broth, M9CAG medium yielded superior expression levels; therefore, we used M9CAG exclusively. Urea-solubilized protein was purified by Q-sepharose anion exchange chromatography followed by S-sepharose cation exchange chromatography under denaturing conditions. After these first two chromatography steps, the protein was typically greater than 95% pure (Fig. 1).

Urea solubilized S-Tag-p16 was refolded by dilution to 1 M urea at a final protein concentration of 35 μg/mL. Aggregation was observed, as determined by gel filtration, when refolding was carried out at 100 μg/mL or 1 mg/mL. Anion exchange chromatography effectively removed any aggregated protein after refolding. The yield of purified, refolded protein was approximately 65%. The S-Tag was removed by thrombin cleavage to yield p16 with an additional glycine and serine at the N-terminus. The purity of the final protein was determined to be approximately 97% by gel densitometry.

During development of the purification protocol, significant internal cleavage was noted at pH 5 and below. This cleavage occurred at a particularly labile Asp-Pro bond, as determined by amino acid sequencing. This observation is surprising, as most Asp-Pro bonds are only labile under highly acidic conditions. Due to these findings, the time that the protein was exposed to an acidic environment was kept to a minimum.

Electrospray mass spectrometry of S-Tag-p16 and S-Tag-cleaved p16 yielded molecular masses that were within 1 dalton of that expected, 18,757 and 15,947, respectively. Additionally, both samples consistently yielded a minor component approximately 29 Da less than the expected mass. Amino acid sequencing from an internal site yielded sequence identical to that of the published p16 sequence.

To test functionality, the ability of the recombinant p16 to interact with CDK4 was assessed. Immunoprecipitations using a reticulocyte lysate-generated, ^{35}S-labeled CDK4, recombinant p16 and p16 antibody showed that recombinant p16 bound specifically to CDK4, while little or no binding was observed using a mutant form of CDK4 (Fig. 2). The intact fusion protein also binds to CDK4 as when incubated with S-Tag-p16, CDK4 is precipitated using S-protein agarose. There was no difference observed in the amounts of S-Tag-p16 and S-Tag-cleaved p16 precipitated with p16 antibody. This indicates that the S-Tag on the N-terminus does not interfere with the ability of p16 to interact with CDK4.

These recombinant proteins were also found to be potent inhibitors of CDK4/cyclin D1-mediated RB phosphorylation when tested in an in vitro kinase assay (IC_{50} in the low nM range), providing additional proof of their function (data not shown).

Circular dichroism (CD) was used to measure the secondary structure of p16. CD spectra for both S-Tag-p16 and p16 are characteristic of an α-helical protein based on the minima at 222 and 208 nm and the maximum at 192 (Fig. 3). However, the low signal intensity at 222 nm as well as the proportion of the 208 nm and 222 nm minima suggests additional secondary structure (Yang et al., 1986). Spectral deconvolution yielded 33% α-helix, 22% β-sheet, 19% turn, and 27% other (including components of aromatic and random coil CD) according to the method of Hennessey and Johnson (1981) with a variable selection of protein spectra from a 33-spectra basis set (Manavalan & Johnson, 1987) (Table 1). The

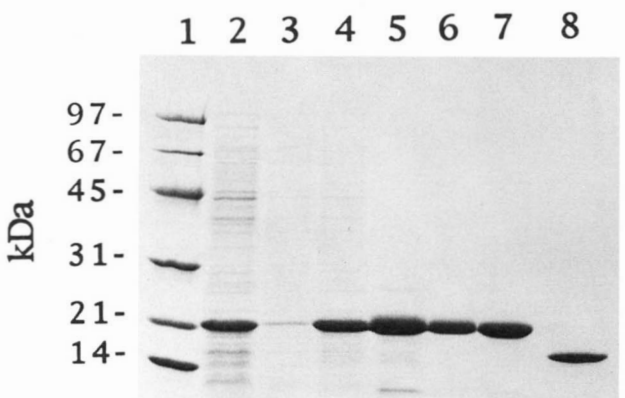

Fig. 1. Purification of p16 analyzed by SDS-PAGE (4–20% gradient gel). Molecular weights in kDa indicated on the left. Lane 1, molecular weight markers; Lane 2, *E. coli* whole-cell lysate; Lane 3, supernatant of lysate; Lane 4, inclusion body pellet from lysate; Lane 5, urea solubilized inclusion body pellet, cleared by centrifugation; Lane 6, sample after S-sepharose cation exchange chromatography; Lane 7, S-Tag-p16, after final purification; Lane 8, final purified p16 after removal of the S-Tag by thrombin cleavage.

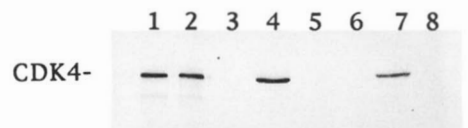

Fig. 2. p16 binding activity using ^{35}S-CDK4 followed by SDS-PAGE (13% gel). The migration of CDK4 is indicated on the left. Lane 1, reticulocyte (retic) lysate expressed CDK4; Lane 2, retic lysate expressed mutant CDK4; Lane 3, CDK4 precipitated with S-Protein agarose in the absence of added p16; Lane 4, CDK4 + S-Tag-p16, precipitated with S-Protein agarose; Lane 5, mutant CDK4 + S-Tag-p16 precipitated with S-Protein agarose; Lane 6, CDK4 immunoprecipitated with p16 antibody in the absence of added p16; Lane 7, CDK4 + p16, immunoprecipitated with p16 antibody; Lane 8, mutant CDK4 + p16 immunoprecipitated with p16 antibody.

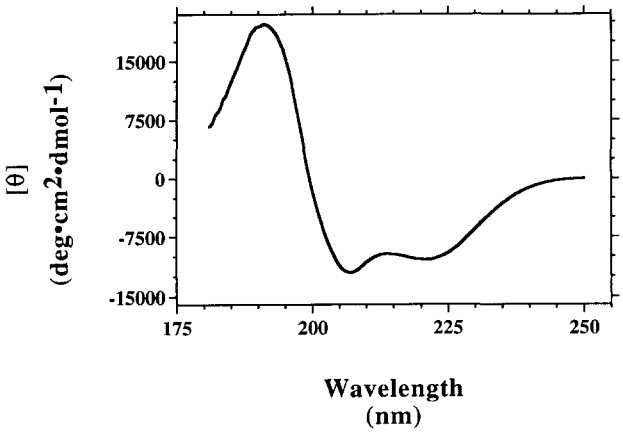

Fig. 3. CD spectrum of p16. Conditions: 20 μM p16, 5 mM Na phosphate, pH 7.4, 10 mM NaF.

total amount of structure accounted for with these results was 100%, with a root mean square error of 0.16. A means to assess the accuracy of this prediction is to apply other deconvolution algorithms and look for consistency. We obtained similar values deconvoluting the data using several different methods as well as different basis sets (Table 1), suggestive of a high level of accuracy for these results. Interestingly, when a poly-proline type II helix was included in the consideration using the algorithm reported by Woody and colleagues (Sreerama & Woody, 1994), a portion of the random (other) structure was given this assignment. This indicates that p16 may contain this type of extended helical structure.

The quaternary structure of p16 was measured using sedimentation equilibrium ultracentrifugation. To accurately determine the oligomeric state of functional p16, sedimentation equilibrium was applied at different protein concentrations. Results for 5, 15, and 45 μM samples indicated that p16 is primarily monomeric for the low end of the concentration range, as typically used in this study.There was, however, indication of some self-association at higher concentrations.

The data obtained with the 5 μM sample were well fit as a single species with the molecular weight (M_r) fixed to that of monomeric p16 (15,947), as judged by the randomness of the residuals from the global four-speed curve fit (data not shown).

The apparent molecular weights from the 15 and 45 μM samples, as determined by a single species analysis where the M_r was treated as an adjustable parameter, were greater than that of monomeric p16 (Table 2). The increase in M_r with increasing protein concentration is indicative of self-association; therefore, we considered monomer-oligomer equilibrium models. To evaluate the possible models, the square root of variances (SRV), which is a measure of the goodness of fit, were compared for each model (Table 2). Consideration of monomer-oligomer models resulted in an improved SRV, an indication that these models fit better than a single species, monomer model. This trend in SRV was even more pronounced for the fits to the data obtained with the 45 μM sample. The data for a monomer-dimer model were well fit, as judged by the randomness of the residuals (Fig. 4). As is apparent, the single species analysis, fixing the M_r to monomeric p16, resulted in nonrandom residuals, as well as a significant increase in the SRV (Table 2).

Because of the weak self-association, we cannot absolutely identify the n-mer species. However, based on a comparison of SRVs, our results suggest a monomer-dimer equilibrium with a K_d of approximately 270 μM. There are three reasons for this conclusion. First, the 15 and 45 μM data show an increasing, albeit small, trend in the observed SRV when comparing monomer-dimer, monomer-trimer, and monomer-tetramer equlibria. Second, the residuals appear to be most random for the monomer-dimer fits. This is especially true when compared to the monomer-tetramer fits. Third, the best agreement in K_ds across the different protein concentrations is observed for a monomer-dimer model. This third point is, in fact, considered to be the most rigorous test of oligomeric state determination. Similar results were obtained for S-Tag-p16 analyzed at the same concentrations (data not shown).

To gain further insight into the structure of p16, we evaluated its stability using thermal unfolding as monitored by CD. Thermal denaturation provides information about the nature of the tertiary structure. The midpoint of the thermal unfolding transition, or Tm,

Table 1. *Estimation of the secondary structure content of p16*

Secondary structure	Johnson[a]	Fasman[b]	Woody[c]	Woody[d]	Woody[e]
α-Helix	33%	37%	30%	34%	35%
Antiparallel β-sheet	17%	19%[f]	13%	10%	13%[f]
Parallel β-sheet	5%		6%	5%	
Turn	19%	8%	18%	26%	23%
Other	27%	36%	32%	24%	19%
Poly(pro)II helix	–	–	–	–	9%

[a] Secondary structure as predicted using the singular value decomposition method (Hennessey et al., 1981) in conjunction with variable selection (Manavalan & Johnson, 1987).

[b] Secondary structure as estimated using the convex constraint algorithm (Perczel et al., 1992) using the basis set as described by Brahms and Brahms (1980).

[c] Secondary structure as estimated using the method of Sreerama and Woody (1993) in conjunction with the basis set as described by Manavalan and Johnson (1987).

[d] Method c in conjunction with the basis set derived from the Kabsch and Sander method for assigning secondary structure from X-ray data (Kabsch & Sander, 1983).

[e] Method d with the inclusion of the poly(Pro)II helix in the deconvolution (Sreerama & Woody, 1994).

[f] This value represents the contribution of anti-parallel and parallel β-sheet.

Table 2. *Analysis of sedimentation equilibrium data*[a]

		SRV[b]						ln K_a[c]		
[p16][d]	Mr[e]	SS[f]	2[g]	1[h]	1 ↔ 2[i]	1 ↔ 3[i]	1 ↔ 4[i]	1 ↔ 2[i]	1 ↔ 3[i]	1 ↔ 4[i]
5 μM	16,100 ± 1,800	5.6	13.5	5.6	5.6	5.6	5.6	—	—	—
15 μM	17,200 ± 900	4.1	20.3	4.6	4.1	4.1	4.2	8.18 ± 0.5	18.0 ± 0.6	28.0 ± 0.75
45 μM	18,800 ± 1,600	4.0	22.3	7.1	3.9	4.0	4.2	8.23 ± 0.3	17.0 ± 0.35	26.0 ± 0.4

[a] Data were collected at 4 °C, scanning at 239 nm for the 5 μM sample and 280 nm for the 15 and 45 μM samples. Data for 5 and 15 μM samples were analyzed using global fits to 15,000, 20,000, 25,000, and 35,000 rpm while 45 μM data were analyzed using a fit to 15,000 rpm. We used the HID program from the Analytical Ultracentrifugation Facility at the University of Connecticut.
[b] Square root of variance (× 10^3).
[c] Natural log of the association constant.
[d] Concentration of protein loaded.
[e] The apparent molecular weight as determined by a single species analysis.
[f] Single species.
[g] Dimer.
[h] Monomer.
[i] Monomer ↔ n-mer equilibria of increasing order up to $n = 4$.

was 42 °C, and the sigmoidal nature of the transition indicates that the thermal unfolding is cooperative (data not shown). In addition, thermal unfolding is largely reversible, suggesting that the thermodynamic stability of p16 can be measured.

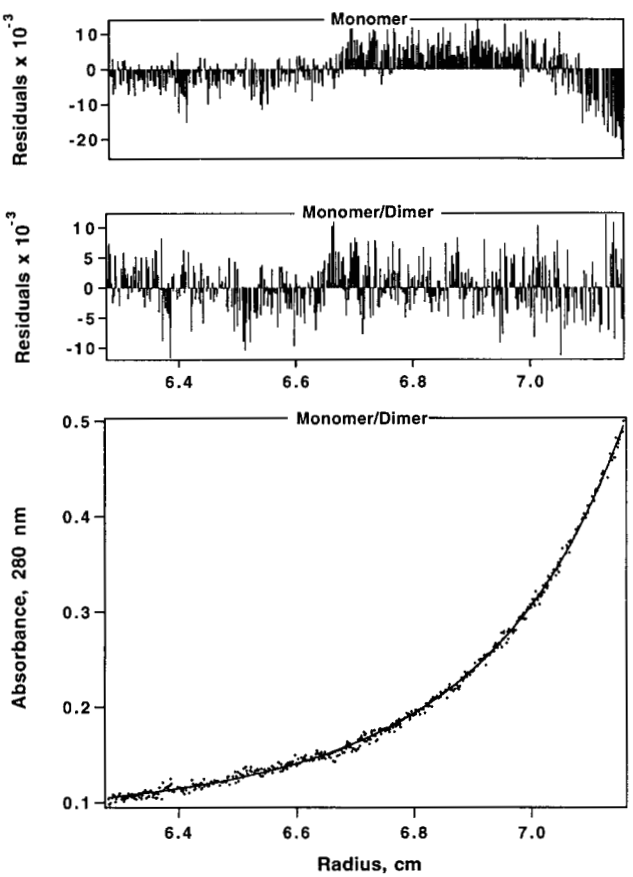

Fig. 4. Sedimentation equilibrium analysis of 45 μM p16. The sample was run at 4 °C in 5 mM Na phosphate, pH 7.4, 100 mM NaCl. Data were collected at 280 nm with a rotor speed of 15,000 rpm. The monomer-dimer fit is shown as well as the residuals comparing the single species monomer with the monomer-dimer residuals.

Thermodynamic stability was measured by chemical, rather than thermal, denaturation using urea as a denaturant since p16 unfolding is fully reversible with urea. The urea denaturation profile (Fig. 5) indicates cooperativity, similar to that seen by thermal unfolding and exhibits low stability with a [urea] denaturation midpoint at 1.73 M.

We fit the urea denaturation data with a two-state function (folded monomer and unfolded monomer) to obtain the associated free energy of stability ($\Delta G = -2.3$ kcal/mol). The fit using such a function is shown in Figure 5, and includes a linear baseline correction for the unfolded state (Fairman et al., 1995).

Two criteria must be met to satisfy the requirements for such a thermodynamic treatment. First, the dimer state cannot be significantly populated; this is the case for the protein concentration used here (9.3 μM) considering a K_d of 270 μM. Second, as described above, the two states must be in reversible equilibrium.

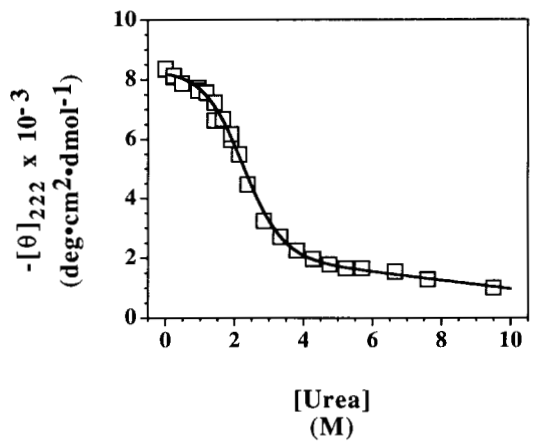

Fig. 5. Urea denaturation of p16. The data points shown as a function of urea concentration with the theoretical fit to a two-state unfolding function. The signal was recorded at 222 nm in 10 mM MOPS, pH 7.5, 50 mM NaF. Data points represent an averaging time of 3 min. The following values were obtained from the fit: transition slope = −1.33 ± 0.061 kcal/mol/M; log K_d = −1.69 ± 0.15; $[\theta]_{folded}$ = −8.32 ± 0.13 deg cm^2/dmol (× 10^3); $[\theta]_{unfolded}$ = −2.39 ± 0.33 deg cm^2/dmol (× 10^3); baseline slope = −0.142 ± 0.049; and a calculated ΔG of −2.3 ± 0.21 kcal/mol.

This is shown by our ability to refold p16 from inclusion bodies and is further supported by a comparison of inhibitory activity to solubly expressed protein (data not shown).

The low stability exhibited by p16 is most likely characterized by an ill-defined tertiary structure. Two explanations are likely: 1) the p16 structure behaves as a molten globule, a protein-folding intermediate; or 2) the p16 structure is highly dynamic. To investigate whether p16 has characteristics of a molten globule, a well-known folding intermediate, 1-anilinonaphthalene-8-sulfonic acid (ANS) binding, was used. ANS, a specific probe for the molten globule state, binds to apolar regions of molten globule proteins while it does not bind with an appreciable affinity to folded or unfolded proteins. The emissions spectrum of ANS is very sensitive to solvent polarity, affecting both the fluorescence intensity and the maximum wavelength. While addition of p16 resulted in a shift in the maximum wavelength and intensity, a binding constant could not be accurately established for the dissociation constant due to the low binding affinity. However, a lower limit was determined with a $K_d > 250$ μM. In addition, molten globule states typically exhibit broad thermal unfolding transitions, whereas thermal unfolding of p16 is highly cooperative. These results, taken together, indicate that p16 does not exhibit characteristics of a molten globule protein. The more likely explanation for the low stability is that p16 is highly dynamic. Fluorescence results are consistent with this conclusion, as the tryptophan fluorescence spectrum is largely unperturbed when comparing p16 under native or unfolded conditions.

We also measured protein dynamics using hydrogen–deuterium exchange rates as monitored by NMR. The rate at which amide protons exchange with solvent deuterons is an indication of the dynamics of a protein as well as the number of residues that are involved in the folding core (Kim et al., 1993). A sample of S-tag-p16 was reconstituted in deuterated H_2O and the NMR spectrum was compared to the NMR spectrum established for an identical sample reconstituted in H_2O. Spectra were collected every 20 min for 16 h. The results showed that that all but six protons were exchanged within 20 min, the time required to acquire the first spectrum. The peaks that did persist were fully exchanged after 3 h. The exchange rate for the measurable peaks is consistent with the ΔG measured by urea denaturation (Kim et al., 1993). These results show that the majority of backbone NH groups of p16 are in very rapid exchange, therefore suggesting that p16 is a highly dynamic protein. The conclusion is that this dynamic character leads to the low stability observed by urea denaturation.

Discussion

We have expressed, refolded, and purified p16 from *E. coli* to study its structural characteristics in order to gain insight into its role as an inhibitor of CDK4 and CDK6 holoenzyme activity. Recombinant, refolded p16 was functionally active, as judged by CDK4 binding (Fig. 1). No difference in this binding activity was observed when compared to soluble p16 expressed either in *E. coli* or in baculovirus (data not shown).

We have used circular dichroism (CD) to determine the secondary structure of p16. Although the overall shape of the CD spectrum is characteristic of proteins that are α-helical, the deconvolution results, as well as the intensity of the band at 222 nm, suggest that only 33% is α-helical and that a significant proportion of β-sheet is present.

It has been suggested that p16 contains four ankyrin-like repeats (Serrano et al., 1993). These motifs, which are found in proteins of diverse function, have been proposed to be involved in specific protein–protein interactions (Michaely & Bennett, 1992). Therefore, the presence of ankyrin-like repeats in p16 would be consistent with its function. To try to identify the position of the units of secondary structure measured by CD, we aligned the amino acid sequence of p16 with the consensus sequence generated from 62 different ankyrin-like repeats from human proteins (Bork, 1993) (Fig. 6A). We generated the human ankyrin consensus sequence using a previously described method based on conservation of sequence (Anthony-Cahill et al., 1992). It is striking to note the similarity of the consensus sequence pattern to that generated by Michaely and colleages (Michaely & Bennett, 1992). Figure 6B shows graphically the extent to which each position's amino acid type is conserved, thus providing a more quantitative intepretation of the consensus sequence.

The sequence conservation between p16 and ankyrin repeats is not striking; however, this is also true of ankyrin repeats themselves. There are two clusters within each repeat that are well conserved. When we examine the amino acid conservation and hydrophobicity plots closely, a weak pattern emerges. There is a periodicity in the two conserved clusters, and this periodicity also occurs in the hydrophobicity plot (Fig. 6B). Highly conserved hydrophobic amino acids are found every three to four residues in these two regions. This periodicity is typical of a canonical helix with one face buried, suggesting that these portions of an ankyrin repeat are amphipathic α-helices. These clusters also show a propensity to be α-helical as shown in the lower panel of Figure 6B. Interestingly, these residues, or residues that are similarly hydrophobic, are highly conserved among the four ankyrin-like repeats of p16. If this prediction were true, one would expect the helix content of p16 to be approximately 35%. The results from the deconvolution of the CD spectrum are in agreement with this value. Additionally, in this alignment there are amino acids within each ankyrin repeat of p16 that can be identified as potential helix starting and terminating residues, for example, proline and glycine residues, respectively.

An attempt was made to assign the fraction of β-sheet to the sequence as well. Some evidence from the sequence analysis points to residues 26–31 of the ankyrin motif as a potential β-sheet. Interestingly, this predicts 16% sheet, which is in good agreement with the CD determined value for anti-parallel β-sheet. Our sequence analysis of the ankyrin consensus and, by extension, the p16 sequence generates a structural model similar to that proposed by Michaely & Bennett (1992).

Since ankyrin repeat proteins are diverse in their function, with their target proteins being quite different, the conservation in sequence most likely reflects important structural elements (Gay & Ntwasa, 1993). It has been suggested that the less conserved positions in the ankyrin repeat sequence are responsible for specificity in protein–protein interactions (Gay & Ntwasa, 1993).

Although ankyrin-like repeats are predicted to be largely helical, the only structural data available are the CD spectra of two different proteins. The first of these, the 85,000 Da membrane-interacting domain of ankyrin, part of a multigene family, contains 24 ankyrin repeats and is approximately 16% helical by CD (Michaely & Bennett, 1993). The second protein, a 53,000 Da *Drosophila* protein called cactus (Gay & Ntwasa, 1993), contains six ankyrin-like repeats and was reported to be 77% helical by CD. Closer examination of the CD spectrum for cactus suggests that the

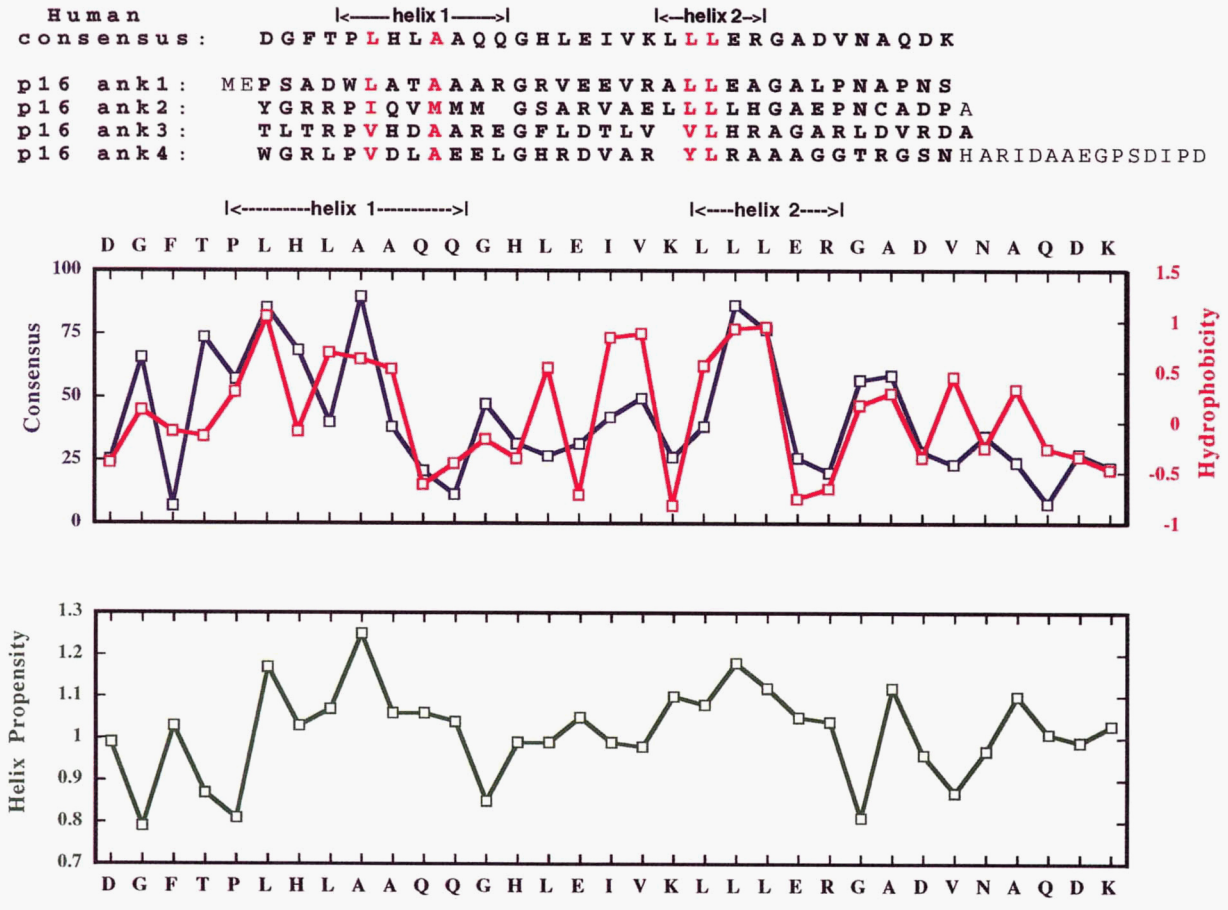

Fig. 6. A: An alignment of the p16 amino acid sequence with the consensus sequence for ankyrin-like repeats. The human consensus was generated using the ankyrin-like repeat sequences of 62 different human proteins using a method developed by Anthony-Cahill et al. (1992). The alignment shows the complete amino acid sequence of p16. The residues in bold letters are those that we have identified as part of the ankyrin-like repeat consensus sequence. Above the human consensus we show two predicted helices based on conservation of sequence, hydrophobicity, and periodicity. The highly conserved hydrophobic residues are shown in red. **B:** The human ankyrin consensus sequence. The sequence conservation (in blue) and hydrophobicity (in red) are overlaid. The helix propensity is shown in green.

helix content reported is inconsistent with the shape and intensity of the this spectrum. Our CD results suggest that the secondary structure of p16 more closely compares to that of the 85,000 Da membrane domain of ankyrin, the first of the ankyrin repeat proteins to be described (Lux et al., 1990). It is clear, however, that more structural information needs to be obtained for a variety of these proteins before any generalizations can be made.

Our urea denaturation and thermal unfolding results, as monitored by CD, indicate that p16 is cooperative in its unfolding. If we assume a two-state model, the ΔG would be on the order of -2.3 kcal/mol, a value smaller than typically associated with a compact, stable structure. Additionally, tryptophan fluorescence experiments suggest that p16 may not have a well-defined tertiary structure (data not shown). Rapid hydrogen-deuterium exchange as monitored by NMR indicates that p16 is a dynamic protein, and it is conceivable that its role in binding CDK4 may require this characteristic.

Although the urea denaturation and thermal unfolding results suggest that the ankyrin repeats of p16 encode a structural domain, we cannot conclude that four ankyrin repeats defines an autonomous folding domain without information about the oligomeric state of the protein. It is not clear from the literature what number of ankyrin repeats does define an autonomous folding domain; it has been suggested that four to seven are minimally required, although there are exceptions to this rule (Michaely & Bennett, 1993). Our sedimentation equilibrium results suggest that p16 is largely monomeric at low micromolar concentrations (the K_d for the dimer is 270 μM). CD spectra of monomeric and dimeric p16 are identical, providing further evidence that four ankyrin repeats are, indeed, sufficient to define an autonomous folding domain.

In order for the observed self-association to be an important aspect of its function, the intracellular protein concentration would have to be significantly higher. There is no information available on the cellular concentration of p16, but even if it is present at monomeric concentrations, it does not rule out cooperativity of binding, that is, binding of one monomer to CDK4 driving the binding of an additional monomer.

Part of the information for p16's tertiary structure may come from its interaction with CDK4. We are currently studying the p16/CDK4 complex to determine if, upon binding, a change occurs

in the biophysical parameters that we have established. There are many examples providing precedence for the coupling of protein binding to local folding (Burley, 1994; Spolar & Record, 1994), and it will be important to find whether p16 falls into this category of induced-fit proteins.

Experimental procedures

Cloning

p16 DNA, encoding amino acids 1–148 (Serrano et al., 1993), was generated by standard PCR methods (Saiki et al., 1988) using Vent DNA polymerase (New England Biolabs). The PCR template (provided by K. Coleman, Bristol-Myers Squibb) was a human cDNA clone isolated from a HeLa cell cDNA library (Clontech) and cloned into the pBC SK phagemid (Stratagene). The open reading frame of this clone included an additional 78 nucleotides upstream of the published sequence. Two PCR primers (Genosys), a 5′ primer (5′-CGTAGGATCCACCATGGAGCCTTCGGCTGACTGG CTGGCCAC) and a 3′ primer (5′ GTGAGGATCCTCAGTCAGT CAATCGGGGATGTCTGAGGGACCTTCCGCGGCATC) were used to generate p16 without the additional upstream nucleotides. These primers also provided *Bam*HI and *Nco*I sites upstream of the start site and a *Bam*HI site downstream of the UGA stop codon at the 3′ end. *Nco*I and *Bam*HI digested PCR product was ligated into the *Nco*I and *Bam*HI sites of pet-29a (Novagen), a vector whose expression is driven by the T7 promoter and which provides an N-terminal S-Tag fusion. Clones were sequenced by dideoxy chain termination (Sanger et al., 1977) to confirm that the sequence matched that of the published p16 sequence.

Expression and purification

We used *E. coli* strain BL21(DE3) for plasmid propagation in M9 minimal media containing casamino acids, trace minerals, and glucose (M9CAG). Kanamycin was added at 25 μg/mL. One liter of media was inoculated with 20 mL of an overnight culture of freshly transformed BL21(DE3). Protein expression was induced by the addition of 0.4 mM IPTG when the cell density, measured at OD_{600}, reached 1.0. Cells were harvested after 4 h and resuspended in 30 mL of Buffer A (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM DTT) including 1X Complete Protease Inhibitor (Boehringer Mannheim). We exposed this suspension to one freeze/thaw cycle at −80 °C and completed cell lysis using a French Press (SLM Instruments, Inc.).

The pellet was recovered after centrifugation at 50,000 × *g* for 20 min. The gelatinous, top layer of the pellet was removed and the remaining inclusion body pellet was resuspended with a Dounce homogenizer (Wheaton) in 30 mL of Buffer A containing protease inhibitor. The inclusion body pellet was recovered by an additional centrifugation step and was solubilized in 30 mL of Buffer A containing protease inhibitor and 8 M urea. The suspension was incubated on a rotator for 1 h at room temperature and then overnight at 4 °C. The sample was then centrifuged at 50,000 × *g* for 20 min to remove any unsolubilized protein. After adjusting the pH to 8.8, the supernatant was fractionated on a 20 mL bed of Q-sepharose anion exchange resin (Pharmacia) equilibrated with Buffer A, pH 8.8, 8 M urea. We eluted the fusion protein with a 400 mL linear gradient of 0 to 300 mM NaCl in Buffer A, pH 8.8, 8 M urea. S-Tag-p16 typically eluted at 100 mM NaCl. Fractions containing S-Tag-p16, as judged by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie staining, were pooled. The pH of the sample was adjusted to 4.8 and dialyzed overnight at 4 °C against 40 volumes of Buffer B (50 mM Na acetate, pH 4.8, 1 mM EDTA, 5 mM DTT, 8 M urea) to lower the NaCl concentration. We passed the sample over a 30 mL bed of S-sepharose cation exchange resin (Pharmacia) equilibrated with Buffer B. S-Tag-p16 was eluted with a 600 mL linear gradient of 0 to 250 mM NaCl in Buffer B. S-Tag-p16 typically eluted at 140 mM NaCl. Fractions were pooled and approximate protein concentrations were determined with the Coomassie Plus protein assay (Pierce). The pH of the sample was then adjusted to 8.4.

Urea solubilized S-Tag-p16 was refolded at a final protein concentration of 35 μg/mL by diluting into Buffer A, pH 8.4 containing 1 M urea and incubating approximately 15 h at 4 °C.

After adjusting the pH to 8.8, the sample was passed over a 10 mL bed of Fractogel EMD anion exchange resin (EM Separations) equilibrated with Buffer A, pH 8.8 containing 1 M urea. We included this step to separate any aggregated protein that may have formed. The column was washed with two column volumes of Buffer A, pH 8.8 without urea. S-Tag-p16 was eluted with a 200 mL linear gradient of 0 to 300 mM NaCl in Buffer A, pH 8.8. S-Tag-p16 typically eluted at 100 mM NaCl. Pure fractions were pooled and concentrated, if necessary, with a 3,000 MW cutoff Centriprep (Amicon) to 3–5 mg/mL. The sample was then dialyzed against 500 volumes of 5 mM Na phosphate, pH 7.5, 100 mM NaCl, 1 mM DTT, with a continuous flow microdialysis chamber (Pierce). Samples prepared for CD were exchanged with 5 mM Na phosphate, pH 7.4, 10 mM NaF immediately before analysis.

Thrombin cleavage

To remove the S-Tag, we cleaved the fusion protein in solution as the site appeared to be inaccessible when bound to S-Protein agarose (Novagen). Cleavage was carried out according to the manufacturer's directions using biotinylated thrombin (Novagen). Complete cleavage occurred after a 12-h incubation at 4 °C as judged by SDS-PAGE. Biotinylated thrombin was removed by adding 100 μL of a 50% slurry of streptavidin agarose (Novagen), incubating the sample on a rotator for 30 min at room temperature and removing the streptavidin agarose by centrifugation. Residual resin was removed by filtering the sample through a 0.45 micron microcentrifuge filter unit (Millipore). A two-step procedure was used to remove free S-Tag. First, the sample was dialyzed against 500 volumes of 5 mM Na phosphate, 100 mM NaCl, 5 mM DTT to decrease the amount of the free S-Tag. We removed any remaining free S-Tag, as well as any trace amounts of uncleaved material, with S-Protein resin according to manufacturer's directions. Again, any residual resin was removed with a 0.45 micron spin filter.

We determined protein concentrations by two or more of the following methods: absorbance at 280 nm in 6 M guanidine-HCl, 10 mM MOPS, pH 7.5 using an extinction coefficient of 13,940 (Edelhoch, 1967), amino acid analysis, or with the Coomassie Plus protein assay using bovine serum albumin as the standard.

Western blot

Western blots were carried out with an ECL chemiluminescence kit according to manufacturer's directions (Amersham) using p16 antibody (Santa Cruz Biotechnology, C-20 antibody).

Amino acid sequencing

Amino acid sequencing was performed by the Princeton University Synthesis/Sequencing Facility.

Immunoprecipitations

The TNT coupled reticulocyte lysate system was used according to manufacturer's directions (Promega) to generate ^{35}S-labeled CDK4 protein. The standard immunoprecipitation (IP) reaction used to assess the ability of purified p16 to interact with CDK4 contained 5 μL of the CDK4 reticulocyte lysate, 5 μg of p16 or S-Tag-p16, and 750 μL of TTBS (0.1 M Tris, pH 7.5, 0.9% NaCl, 0.1% Tween-20) containing 10% nonfat milk. All steps were carried out at 4 °C on a rotator unless otherwise noted. The reactions were incubated for 90 min, after which 50 μL of p16 antibody was added and incubated for 1 h. TTBS (25 μL) equilibrated Protein A sepharose was added and incubated for 1 h. The reactions were centrifuged momentarily at 14,000 × *g* to pellet the resin, and the supernatant was discarded. The resin was washed three times with 200 μL of TTBS at each wash and recovered by centrifugation as above. SDS-PAGE loading buffer (50 μL) was added after the final wash step, the samples were vortexed, boiled for 3 min, vortexed again, and then centrifuged to pellet the resin away from the sample. Binding was assessed by SDS-PAGE and autoradiography. If S-Protein agarose was used to precipitate S-Tag-p16, then the p16 antibody step was eliminated and 25 μL of TTBS equilibrated S-Protein agarose was added in place of Protein A sepharose.

Analytical ultracentrifugation

We carried out sedimentation equilibrium experiments at 4 °C on a Beckman Model XLA analytical ultracentrifuge using an AN-60-Ti rotor (Giebeler, 1992). Most experiments were carried out using the standard 12 mm pathlength, six-channel, charcoal-filled Epon cells with quartz windows. Continuous radial scanning at 239 nm was used for 5 μM samples, while 15 and 45 μM samples were analyzed at 280 nm. Two channel, 3-mm pathlength cells were used for samples at 45 μM, which were scanned at a single speed of 15,000 rpm. The 12-mm pathlength cells were scanned at multiple speeds. The cells were scanned every 0.001 cm, and 50 scans were averaged. The standard buffer was 5 mM Na phosphate, pH 7.4, 100 mM NaCl, 1 mM DTT. The default value of 1.0017 g/mL was used for the density of the solvent, and a partial specific volume of 0.73 mL/g was calculated from the weight average of the partial specific volumes of the individual amino acids (Cohn & Edsall, 1943).

Circular dichroism

All CD spectra were recorded with an Aviv Model 62DS spectropolarimeter equipped with a temperature controller. The sample spectrum, an average of 15 repeats, was recorded at 25 °C in the far UV region between 181 and 250 nm, with a step size of 0.25 nm, at a bandwidth of 1.5 nm, and an averaging time of 2 s. A buffer blank using the same parameters was subtracted to yield the corrected sample spectrum. Spectra were collected in 0.1-mm pathlength quartz cells (Hellma) and the sample buffer consisted of 5 mM Na phosphate, pH 7.4, 10 mM NaF.

For urea denaturation experiments, the signal at 222 nm was recorded at 25 °C for 500 s. The concentration of p16 in these experiments was 9.3 μM. The urea buffer contained 10 mM MOPS, pH 7.5, 50 mM NaF. Samples were incubated for 1 h prior to analysis, and no change in signal was observed when the samples were incubated for an additional 12 h. Data points represent a time average of 3 min. The data were fit applying a specialized algorithm (Fairman et al., 1995) encoded in MLAB (Civilised Software) (Knott, 1979).

Fluorescence

Fluorescence spectra were measured using a Spex Fluorolog-2 spectrofluorometer. Varying concentrations of S-tag-p16 were titrated into 10 μM 1-anilinonaphthalene-8-sulfonic acid (ANS) in a 1 × 10 cm quartz cuvette with an excitation wavelength of 370 nm and an emission wavelength of 450 nm. Slitwidth settings of 0.5 and 2 mm were used for excitation and emissions, respectively.

Amide exchange

Amide exchange was monitored by NMR with S-tag-p16 in 5 mM sodium phosphate, 25 mM NaF, 2 mM DTT, pH 7.0 lyophilized, and reconstituted in an identical volume of cold H_2O or cold D_2O. A reference 1D proton spectrum was taken in H_2O buffer at 20.0 °C. The water resonance was suppressed by the non-saturative WATERGATE technique (Piotto et al., 1992). Within 25 min of the resuspension, data collection had begun at 20.0 °C. Forty-eight individual 1D proton spectra were collected in a continuous fashion over a total time of 16 h; each spectrum took 20 min to collect. These spectra were collected in the same manner and under the same conditions as was the reference spectrum.

Acknowledgments

We thank Sandy Farmer for the NMR experiment monitoring amide exchange and for discussions involving NMR; Kevin Coleman, Barri Wautlet, and Pam Brinkley for the initial cDNA clone, for the CDK4 DNA used for immunoprecipitations, and for comparing our recombinant p16 to baculovirus expressed p16 in a CDK4 binding assay; Kevin Webster and Janet Mulheron for performing the kinase inhibition assay; Clifford Klimas for amino acid analysis; and Mark Hail for mass spectrometry analysis. We thank Norma Greenfield for providing the CD deconvolution package including her programming interface and the algorithms used in this study and Preston Hensley for the Iger-based nonlinear fitting algorithm used to generate the sedimentation equilibrium figure. We also thank Thom Lavoie, Gary Matsueda, and Jannette Carey for helpful advice and discussions during the course of this work.

References

Anthony-Cahill SJ, Benfield PA, Fairman R, Wasserman ZR, Brenner SL, Stafford III WF, Altenbach C, Hubbell WL, DeGrado WF. 1992. Molecular characterization of helix-loop-helix peptides. *Science 255*:979–983.

Bork P. 1993. Hundreds of ankyrin-like repeats in functionally diverse proteins: Mobile modules that cross phyla horizontally? *Proteins: Struct Funct Genet 17*:363–374.

Brahms S, Brahms J. 1980. Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J Mol Biol 138*:149–178.

Burley SK. 1994. Plus ca change, plus c'est la meme chose. *Struct Biol 1*:207–208.

Chellappan SP, Hiebert S, Mudryj M, Horowitz JM, Nevins JR. 1991. The E2F transcription factor is a cellular target for the Rb protein. *Cell 65*:1053–1061.

Cohn EJ, Edsall JT. 1943. In: *Proteins, amino acids and peptides as ilons and dipolar ions.* New York: Reinhold Publishing Corporation. pp 370–377.

Draetta GF. 1994. Mammalian G_1 cyclins. *Curr Opin Cell Biol 6*:842–846.

Edelhoch H. 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry 6*:1948–1954.

Elledge SJ, Harper JW. 1994. Cdk inhibitors: On the threshold of checkpoints and development. *Curr Opin Cell Biol 6*:847–852.

Fairman R, Chao H-G, Mueller L, Lavoie TB, Shen L, Novotny J, Matsueda G. 1995. Characterization of a new four-chain coiled-coil: Influence of chain length on stability. *Protein Sci 4:*1457–1469.

Gay NJ, Ntwasa M. 1993. The *Drosophila* ankyrin repeat protein cactus has a predominantly α-helical secondary structure. *FEBS Lett 335*:155–160.

Giebeler R. 1992. The Optima XL-A: A new analytical ultracentrifuge with a novel precision absorption optical system In: SE Harding, AG Rowe, JC Hortons, eds. *Analytical uUltracentrifugation in biochemistry polymer science*. Cambridge: Royal Society of Chemistry. pp 16–25.

Hennessey JP, Johnson WC Jr. 1981. Information content in the circular dichroism of proteins. *Biochemistry 20*:1085–1094.

Hiebert SW, Chellappan SP, Horowitz JM, Nevins JR. 1992. The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. *Genes Dev 6*:177–185.

Hunter T, Pines J. 1994. Cyclins and cancer: Cyclin D and Cdk inhibitors come of age. *Cell 79*:573–582.

Kabsch W, Sanders C. 1983. Dictionary of protein secondary structures: Pattern recognition of hydrogen-bonded and geometric features. *Biopolymers 22:* 2577–2637.

Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavtigan SV, Stockert E, Day RSI, Johnson BE, Skolnick MH. 1994. A cell cycle regulator potentially involved in the genesis of many tumor types. *Science 264*:436–440.

Kim K-S, Fuchs JA, Woodward, CK. 1993. Hydrogen exchange identifies native-state motional domains important in protein folding. *Biochemistry 32*:9600–9608.

Knott GD. 1979. MLAB—A mathematical modeling tool. *Comput Programs Biomed 10*:271–280.

Lux S, John K, Bennett V. 1990. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue differentiation and cell cycle control proteins. *Nature 344*:36–42.

Manavalan P, Johnson WCJ. 1987. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal Biochem 167*:76–85.

Michaely P, Bennett V. 1992. The ANK repeat: A ubiquitous motif involved in macromolecular recognition. *Trends Cell Biol 2*:127–129.

Michaely P, Bennett V. 1993. The membrane-binding domain of ankyrin contains four independently folded subdomains, each comprised of six ankyrin repeats. *J Biol Chem 268*:22703–22709.

Moulton T, Samara G, Chung W-Y, Yuan L, Desai R, Sisti M, Bruce J, Tycko B. 1995. MTS1/p16/CDKN2 lesions in primary glioblastoma multiforme. *Am J Pathol 146*:613–619.

Nevins JR. 1992. E2F: A link between the Rb tumor suppressor protein and viral oncoproteins. *Science 258*:424–429.

Nobori T, Miura K, Wu DJ, Lois A, Takabayashi K, Carson DA. 1994. Deletion of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. *Nature 368*:753–756.

Okamoto A, Hussain SP, Hagiwara K, Spillare EA, Rusin MR, Demetrick DJ, Serrano M, Hannon GJ, Shiseki M, Zariwala M. 1995. Mutations in the *p16^{INK4}/MTS1/CDKN2, p15^{INK4B}/MTS2,* and *p18* genes in primary and metastatic lung cancer. *Cancer Res 55*:1448–1451.

Perczel A, Park K, Fasman GD. 1992. Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: A practical guide. *Anal Biochem 203*:83–93.

Piotto M, Saudek V, Sklenar V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR 2*:661–665.

Ranade K, Hussussian CJ, Sikorski RS, Varmus HE, Goldstein AM, Tucker MA, Serrano M, Hannon GJ, Beach D, Dracopoli NC. 1995. Mutations associated with familial melanoma impair p16^{INK4} function. *Nat Genet 10*:114–116.

Richards FM, Vithayathil PJ. The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components. 1955. *J Biol Chem 234*:1459–1465.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science 239*:487–491.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA 74*:5463.

Serrano M, Hannon GJ, Beach D. 1993. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature 366*:704–707.

Shapiro GI, Edwards CD, Kobzik L, Godleski J, Richards W, Sugarbaker DJ, Rollins BJ. 1995. Reciprocal Rb inactivation and p16^{INK4} Expression in primary lung cancers and cell lines. *Cancer Res 55*:505–509.

Sherr CJ. 1994. Phase progression: Cycling on cue. *Cell 79*:551–555.

Sherr CJ. 1995. D-type cyclins. *Trends Biochem Sci 20*:187–190.

Spolar RS, Record MT Jr. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science 263*:777–784.

Sreerama N, Woody RW. 1993. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem 209*:32–44.

Sreerama N, Woody RW. 1994. Poly(Pro)II helices in globular proteins: Identification and circular dichroic analysis. *Biochemistry 33:*10022–10025.

Washimi O, Nagatake M, Osada H, Ueda R, Koshikawa T, Seki T, Takahashi T, Takahashi T. 1995. In vivo occurences of p16 (MTS1) and p15 (MST2) alterations preferentially in non-small cell lung cancers. *Cancer Res 55*:514–517.

Weinberg RA. 1995. The retinoblastoma protein and cell cycle control. *Cell 81*:323–330.

Yang R, Gombart AF, Serrano M, Koeffler HP. 1995. Mutational effects on the p16^{INK4a} Tumor suppressor protein. *Cancer Res 55*:2503–2506.

Yang JT, Wu CSC, Martinez HM. 1986. Calculation of protein conformation from circular dichroism. *Methods Enzymol 130*:208–225.