AD_____

Award Number: DAMD17-98-1-8270

TITLE: Structural Basis of CDK4 Inhibition by $p18^{INK4C}$

PRINCIPAL INVESTIGATOR: Ravichandran N. Venkataramani
Ronen Marmorstein, Ph.D.

CONTRACTING ORGANIZATION: The Wistar Institute
Philadelphia, Pennsylvania 19104

REPORT DATE: May 2001

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2001	Annual Summary (1 May 98 - 30 Apr 01)

4. TITLE AND SUBTITLE

Structural Basis of CDK4 Inhibition by $p18^{INK4C}$

5. FUNDING NUMBERS

DAMD17-98-1-8270

6. AUTHOR(S)

Ravichandran N. Venkataramani
Ronen Marmorstein, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Wistar Institute
Philadelphia, Pennsylvania 19104
Email - venkatar@mail.med.upenn.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Report contains color graphics

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

Our research primarily revolves around the study of INK4 proteins. INK4 proteins are key regulators of cell cycle progression acting at the G1-S phase transition. INK4 proteins act upstream of pRb tumor suppressor and thereby, control the expression of S-phase specific genes. We have determined the structure of $p18^{INK4c}$ to 1.95 Å. We present the structure here and its implications for ankyrin repeat proteins and INK4 protein function. Based on the insight gained from our structure of $p18^{INK4c}$, we have designed several mutants to study the impact of the mutation of specific residues on the efficacy of the INK4 proteins. We have determined that atleast three mutants - F71N, F82Q, F92N have significantly higher thermostability than the wild-type protein. We present the 3D structures of these proteins as well as their efficacy *in vivo* in this report.

14. SUBJECT TERMS

Breast Cancer, TuMor Suppressor, Oncogene, $p18^{INK4C}$

15. NUMBER OF PAGES

61

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Cover.....1

SF 298.....2

Table of Contents.....3

Introduction.....4

Body.....4

Key Research Accomplishments.....17

Reportable Outcomes.....18

Conclusions.....19

References.....20

Appendices.....22

Introduction

Progression through the cell cycle is monitored at the G1-S phase checkpoint by active complexes between cyclin dependent kinases (CDK) 4 and 6 and the D-type cyclins (D1, D2, D3). The INK4 family of proteins play a key role in inhibiting the G1-S phase cell cycle transition by specifically inhibiting the kinase activity of CDK4-Cyclin D and CDK6-Cyclin D complexes. The INK4 (Inhibitors of CDK4) family of proteins consists of four known members: $p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$, $p19^{INK4d}$, which share 40% sequence identity overall and structural homology[1-5]. The INK4 family members have indistinguishable CDK/cyclin inhibitory activity when assayed *in vitro*[6], but appear to have distinct functions *in vivo*. For example, the locus on 9p21 (MTS1) expressing $p16^{INK4a}$ is a frequent target of genetic alterations in cancer, while the genes encoding the other INK4 proteins are less commonly mutated in cancer. Specifically, gene deletions in MTS1 and promoter methylation often lead to transcriptional silencing, and point mutations are frequently associated with several different types of cancer including breast carcinomas [4,7-10]. Nearly 55 different residues are targeted in missense mutations of $p16^{INK4a}$ (10). Homozygous deletions of $p15^{INK4b}$ have also been identified in a more limited number of cancer cell lines[11], and point mutations of $p18^{INK4c}$ have been associated with some breast carcinomas[12]. Homozygous deletions of the $p18^{INK4c}$ are also observed in oligodendrogliomas[13], however, these deletions are rare. Polymorphisms of the $p19^{INK4d}$ gene are found in a small percentage of osteosarcomas[14]. Together, the data on tumor-derived mutations suggest a strong correlation between the INK4 proteins and carcinogenesis while also suggesting that the indiviual family members might have tissue-specific, and/or non-redundant activities[15]. Thus, the study of the structural mechanisms underlying the Cyclin D1-CDK4 inhibitory action of the INK4 family of tumor suppressors could lead to valuable insights into the molecular mechanisms underlying tumorigenesis. This understanding could lead to the design of small molecule drugs that can be used in the treatment of INK4-mediated cancers which include breast carcinomas.

Body

The specific aims outlines in the approved statement of work are

1. Determine the structure of $p18^{INK4c}$
2. Prepare mutants of $p18^{INK4c}$ defective in CDK4/6 inhibition
3. Determine the X-ray crystal structure of $p18^{INK4c}$ mutants defective in CDK4/6 inhibition
4. Prepare CDK4-$p18^{INK4c}$ and/or CDK6-$p18^{INK4c}$ complex for structure determination

To briefly summarize our accomplishments, we have determined the crystal structure of $p18^{INK4c}$ to 1.95 Å (Aim #1)[16]. Other research groups pre-empted our efforts to determine the INK4-CDK complex structure (Aim #4)[17,18]. However, based on the insights gained from the crystal structures of free and complexed $p18^{INK4c}$, we have designed and tested several mutants to test the importance of particular residues and the thermostability of INK4 proteins in mediating cell cycle inhibition (Aim #2). We have engineered more thermostable mutants of $p18^{INK4c}$ and have determined the crystal structures of some of these mutants to understand the basis of this increased thermostability (Aim #3). From these studies we find a good correlation between the increased thermostability of $p18^{INK4c}$ to its efficacy as a cell cycle inhibitor[Venkataramani, 2001 #141].

Specific Aim # 1: Determine the structure of $p18^{INK4c}$

Introduction

INK4 proteins share a high degree of sequence identity with each other (nearly 40%). Moreover, they are entirely made up of ankyrin repeats (four repeats in $p16^{INK4a}$ and $p15^{INK4b}$, five repeats in $p18^{INK4c}$ and $p19^{INK4d}$). Ankyrin repeats are about 33 amino acids long and occur in at least four consecutive copies. They are involved in protein-protein interactions. The core of the repeat seems to be a helix-loop-helix structure. They are found in more than 1000 proteins and found in organisms ranging from viruses to humans[19,20]. The crystal structure of $p18^{INK4c}$ will allow us not only to explain the function of INK4 proteins but also help us understand the ankyrin repeat architecture.

Structure determination

The $p18^{INK4c}$ protein was expressed and purified recombinantly from *E.coli*. The protein was crystallized at 4°C. The structure was determined using multiple isomorphous replacement (MIR) techniques using phases from three heavy atom derivatives – Selenomethionine, Mercury and Platinum derivatives. An initial model was built and refined to a final R-factor 20.5 % of and R_{free} of 28.4 %. The results were published[16] and the coordinates deposited in the PDB database (Accession Code: 1IHB).

Structure of $p18^{INK4c}$

The crystal structure of $p18^{INK4c}$ reveals an elongated molecule comprised of five ankyrin repeats. Each ankyrin repeat contains a β-strand loop helix-loop-helix extended loop β-strand motif. Each ankyrin repeat forms a N- terminal and C-terminal face that mediate van der Waals interactions with the neighboring repeats. The neighboring repeats also interact with each other through non-conserved hydrogen bonds in the β-strand regions. The ankyrin repeats stack on top of each other to form an elongated overall structure in which the helical region is along one side of the protein and the β-strand region is along the opposite side. The extended structure provides a large surface area for mediating protein-protein interactions. There are several tight turns within each ankyrin repeat. Two 90° turns occur at the transition from loop to helix1 or helix 2 to extended loop. Two 180° turns occur at the transition from one repeat to the next (between the β-strand 2 of the first repeat and the β-stand 1 of the second repeat) and between the two helices of each repeat.

Implications of the Structure of $p18^{INK4c}$ for Ankyrin Repeat Proteins

The structure of $p18^{INK4c}$ explains the role played by residues conserved in all ankyrin repeats. Generally, ankyrin repeats are poorly conserved and tolerate many substitutions in their consensus sequence. However, we have found that most of the conserved residues play important roles in stabilizing the ankyrin repeat structure.

1. **Residues mediating tight turns in the structure:** Ankyrin consensus positions 5 and 25 mediate the 90° turns in the structure, typically the residues in this position are glycines or small hydrophobic residues that can adopt a positive ϕ and ψ torsion angles. Similarly, positions 2 and 13 that mediate the 180° turns in the structure also typically have glycine ensuring a residue that can adopt a positive ϕ and ψ torsion angles. Finally, residues at position 15 mediating the transition into the second helix, typically has small hydrophobic residues that can mediate the turn.

2. **Residues involved in hydrogen bonding:** The histidines at position 7 in helix 1 hydrogen bonds with backbone NH of residues in the β-strand regions that stabilizes the β-turn regions. Aspartates at position 32 terminate the second β-strand by hydrogen bonding to the residue in position 1 of the next repeat. This interaction also stabilizes the β-turn region of the ankyrin repeat structure.
3. **Residues involved in van der Waals interactions:** Several residues mediate van der Waals interactions between ankyrin repeats. However, residues at positions 6,9,10,17,18,21 and 22 on the two helices are invariably involved in these interactions. These residues are located on the two helical faces that mediate van der Waals interactions in the interior of the protein or with the N- and C-terminal repeats. The residues at these positions are typically small hydrophobic residues.

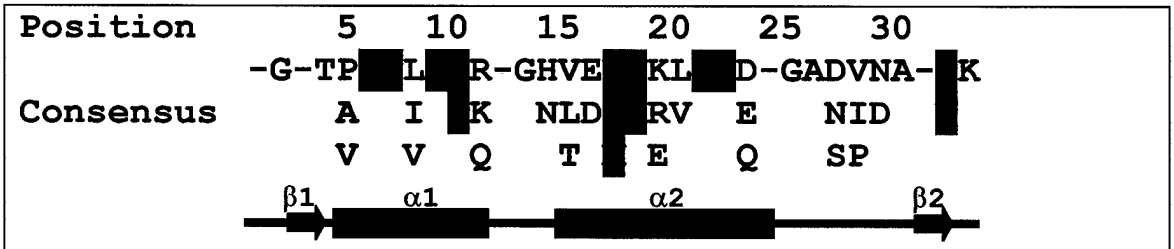


Figure 1 The ankyrin repeat consensus[21] highlighting the conserved residues that have specific roles.
The residues highlighted in yellow are residues involved in mediating steric requirements, residues in purple are residues involved in hydrogen bonding interactions. Finally, the residues highlighted in red are residues involved in mediating van der Waal interactions. The corresponding secondary structure is shown below the consensus.

Thus, 19 out of 33 residues in each repeat are available for protein-protein interaction. However, in reality some of these 19 residues from each repeat will be involved in stabilizing the structure by providing additional non-conserved van der Waals and hydrogen bonding interactions. The remainder form the interaction surface. Since, up to a third of the residues from each repeat can be involved in protein-protein interaction, ankyrin repeats offer an extensive surface area for macromolecular interaction.

Implications of the structure of $p18^{INK4c}$ for the INK4 proteins

Sequence alignment of INK4 proteins reveals an overall high degree of homology and sequence identity. Moreover, our biophysical characterization of $p18^{INK4c}$, $p16^{INK4a}$ and $p19^{INK4d}$, revealed that they have similar composition of secondary structure. Based on this sequence and structural homology, we studied the roles $p16^{INK4a}$ tumor derived mutations and the INK4 conserved residues.

Mapping of $p16^{INK4a}$ tumor derived mutations onto the structure of $p18^{INK4c}$ reveals that most of these mutations localize to residues involved in ankyrin repeat conformation or inter-repeat interaction. This suggests that a large percentage of $p16^{INK4a}$ mutations result in reduced $p16^{INK4a}$ stability. Indeed a majority of these mutations studied *in vitro* reveals that the mutated proteins are non-functional and less stable (prone to aggregation)[12,22-33]. However, a subset of these mis-sense mutations are found to be clustered on the surface around the α3 helix and α5-turn-α6 regions of the structure (Figure 2, Page 11).

The structure of $p18^{INK4c}$ shows that about 80% of these INK4 conserved residues play a role in either stabilizing the ankyrin repeat structure or facilitating complementary interactions between neighboring repeats and therefore, buried inside the protein. However, the residues that are solvent exposed in the $p18^{INK4c}$ molecule are clustered around the α3 helix and the α5-turn-α6 region (Figure 3, Page 12). This correlates well with the surface cluster of $p16^{INK4a}$ mutations suggesting that the residues in these regions form the CDK binding surface.

These predictions were confirmed when the complex structures of CDK-INK4 proteins became available [17,18,34]. These structures confirmed that the regions of the α3 helix and the α5-turn-α6 regions were involved in CDK6 binding. Moreover, the residues involved in binding CDK6 correlate well with those predicted to be involved [35]. These residues include Ala 14, Arg 15, Asn 35, Val 44, Lys 46, Asp 67, Asp 76, Arg 79, Ala 80, Phe 82, and Asp 84 (Figure 3, Page 12).

Aim 2: Prepare mutants of $p18^{INK4c}$ defective in CDK4/6 inhibition
Aim 3: Determine the X-ray crystal structure of $p18^{INK4c}$ mutants defective in CDK4/6 inhibition

Due to the relatedness of Aims 2 and 3, the experiments designed to investigate these aims and their result are reported together here. **The rationale behind specific aims 2 and 3 as stated in the original proposal were to identify the importance of residues involved in CDK4/6 binding and inhibition through mutagenesis.** Originally, we planned on making mutants of $p18^{INK4c}$ defective in CDK4/6 inhibition. However, based on the structure of $p18^{INK4c}$ (16) and the structure of the complex[34], we realized that we might be able to engineer mutations that will increase the efficacy of the INK4 proteins. **This mutational analysis will further the understanding of INK4 mediated inhibition of CDK4 and CDK6 more than just the testing of defective mutants.** The design of these mutations and the results are reported here.

Introduction

The structure of the complexes of CDK-INK4 show that the INK4 proteins do not undergo a conformational change (while inflicting a drastic confromational change in the CDK) [34]. This suggests that the stability and the rigidity of the INK4 proteins are important for their function. Indeed, structure-function studies of the INK4 proteins have revealed that the vast majority of tumor-derived $p16^{INK4a}$ mutations reduce the thermodynamic stability of $p16^{INK4a}$ suggesting a correlation between reduced thermodynamic stability and decreased cell cycle inhibitory activity of INK4 proteins. In order to understand the correlation between thermostability and activity of INK4 proteins, we designed several mutants to study the effect of thermostability on the structure and function of INK4 proteins.

Rationale for design of $p18^{INK4c}$ mutants

In total, we designed 9 single site mutations in $p18^{INK4c}$. Table 1 (Page 13) lists the residues targeted for the mutations, the rationale for the mutation, and the predicted change in thermostability. Of the 9 mutations, six were predicted to increase thermostability, these include, W5R, F37H, R55V, F71N, F82Q, and F92N. Of these six mutations, three (F37, F71, F82) are residues that interact with CDK [34]. The other three mutations (H75F, H108L, T85) were

predicted to decrease the thermostability of the $p18^{INK4c}$ proteins, serving as controls for the experiments.

The selection of the mutations was based on the assumption that the following changes would increase the stability of the $p18^{INK4c}$ protein:

1. Changing exposed hydrophobic residues to hydrophilic residues
2. Changing buried hydrophilic residues to hydrophobic residues, and
3. Changing residues at the interface of neighboring ankyrin repeats to increase inter-ankyrin repeat interaction (by introducing hydrogen bonds or van der Waals interactions)

Purification, Biophysical Characterization, Crystallization and Structure Determination of $p18^{INK4c}$ mutants

The $p18^{INK4c}$ mutants were purified in a similar fashion as the native protein[35,36]. However, we were unable to purify one mutant H75F. The thermostability of the native and mutant $p18^{INK4c}$ proteins was determined through Circular Dichroism monitored thermal denaturation[35,36]. Three of the more thermostable $p18^{INK4c}$ mutants were targeted for crystallization. These were F71N, F82Q, F92N. These mutant proteins crystallized isomorphous to the native protein and their structure was determined by examining difference fourier maps[35,36]. The mutant proteins were tested *in vivo* with transient transfections of U2OS cells and studying their effect on cell cycle progression at 24 hours[35,36].

Thermodynamic Stability of $p18^{INK4c}$ mutants

The melting temperature and the free energy of unfolding (ΔG_{UF}) are reported in Table 2 (Page 14). We also report the change in free energy of unfolding of the mutants relative to the native protein, ($\Delta\Delta G_{UF}$). Of the 8 mutants we purified, 5 were more thermostable than the native protein while 3 was thermodynamically less stable. The more stable $p18^{INK4c}$ mutants, there were two distinct classes. The first class of $p18^{INK4c}$ mutant proteins includes F71N, F92N, and F82Q with melting temperatures that are approximately 4°C higher than the native protein. On a relative scale, these mutants are the most stable class of $p18^{INK4c}$ mutants. The second class of thermostable proteins includes F37H and R55V with melting temperatures of 0.6-0.8°C higher than the native protein. The less stable mutants contains W5R, T85F and H108L with melting temperatures from 4°C to 17°C less than the native protein.

The $\Delta\Delta G_{UF}$ calculated from the CD thermal denaturation data reveals that the most stable class of $p18^{INK4c}$ mutants (F71N, F92N, and F82Q) have $\Delta\Delta G_{UF}$ of 3 kcal mol^{-1} more stable than the native protein. The next group of stable mutants (F37H and R55V) are approximately 1 kcal mol^{-1} more stable than the native protein and the less stable class of $p18^{INK4c}$ mutants is approximately 2 to 6 kcal mol^{-1} less stable than the native protein.

Structural and Functional Characterization of F71N, F82Q, and F92N

F71N

The 2.25 Å crystal structure reveals that asparagine 71 directly hydrogen bonds with arginine 79 in the α5 helix in the middle of the ankyrin repeat and also makes a water mediated hydrogen bond with Asp 100 in the β5 strand at the end the ankyrin repeat (Figure 4*a* and *b*). Taken together, the F71N $p18^{INK4c}$ mutation allows new intra-ankyrin repeat interactions that stabilize the tight turns in the structure resulting in the increased stability of the mutant protein.

Cell cycle assays *in vivo* show that the F71N $p18^{INK4c}$ mutant is also a more active cell cycle inhibitor (Figure 5). Indeed, this mutant protein is the best cell cycle inhibitor of those tested in this study. The structure of the $p18^{INK4c}$ /CDK-cyclin complex reveals that although phenylalanine 71 does not interact with CDK, it is nonetheless at the binding interface. Therefore, it is possible that an asparagine substitution at this position would also introduce favorable $p18^{INK4c}$-CDK contacts. Modeling studies suggests that an asparagine at position 71 would be in position to interact with a backbone NH of CDK (glycine 36 of CDK6).

F82Q

The 2.0 Å crystal structure of the mutant reveals that the glutamine at position 82 hydrogen bonds to the backbone nitrogen of glycine 48 on the turn between the α3 and α4 helices of ankyrin repeat 2. The glutamine mutant also makes a water-mediated hydrogen bond to arginine 117 in helix α8 in ankyrin repeat 4 (Figure 4*c* and *d*). Therefore, the glutamine mutation in position 82 of $p18^{INK4c}$ appears to increase the thermodynamic stability of the $p18^{INK4c}$ protein by increasing inter-ankyrin repeat interaction.

Our functional studies of the F82Q $p18^{INK4c}$ mutant protein reveals that, in addition to being thermodynamically more stable, it is also a better cell cycle inhibitor than the native protein when analyzed *in vivo* (Figure 5). The structure of the INK4/CDK complexes reveals that phenylalanine 82 makes a van der Waals contact with the CDK. Indeed, the fact that this residue is strictly conserved within the INK4 family supports its functional importance. The observation that the F82Q $p18^{INK4c}$ mutant has increased cell cycle inhibitory activity relative to the native protein suggests that the glutamine residue must make somewhat compensatory interactions with CDK. Possibly, the aliphatic region of the glutamine may mimic the van der Waals interaction that is mediated by the native phenyalanine residue. In addition, modeling studies suggests that glutamine 82 is in position to make a backbone NH hydrogen bond with the CDK (Ser 155 of CDK6).

F92N

The 2.0 Åcrystal structure of the mutant reveals that asparagine 92 participates in a water-mediated hydrogen bond with arginine 54 that is located on the α4 helix on the second ankyrin repeat (Figure 4*e* and *f*). In addition, this new interaction might stabilize the turn from α-helix 6 to the adjacent loop. Therefore, the mutation F92N appears to facilitate new inter-ankyrin repeat interactions that stabilize a tight turn in the structure resulting in the increased stability of the mutant.

Functional analysis of the F92N $p18^{INK4c}$ mutant protein reveals that it is only marginally more active than the native proteins in cell-cycle inhibition *in vivo* (Figure 5). Interestingly, unlike the F71N and F82Q $p18^{INK4c}$ mutants that were in position to mediate interactions with CDK, the F92N mutation is unlikely to effect interaction with CDK since it is located too far from the binding interface. Presumably, the increased thermodynamic stability alone of the F92N $p18^{INK4c}$ mutant is not sufficient for increased cell cycle inhibitory activity *in vivo*.

Structural and Functional Characterization of other mutants

The structural and functional characterization of the other mutants have been described in detail in the attached manuscript[36].

Conclusions

We have used a structure-based approach to successfully prepare $p18^{INK4c}$ proteins containing single site mutations that are thermodynamically more stable *in vitro,* and that are more potent cell cycle inhibitors *in vivo* than the native protein. In general, we find a strong correlation between increased thermostability *in vitro* and the increased potency of cell cycle inhibition *in vivo*. However, we have also identified some interesting exceptions.

In this study, several factors that can play a role in the *in vivo* cell cycle activity of the $p18^{INK4c}$ mutant proteins are not controlled for, including their relative abilities to interact with CDK and their half-life *in vivo*. Therefore, it is not surprising that there is not a strict correlation between the degree of thermodynamic stability *in vitro* and the ability to inhibit cell cycle progression *in vivo* by the $p18^{INK4c}$ mutant proteins. Indeed, the importance of some of these other factors is supported by our finding that between thermodynamic stability *in vitro* and cell cycle inhibitory activity *in vivo* of two of the nine $p18^{INK4c}$ mutant proteins that we prepared (W5R and H75F) are not correlated.

Nonetheless, this study does show that the thermodynamic stability of $p18^{INK4c}$ can be exploited to create more active cell cycle inhibitory proteins, and that a structure-based approach can be used to facilitate this process. Indeed, we are now in a position to combine $p18^{INK4c}$ mutants that show increased thermodynamic stability and cell cycle inhibitory activity to create even more stable and active molecules.

Aim #4 Prepare CDK4-$p18^{INK4c}$ and/or CDK6-$p18^{INK4c}$ complex for structure determination

Other research groups completed specific aim #4 of our original statement of work[17,18,34].

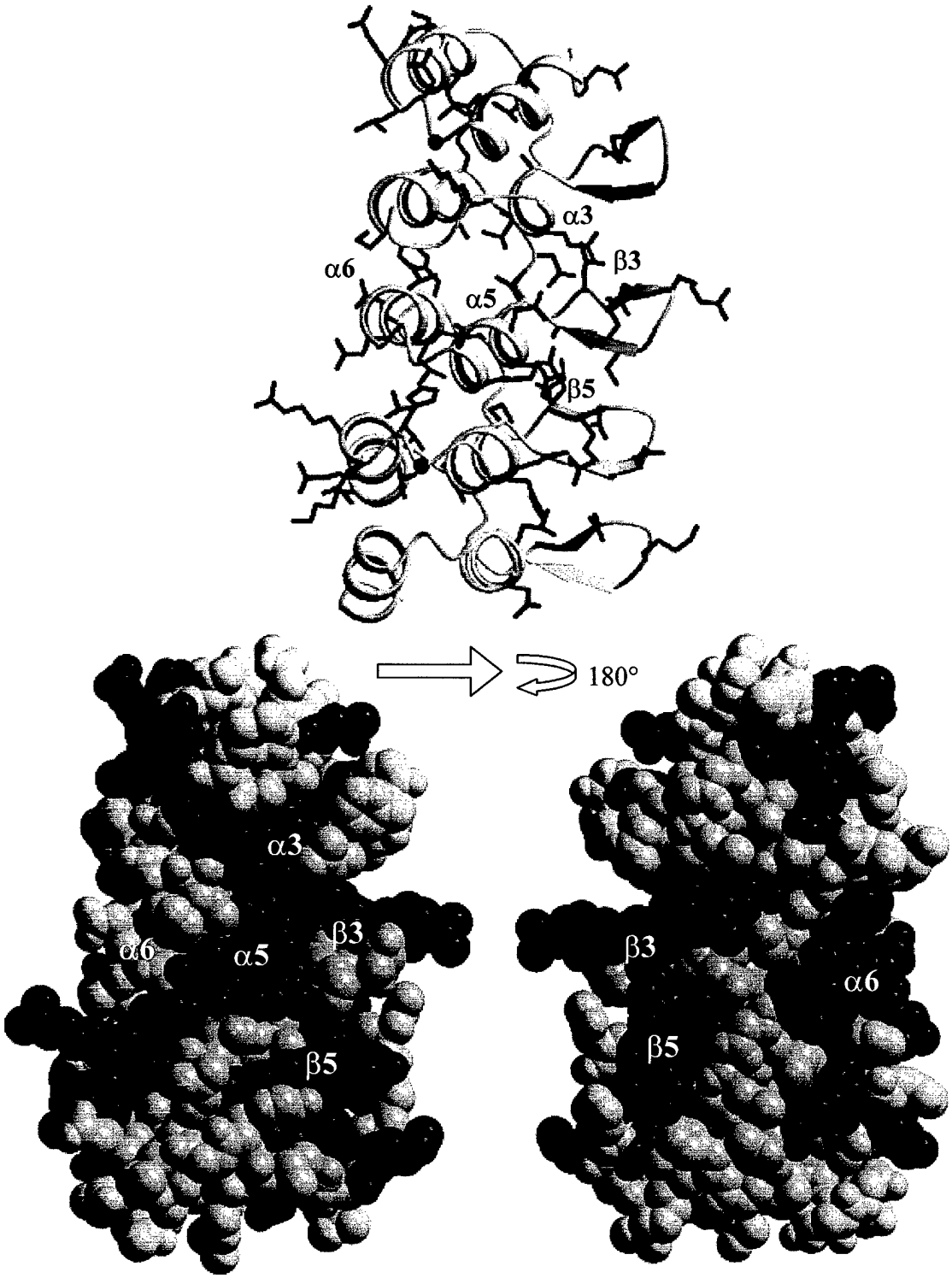


Figure 2 Mapping of tumor-derived mutations of $p16^{INK4a}$ highlights residues clustered around the α3, α5, α6 helices

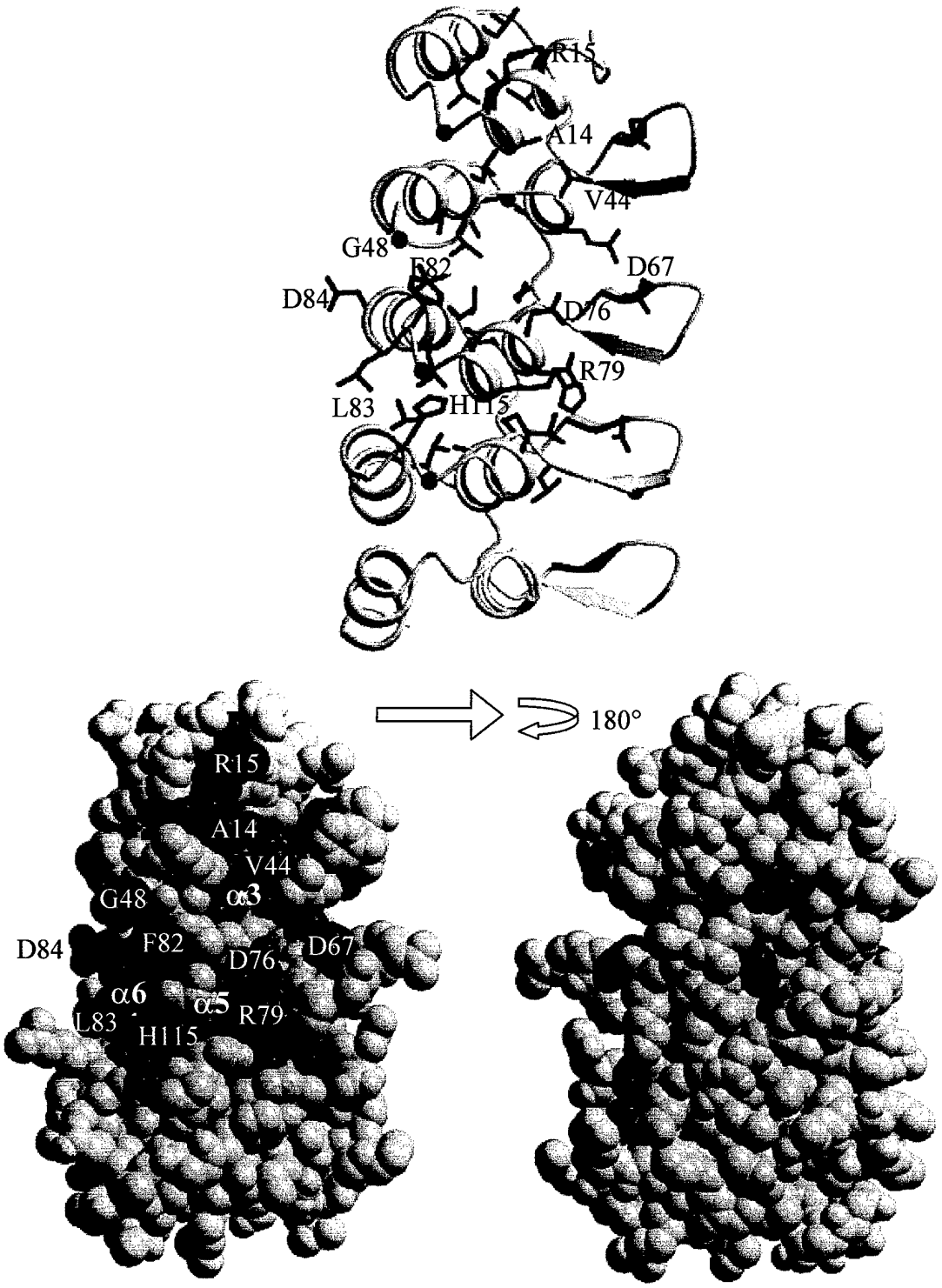


Figure 4 Mapping of INK4 conserved residues on to the structure of $p18^{INK4c}$ reveals several residues on the surface, including clusters around the α3, α5, and α6 helices.

Table 1 Properties of Residues Targeted for Mutations

Residue	Rationale for mutation	Mutation	Predicted change in thermostability	Interaction with residues in CDK6[1]	Conservation among INK4 proteins
Trp 5	Exposed Hydrophobic	Arg	Increase	No	No
Phe 37	Exposed Hydrophobic	His	Increase	Yes, non-polar side chain contacts	Yes
Arg 55	Buried Hydrophilic	Val	Increase	No	No
Phe 71	Exposed Hydrophobic, Possible inter-repeat contacts	Asn	Increase	Yes, non-polar side chain contacts	No
His 75	Buried polar	Phe	Decrease	No	Yes
Phe 82	Exposed Hydrophobic, Possible inter-repeat contacts	Gln	Increase	Yes, non-polar side chain contacts	Yes
Thr 85	Buried polar	Phe	Decrease	No	Yes
Phe 92	Exposed Hydrophobic, Possible inter-repeat contacts	Asn	Increase	No	No
His 108	Buried polar	Leu	Decrease	No	Yes

[1] Based on $p18^{INK4c}$ interactions with CDK6 as reported in the ternary complex with viral cyclin 34

Table 2 Thermodynamic parameters of $p18^{INK4c}$ and mutants

Protein	T_m [1] (°C)	p-value [2]	ΔG_{UF} [3] (kcal mol^{-1})	$\Delta\Delta G_{UF}$ [4] (kcal mol^{-1})
F71N	45.96 ± 0.03	0.001	9.32 ± 0.06	3.45
F92N	45.81 ± 0.22	0.001	8.94 ± 0.10	3.08
F82Q	45.31 ± 0.14	0.001	8.52 ± 0.07	2.66
F37H	42.28 ± 0.11	0.1	7.20 ± 0.06	1.34
R55V	42.15 ± 0.01	0.1	7.09 ± 0.07	1.23
Native $p18^{INK4c}$	**41.53 ± 0.56**	**-**	**5.86 ± 0.11**	**-**
W5R	37.86 ± 0.05	0.001	3.52 ± 0.03	-2.35
T85F	33.93 ± 0.08	0.001	2.48 ± 0.07	-3.38
H108L	24.55 ± 0.23	0.001	-0.46 ± 0.07	-6.32
H75F	-	-	-	-

[1] Midpoint of thermal unfolding transition
[2] The probability that the difference of T_m between native and mutant is significant (99.9% corresponds to p-value of 0.001 and 90.0% corresponds to p-value of 0.1).
[3] Free energy of unfolding at 27 °C (300 K) with estimated error
[4] Change in free energy of unfolding with respect to Native $p18^{INK4c}$

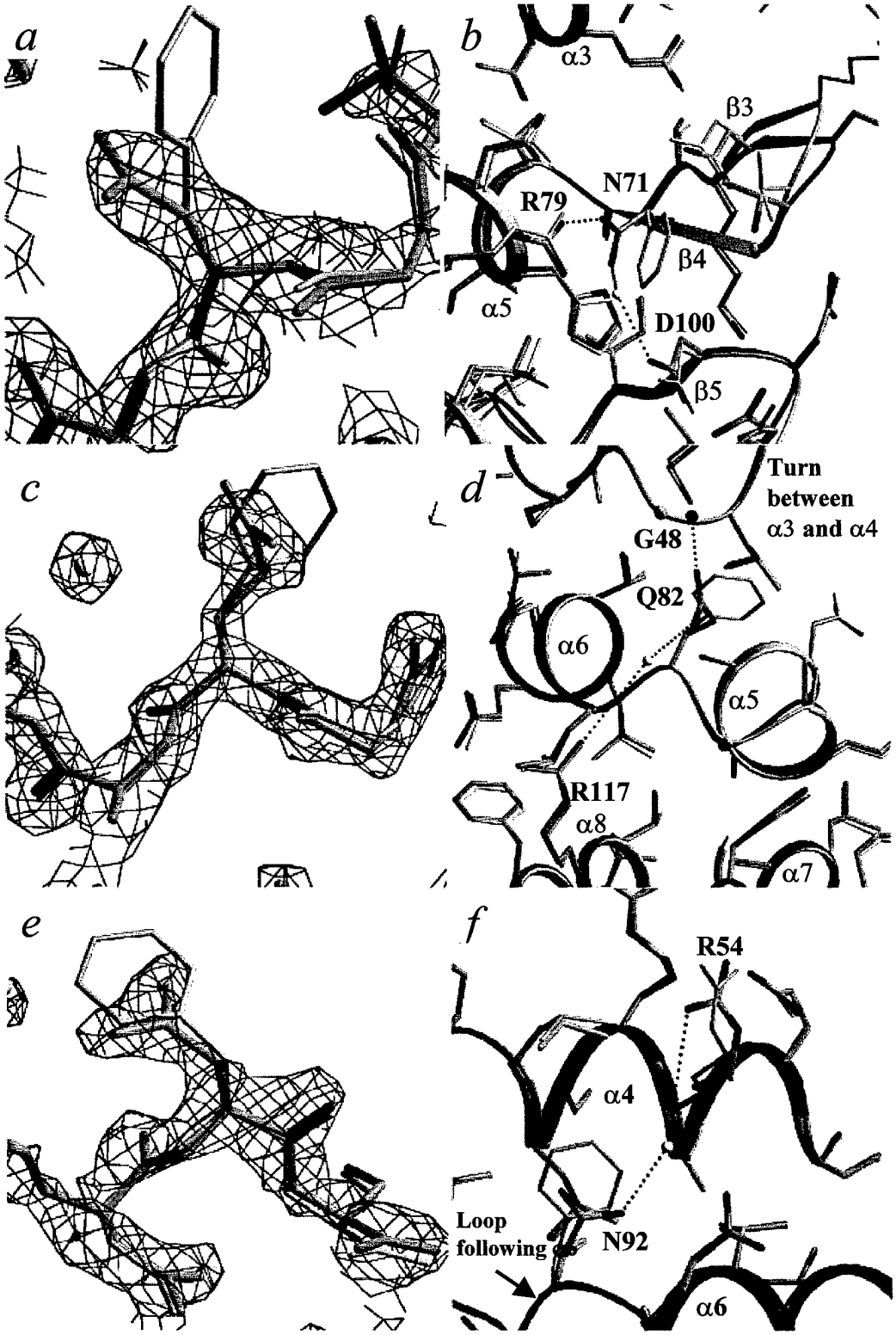


Figure 5 Crystal Structures of F71N (a, b), F82Q (c,d) and F92N (e,f) mutants of p18^{INK4c}
Figures on the left show simulated annealing omit maps of the mutated region. Figures on the right show the mutant in green with CPK coloring in comparision with the native protein in gray. Waters are shown as yellow balls and backbone NH as blue balls.

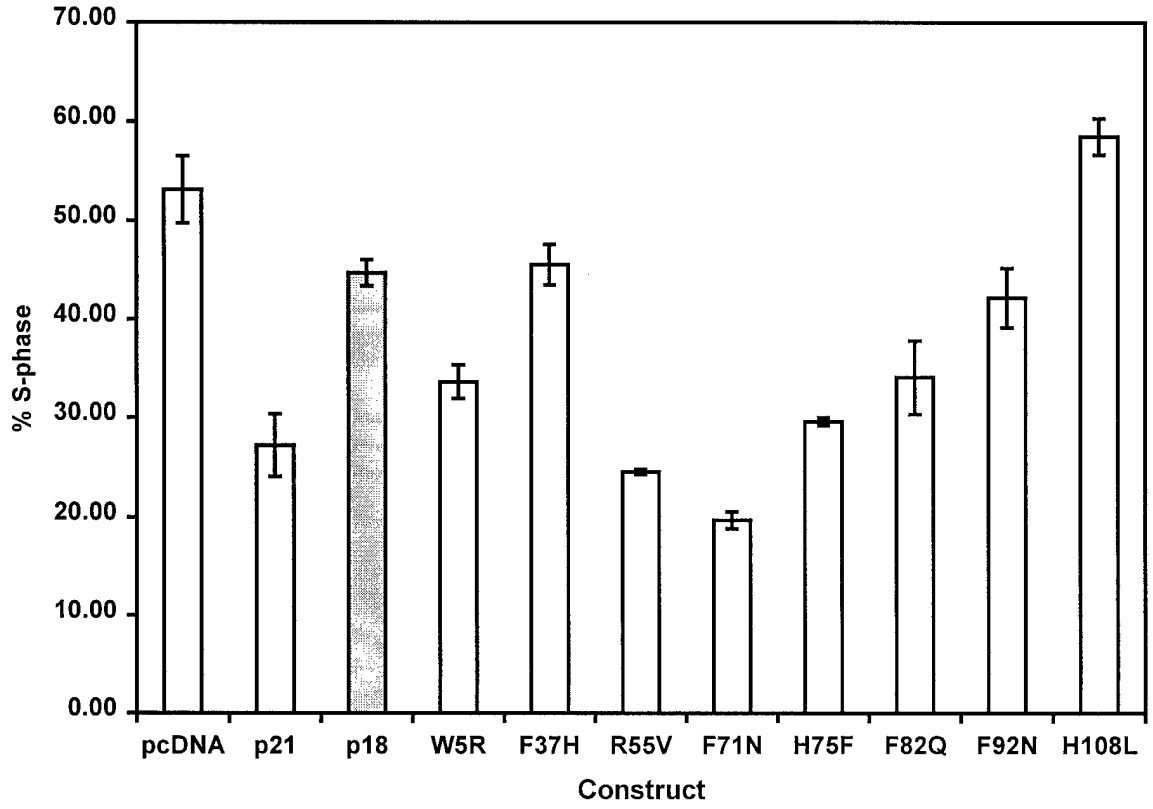


Figure 5 ***In vivo*** **cell cycle assays for p18^{INK4c} and mutants**
The comparison of p18^{INK4c} with the mutants showing the percent of U2OS cells in S-phase after 24 hours of transient transfections (along with 95% Confidence Intervals). The vector pcDNA and CDK inhibitor, p21 are shown as negative and positive controls respectively. The native p18^{INK4c} is shaded in gray while the mutants are arranged in the order of their position in the p18^{INK4c} protein sequence. The transfection experiments were performed four times, and in each experiment native or mutant expression vectors was transfected in triplicate flasks. For each cell cycle analysis from each individual transfection, 10,000-20,000 GFP-positive cells were analyzed. The data presented represents the aggregate analysis using all experiments performed, including the confidence intervals (calculated from standard deviation estimates) of the mean.

Key research accomplishments

- Purified $p16^{INK4a}$, $p18^{INK4c}$ and $p19^{INK4d}$ to homogeneity
- Tested the purified INK4 proteins biophysically to determine that all INK4 proteins have similar secondary structural composition
- Determined the crystal structure of $p18^{INK4c}$ to 1.95 Å
- Used the crystal structure to explain $p16^{INK4a}$ mis-sense mutants
- Engineered more thermostable $p18^{INK4c}$ mutants
- Determined the structures of three more thermostable $p18^{INK4c}$ mutants, F71N, F82Q, F92N
- Tested mutants in cell cycle assays
- Discovered several mutants of $p18^{INK4c}$ that are more efficacious cell cycle inhibitors

Reportable outcomes

Scientific Papers

Venkataramani, R., Swaminathan, K., and Marmorstein, R. (1998). *Crystal structure of the CDK4/6 inhibitory protein p18^{INK4c} provides insights into ankyrin-like repeat structure/function and tumor derived p16^{INK4} mutations.* Nature Structural Biology *5*, 74-81. **(Copy submitted with first annual report)**

Johnston, K., Clements A., **Venkataramani R.N.**, Trievel R.C., Marmorstein R., *Coexpression of proteins using T7-based expression plasmids: expression of heteromeric cell-cycle and transcriptional regulatory complexes.* Protein Expression and Purification 20 (3):435-43, 2000 **(Copy submitted with first second report)**

Venkataramani, R. N., MacLachlan, T., El-Deiry, W., and Marmorstein, R. (2001). *Structure-based design of p18^{INK4c} proteins with increased thermodynamic stability and cell cycle inhibitory activity.* Submitted to Nature Structural Biology. **(Copy attached)**

Scientific Presentations

Marmorstein R., **Venkataramani, R.,** Swaminathan, K., Structure of the p18^{INK4c} protein: insights into ankyrin-like repeat structure/function and CDK4/6 inhibition by the p16^{INK4a} tumor suppressor. 17th International Cancer Congress, Rio de Janeiro, Brazil, 24-28 August 1998.

Doctoral Disstertation

Venkataramani, R. N. (2001). *The crystal structure and mutational analysis of cell cycle inhibitor, p18^{INK4c}.* In Biochemistry and Molecular Biophysics (Philadelphia: University of Pennsylvania), pp. 149.

Degrees Awarded

To: Ravichandran N. Venkataramani
Degree: Ph.D. in Biochemistry and Molecular Biophysics, May 2001
Institution: University of Pennsylvania

Conclusions

The INK4 family of proteins comprise of four known members: $p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4C}$, and $p19^{INK4d}$. The INK4 proteins function to regulate the G1 to S cell cycle transition by binding to and inhibiting the pRb kinase activity of cyclin dependant kinases 4 and 6 (CDK4/6) and the $p16^{INK4a}$ member of the INK4 protein family is altered in a variety of tumor types including cancers of the breast. We purified to homogeneity $p16^{INK4a}$, $p18^{INK4c}$, and $p19^{INK4d}$. Our initial structural characterization revealed that all three proteins had very similar secondary structure. Thus, the structure of any one of the family members can answer important questions about the function of the entire family. Of the three, we were able to obtain well diffracting crystals of only one protein, $p18^{INK4c}$.

We have determined the crystal structure of $p18^{INK4c}$ to 1.95Å resolution. The structure reveals that INK4 proteins contain four or five ankyrin repeat motifs that contain a β-strand helix-turn-helix β-strand segment that associates with neighboring motifs through β-sheet and helical bundle interactions. The ankyrin repeats stack on top of each other to form an elongated overall structure in which the helical region is along one side of the protein and the β-sheet region is along the opposite side. A mapping of $p16^{INK4a}$ tumor-derived mutations onto the INK4 protein structures reveal that most of these mutations localize to residues involved in ankyrin repeat conformation or inter-repeat interactions, suggesting that a large percentage of $p16^{INK4a}$ mutations decrease the stability of the protein. Moreover, the mapping of the tumor derived mutations along with the mapping of INK4 proteins has led us to predict that the region surrounding the α3, α5, and α6 helices to be important in CDK binding. This prediction was shown to be true by the work of other investigators.

The structure function studies of the INK4 proteins by us and other investigators have revealed that the vast majority of tumor-derived $p16^{INK4a}$ mutations reduce the thermodynamic stability of $p16^{INK4a}$ suggesting a correlation between reduced thermodynamic stability and decreased cell cycle inhibitory activity of INK4 proteins. Based on this correlation, we used $p18^{INK4c}$ as a model system to test the proposal that INK4 proteins with increased thermodynamic stability might have enhanced cell cycle inhibitory activity. Structure-based mutagenesis was used to prepare $p18^{INK4c}$ mutant proteins with predicted increase in thermodynamic stability. Using this approach, we report the generation of three mutant $p18^{INK4C}$ proteins, F71N, F82Q and F92N, that have increased thermodynamic stability as measured by thermal denaturation and enhanced cell cycle inhibitory activity *in vivo,* as measured using transient transfection assays. The X-ray crystal structure of the F71N, F82Q and F92N $p18^{INK4C}$ mutant proteins were also determined to reveal the structural basis for the increased thermostability of these proteins. These studies show that a structure-based approach to increase the thermodynamic stability of INK4 proteins can be exploited to prepare more biologically active molecules with potential applications for the development of molecules to treat $p16^{INK4a}$-mediated cancers.

References

1. Guan, K.-L. *et al.* Growth suppression by p18, a $p16^{INK4/MTS1}$- and $p14^{INK4B/MTS2}$-related CDK6 inhibitor, correlates with wild-type pRb function. *Genes Dev.* **8**, 2939-2952 (1994).
2. Chan, F.K.M., Zhang, J., Cheng, L., Shapiro, D.N. & Winoto, A. Identification of human and mouse p19, a novel CDK4 and CDK6 inhibitor with homology to $p16^{ink4}$. *Mol. Cell. Biol.* **15**, 2682-2688 (1995).
3. Hannon, G.J. & Beach, D. $p15^{INK4B}$ is a potential effector of TGF-β-induced cell cycle arrest. *Nature* **371**, 257-261 (1994).
4. Kamb, A. *et al.* A cell cycle regulator potentially involved in the genesis of many tumor types. *Science* **264**, 436-440 (1994).
5. Serrano, M., Hannon, G.J. & Beach, D. A new regulatory motif in cell cycle control causing specific inhibition of cyclin D/CDK4. *Nature* **366**, 704-707 (1993).
6. Hirai, H., Roussel, M.F., Kato, H.-Y., Ashmun, R.A. & Sherr, C.J. Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. *Molecular and Cellular Biology* **15**, 2672-2681 (1995).
7. Della Ragione, F., Borriello, A., Giordani, L. & Iolascon, A. Cell division cycle alterations in human malignancies. *Cancer J.* **10**, 151-156 (1997).
8. Okuda, T. *et al.* Frequent deletion of p16INK4a/MTS1 and p15INK4b/MTS2 in pediatric acute lymphoblastic leukemia. *Blood* **85**, 2321-2330 (1995).
9. Pollock, P.M., Pearson, J.V. & Hayward, N.K. Compilation of somatic mutations of the CDKN2 gene in human cancers: non-random distribution of base substitutions. *Genes Chromosom. Cancer* **15**, 77-88 (1996).
10. Smith-Sorensen, B. & Hovig, E. CDK2A ($p16^{INK4A}$) somatic and germline mutations. *Human Mutation* **7**, 294-303 (1996).
11. Gemma, A. *et al.* Molecular analysis of the cyclin-dependent kinase inhibitor genes $p15^{INK4b/MTS2}$, $p16^{INK4/MTS1}$, p18 and p19 in human cancer cell lines. *Int. J. Cancer* **68**, 605-611 (1996).
12. Lapointe, J., Lachance, Y., Labrie, Y. & Labrie, C. A p18 mutant defective in CDK6 binding in human Breast Cancer cells. *Cancer Research* **56**, 4586-4589 (1996).
13. Pohl, U., Cairncross, J.G. & Louis, D.N. Homozygous deletions of the CDKN2C/p18INK4c gene on the short arm of chromosome 1 in anaplastic oligodendrogliomas. *Brain Pathology* **9**, 639-643 (1999).
14. Miller, C.W. *et al.* The $p19^{INK4D}$ cyclin dependent kinase inhibitor gene is altered in osteosarcoma. *Oncogene* **15**, 231-235 (1997).
15. Roussel, M.F. The INK4 family of cell cycle inhibitors in cancer. *Oncogene* **18**, 5311-5317 (1999).
16. Venkataramani, R., Swaminathan, K. & Marmorstein, R. Crystal structure of the CDK4/6 inhibitory protein $p18^{INK4c}$ provides insights into ankyrin-like repeat structure/function and tumor derived $p16^{INK4}$ mutations. *Nature Structural Biology* **5**, 74-81 (1998).
17. Botherton, D.H. *et al.* Crystal structure of the complex of the cyclin D-dependent kinase Cdk6 bound to the cell-cycle inhibitor p19INK4d. *Nature* **395**, 244-250 (1998).

18. Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D. & Pavletich, N.P. Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature* **395**, 237-243 (1998).
19. Schultz, J., Milpetz, F., Bork, P. & Ponting, C.P. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**, 5857-64. (1998).
20. Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P. & Bork, P. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* **28**, 231-4. (2000).
21. Bork, P. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? *Proteins* **17**, 363-374 (1993).
22. Arap, W., Knudsen, E.S., Wang, J.Y.J., Cavenee, W.K. & Huang, H.-J.S. Point mutations can inactivate *in vitro* and *in vivo* activities of $p16^{INK4a}$/CDKN2A in human glioma. *Oncogene* **14**, 603-609 (1997).
23. Harland, M. *et al.* Germline mutations of the CDKN2 gene in UK melanoma families. *Human Molecular Genetics* **6**, 2061-2067 (1997).
24. Koh, J., Enders, G.H., Dynlacht, B.D. & Harlow, E. Tumour-derived p16 alleles encoding proteins defective in cell-cycle inhibition. *Nature* **375**, 506-510. (1995).
25. Lilischkis, R., Sarcevic, B., Kennedy, C., Warlters, A. & Sutherland, R.L. Cancer-associated mis-sense and deletion mutations impair $p16^{INK4}$ CDK inhibitory activity. *International Journal of Cancer* **66**, 249-254 (1996).
26. Parry, D. & Peters, G. Temperature-sensitive mutants of $p16^{CDKN2}$ assocated with familial melanoma. *Molecular and Cellular Biology* **16**, 3844-3852 (1996).
27. Ranade, K. *et al.* Mutations associated with familial melanoma impair $p16^{INK4}$ function. *Nature Genet.* **10**, 114-116 (1995).
28. Reymond, A. & Brent, R. p16 proteins from melanoma-prone families are deficient in binding to CDK4. *Oncogene* **11**, 1173-1178 (1996).
29. Ruas, M., Brookes, S., McDonald, N.Q. & Peters, G. Functional evaluation of tumour-specific variants of p16INK4a/CDKN2A: correlation with protein structure information. *Oncogene* **18**, 5423-34. (1999).
30. Tang, K.S., Guralnick, B.J., Wang, W.K., Fersht, A.R. & Itzhaki, L.S. Stability and folding of the tumour suppressor protein p16. *Journal of Molecular Biology* **285**, 1869-1886 (1999).
31. Tevelev, A. *et al.* Tumor suppressor $p16^{INK4}$: structural charcterization of wild-type and mutant proteins by NMR and circular dichroism. *Biochemistry* **35**, 9475-9487 (1996).
32. Yang, R., Gombart, A.F., Serrano, M. & Koeffler, P.H. Mutational effects on the $p16^{INK4a}$ tumor suppressor protein. *Cancer Research* **55**, 2503-2506 (1995).
33. Zhang, B. & Peng, Z.-y. Defective folding of mutant $p16^{INK4}$ proteins encoded by tumor-derived alleles. *Journal of Biology Chemistry* **271**, 28734-28737 (1996).
34. Jeffrey, P.D., Tong, L. & Pavletich, N.P. Strcutural basis of inhibition of CDK-cyclin complexes by INK4 inhibitors. *Genes and Development* **14**, 3115-3125 (2000).
35. Venkataramani, R.N. University of Pennsylvania (2001).
36. Venkataramani, R.N., MacLachlan, T., El-Deiry, W. & Marmorstein, R. Structure-based design of $p18^{INK4c}$ proteins with increased thermodynamic stability and cell cycle inhibitory activity. *Submitted to Nature Structural Biology* (2001).

SUBMITTED FOR PUBLICATION TO NATURE STRUCTURAL BIOLOGY

Structure-based design of $p18^{INK4c}$ proteins with increased thermodynamic stability and cell cycle inhibitory activity

Ravichandran N. Venkataramani[1,3], Timothy K. MacLachlan[2], Wafik S. El-Deiry[2], and Ronen Marmorstein[1,3,4]

[1]The Wistar Institute, [2]Howard Hughes Medical Institute, [3]Department of Biochemistry and Molecular Biophysics and [4]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States of America

Correspondence should be addressed to R.M. (e-mail: marmor@wista.wistar.upenn.edu).

Abstract

$p18^{INK4C}$ is a member of the INK4 family of proteins that also includes $p15^{INK4b}$, $p16^{INK4a}$, and $p19^{INK4d}$. The INK4 proteins function to regulate the G1 to S cell cycle transition by binding to and inhibiting the pRb kinase activity of cyclin dependant kinases 4 and 6 (CDK4/6). The $p16^{INK4a}$ member of the INK4 protein family is altered in a variety of tumor types including cancers of the skin, esophagus and pancreas. Structure function studies of the INK4 proteins have revealed that the vast majority of tumor-derived $p16^{INK4a}$ mutations reduce the thermodynamic stability of $p16^{INK4a}$ suggesting a correlation between reduced thermodynamic stability and decreased cell cycle inhibitory activity of INK4 proteins. Based on this correlation, we used p18INK4c as a model system to test the proposal that the INK4 proteins with increased thermodynamic stability might have enhanced cell cycle inhibitory activity. Structure-based mutagenesis of $p18^{INK4c}$ was used to prepare mutant proteins with predicted increase in thermodynamic stability. Using this approach, we report the generation of three mutant $p18^{INK4C}$ proteins, F71N, F82Q and F92N, that have increased thermodynamic stability as measured by thermal denaturation and enhanced cell cycle inhibitory activity *in vivo,* as measured using transient transfection assays. The X-ray crystal structure of the F71N, F82Q and F92N $p18^{INK4C}$ mutant proteins were also determined to reveal that structural basis for the increased thermostability of these proteins. These studies show that a structure-based approach to increase the thermodynamic stability of INK4 proteins can be exploited to prepare more biologically active molecules with potential applications for the development of molecules to treat $p16^{INK4a}$-mediated cancers.

Introduction

Progression through the cell cycle is monitored at the G1-S phase checkpoint by active complexes between cyclin dependent kinases (CDK) 4 and 6 and the D-type cyclins (D1, D2, D3). The INK4 family of proteins play a key role in inhibiting the G1-S phase cell cycle transition by specifically inhibiting the kinase activity of CDK4-Cyclin D and CDK6-Cyclin D complexes. The INK4 (INhibitors of CDK4) family of proteins consists of four known members: $p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$, and $p19^{INK4d}$, which share 40% sequence identity overall and structural homology[1-5].

The INK4 family members have indistinguishable CDK/cyclin inhibitory activity when assayed *in vitro*[6], but appear to have distinct functions *in vivo*. For example, the locus on 9p21 (MTS1) expressing $p16^{INK4a}$ is a frequent target of genetic alterations in cancer, while the genes encoding the other INK4 proteins are much less commonly mutated in cancer. Specifically, gene deletions in MTS1 and promoter methylation often lead to transcriptional silencing, and point mutations are frequently associated with several different types of cancer[4,7-10]. Nearly 55 different residues are targeted in missense mutations of $p16^{INK4a}$ (ref.[10]). Homozygous deletions of $p15^{INK4b}$ have also been identified in a more limited number of cancer cell lines[11], and point mutations of $p18^{INK4c}$ have been associated with some breast carcinomas[12]. Homozygous deletions of the $p18^{INK4c}$ are also observed in oligodendrogliomas[13], however, these deletions are rare. Polymorphisms of the $p19^{INK4d}$ gene are found in a small percentage of osteosarcomas[14]. Together, the data on tumor-derived mutations suggest tissue-specific, and/or non-redundant activities for the INK4 proteins[15].

Significant insights into the mechanism of INK4-mediated inhibition of cyclin-CDK complexes and the deleterious effects of tumor-derived mutations on their kinase inhibitory properties comes from structural analyses of the INK4 proteins. The INK4 proteins contain four or five ankyrin repeat motifs that contain a β-strand helix-turn-helix β-strand segment that associates with neighboring motifs through β-sheet and helical bundle interactions. The ankyrin repeats stack on top of each other to form an elongated overall structure in which the helical region is along one side of the protein and the β-sheet region is along the opposite side[16,17]. A mapping of $p16^{INK4a}$ tumor-derived mutations onto the INK4 protein structures reveal that most of these mutations localize to residues involved in ankyrin repeat conformation or inter-repeat interactions, suggesting that a large percentage of $p16^{INK4a}$ mutations decrease the stability of the protein[17]. Indeed, several of these mutations lead to the decreased stability of $p16^{INK4a}$ and $p18^{INK4c}$ proteins when studied *in vitro*[12,18-29]. Together, these studies suggest that the decreased thermostability of $p16^{INK4a}$ correlate with its loss of function and tumorigenic properties.

The structure of INK4 proteins in binary complex with CDK6[30,31] and in ternary complex with CDK6 and a D-type viral cyclin[32] reveal that the INK4 protein binds next to the ATP binding site of the CDK. The resulting distortion of the kinase catalytic cleft prevents ATP binding and thereby inhibits catalysis[33]. While, each of the first 4 ankyrin-repeat units of the INK4 proteins participate in CDK4 interaction, most of the interactions are mediated by the 2nd and 3rd ankyrin-repeat units. In addition, comparison of the INK4 structures in the presence of CDK reveals that the INK4 protein

does not undergo significant structural rearrangement for CDK binding with an overall RMS deviation of less than 1.3 Å.

Based on the structure-function studies described above for the INK4 proteins, we hypothesized that we may be able to use a structure-based approach to prepare mutant INK4 proteins with increased thermostability and increased CDK4 inhibitory activity. The preparation of INK4 proteins with increased CDK4 inhibitory properties would potentially be useful for targeting $p16^{INK4a}$ mediated cancers and/or would provide a proof-of-principle for the design of small molecule compounds that may reactivate tumor derived $p16^{INK4a}$ mutants. Indeed, the design of such compounds for the reactivation of tumor-derived mutants of the p53 tumor suppressor has already been described[34].

We used the $p18^{INK4c}$ protein as a model system to carry out our studies for two main reasons; (1) the high resolution structure of both $p18^{INK4c}$ alone and in complex are available[17,32] therefore facilitating a detailed structure-function analysis, and (2) recombinant $p16^{INK4a}$ is a thermodynamically unstable protein with a free energy of unfolding of 3.1 kcal mol^{-1} (ref. [25]) while recombinant $p18^{INK4c}$ is more thermostable with a free energy of unfolding of 5.9 kcal mol^{-1}(Table 3). Since the INK4 proteins are highly homologous, we expect that the results of our studies with $p18^{INK4c}$ would be transferable to other INK4 proteins such as $p16^{INK4a}$. In total, we designed 9 single site mutations in $p18^{INK4c}$. Six of these mutations were predicted to increase the stability of $p18^{INK4c}$, and of these 6 positions, 4 were predicted to be at the INK4/CDK interface. Each of these 6 mutations was designed to be directly transferable to $p16^{INK4a}$. As a control, three mutations were selected that were predicted to reduce the thermostability

of $p18^{INK4c}$. The selection of mutations were based on the assumption that the following changes would increase the stability of the $p18^{INK4c}$ protein; (1) changing exposed hydrophobic residues to hydrophilic residues, (2) changing buried hydrophilic residues to hydrophobic residues, and (3) changing residues at the interface of neighboring ankyrin repeats to increase inter-ankyrin repeat interaction (by introducing hydrogen bonds or van der Waals interactions).

In this report we describe the preparation, thermodynamic stability and cell cycle inhibitory properties of the 9 mutant $p18^{INK4c}$ proteins that we prepared as well as a crystal structure determination of 3 of the most thermostable mutants. Analysis of the data reveals a general correlation between thermodynamic stability of $p18^{INK4c}$ and cell cycle inhibitory activity. A correlation of these results with the structural analysis of a subset of the $p18^{INK4c}$ mutants reveals the structural basis for the increase $p18^{INK4c}$ stability and cell-cycle inhibitory properties and suggests that improving the thermodynamic stability of INK4 proteins can be exploited to increase its cell cycle inhibitory activity. These studies have implications for the treatment of $p16^{INK4a}$ - mediated cancers as well as implications for the structure-based improvement of the biological activity of other ankyrin-repeat containing proteins.

Results and Discussion

Design of $p18^{INK4c}$ mutant proteins

The high-resolution crystal structure of $p18^{INK4c}$, was used to identify candidate residues for mutation. The strategy for the mutagenesis was to increase the hydrophobicity of the protein interior or to increase the hydrophilicity of the protein exterior. Accordingly, buried hydrophilic residues were changed to hydrophobic residues or exposed hydrophobic residues were changed to hydrophilic residues. In mutating hydrophobic residues to hydrophilic residues, mutations were selected to facilitate interactions with residues in neighboring repeats. Moreover, buried hydrophilic residues were mutated to hydrophobic residues of similar size that were predicted to facilitate favorable interactions with other hydrophobic residues in its vicinity. With the aid of modeling programs such as O [35], we modeled several mutations choosing the ones that best fit the criteria described above. The nine mutations that were made, along with their corresponding degree of conservation and structural properties, are listed in Table 1 and illustrated on a schematic of the overall $p18^{INK4c}$ structure in Fig. 1. Five of the 9 mutations selected are conserved among the INK4 proteins [Phe 37, His 75, Phe 82, Thr 85, and His 108]. Six of the mutations were designed to increase the thermodynamic stability of $p18^{INK4c}$, a subset of these residues are also predicted to be at the INK4-CDK interface based on the INK4 co-crystal structures [Phe 37 (F37H), Phe 71 (F71N), and Phe 82 (F82Q)] (Table 1). Therefore, while one set of mutations are predicted to have effects only on the thermodynamic stability of $p18^{INK4c}$, another subset are predicted to have effects on both the thermodynamic stability of $p18^{INK4c}$ and interaction with CDK. As negative controls, three mutations were also prepared that

were predicted to decrease the thermodynamic stability of $p18^{INK4c}$, these include His 75 (H75F), His 108 (H108L) and Thr 85 (T85F).

Characterization of the $p18^{INK4c}$ mutant protein during purification and crystallization

When expressed in bacteria, some mutants behave significantly differently than the native protein. These mutants include F71N, F82Q, and F92N. Bacterial growths prepared at 37 °C revealed that these mutants are found in the soluble fraction of the bacterial lysate whereas the native protein is found exclusively in the insoluble fraction when the bacteria are grown at 37 °C. Surprisingly, F71N is found exclusively in the soluble fraction, while about 25% of F82Q and F92N are found in the soluble fraction (Table 2). Each of the proteins were purified essentially as described for the native protein and eluted from gel filtration in monomeric form. Following protein purification, all three of these proteins crystallize at room temperature (at 20 °C), while the native protein precipitates when crystal trials are carried out at the same temperature. Despite the different crystal growth temperature, crystals of the three mutant proteins are isomorphous with crystals of the native protein that are obtained from the same crystallization condition, but at 4 °C (Table 2 and Table 4). The increased solubility properties of the F71N, F82Q, and F92N mutant proteins suggested that they might also be more thermostable than the native protein.

Native $p18^{INK4c}$ and each of the other mutant proteins were found exclusively in the insoluble protein fraction from a bacterial preparation that was grown at 37 °C. Upon renaturation and subsequent purification, most of these proteins eluted as monomers from gel filtration. However, the T85F and H108L mutants yielded mixtures

of aggregates and monomeric protein, and the H75F mutant protein eluted exclusively in the void volume from gel filtration. Taken together, these results suggested that the three mutants, H75F, T85F, and H108L might be thermodynamically stable than the native protein.

Thermodynamic stability of $p18^{INK4c}$ protein mutants

Circular Dichroism (CD) thermal denaturation was used to characterize and compare the thermodynamic stabilities of $p18^{INK4c}$ and its mutant proteins. Since the $p18^{INK4c}$ protein has substantial helical content, the CD signal at a wavelength of 222 nm, which is sensitive to helix formation, could be monitored as a function of solution temperature to develop a melting isotherm for each protein. The melting profiles for the native and mutant $p18^{INK4c}$ proteins are compared in Fig. 2 and the thermodynamic parameters that are derived from these profiles are tabulated in Table 3. This data reveals that of the 8 mutant $p18^{INK4c}$ proteins that could be purified to homogeneity, 5 were more thermostable than the native protein while 3 were thermodynamically less stable. Of the thermodynamically more stable $p18^{INK4c}$ mutants, there were two distinct classes. The first class of $p18^{INK4c}$ mutant proteins includes F71N, F92N, and F82Q; and the melting temperature of these mutants are approximately 4 °C higher than the native protein (p-value < 0.001). On a relative scale, these mutants are the most stable class of $p18^{INK4c}$ mutants. The second class of these thermodynamically more stable mutants include F37H and R55V and have reproducible melting temperatures of approximately 0.6-0.8 °C higher than the native protein (p-value < 0.1). The less stable class of mutants contains W5R, T85F, and H108L with melting temperatures from 4 °C to 17 °C less than the native protein (p-value < 0.001). The H75F mutant may belong in

the less stable category of $p18^{INK4c}$ mutants, but it was too unstable to be purified for further characterization (Table 2 and Table 3).

The free energy of unfolding (ΔG_{UF}) for the native and mutant $p18^{INK4c}$ proteins was calculated from the CD thermal denaturation data through linear extrapolation. The most stable class of $p18^{INK4c}$ mutants (F71N, F92N, and F82Q) have values of $\Delta\Delta G_{UF}$ that are approximately 3 kcal mol^{-1} more stable than the native protein (Fig. 2 and Table 3). The next group of stable $p18^{INK4c}$ mutants (F37H and R55V) is approximately 1 kcal mol^{-1} more stable than the native protein and the less stable class of $p18^{INK4c}$ mutants is approximately 2 to 6 kcal mol^{-1} less stable than the native protein (Fig. 2 and Table 3). All the differences in ΔG_{UF} between the native $p18^{INK4c}$ protein and the $p18^{INK4c}$ mutants are within the range of the strength of a hydrogen bond. At room temperature, the strength of a hydrogen bond is thought to be within the range of 2 to 10 kcal mol^{-1} (ref.[36]). Based on differences in ΔG_{UF}, the observed differences could be due to the formation or loss of one or more hydrogen bonding and/or van der Waals interactions.

Structural and functional characterization of the thermodynamically most stable $p18^{INK4c}$ mutant proteins, F71N, F82Q and F92N

Of the nine-$p18^{INK4c}$ mutant proteins prepared in this study, the 3 that showed the greatest increase in thermodynamic stability were F71N, F82Q and F92N (Fig. 2 and Table 3).

F71N The F71N mutant was particularly interesting since it was the only one of the mutant $p18^{INK4c}$ proteins that was completely soluble when expressed in recombinant form in bacteria (Table 2). To understand the structural basis for how a phenyalanine to asparagine change at position 71 of $p18^{INK4c}$ increased its

thermodynamic stability, we determined the crystal structure of the F71N mutant to 2.25 Å resolution (Table 4). Position 71 in $p18^{INK4c}$ is located in the β4 strand and at the beginning of ankyrin repeat 3. The structure shows that asparagine 71 directly hydrogen bonds with arginine 79 in the α5 helix in the middle of the ankyrin repeat and also makes a water mediated hydrogen bond with Asp 100 in the β5 strand at the end the ankyrin repeat (Fig. 3*a* and *b*). Taken together, the F71N $p18^{INK4c}$ mutation allows new intra-ankyrin repeat interactions that stabilize the tight turns in the structure resulting in the increased stability of the mutant protein.

Cell cycle assays *in vivo* show that the F71N $p18^{INK4c}$ mutant is also a more active cell cycle inhibitor (Fig. 5). Indeed, this mutant protein is the best cell cycle inhibitor of those tested in this study. The structure of the $p18^{INK4c}$ /CDK-cyclin complex reveals that although phenylalanine 71 does not interact with CDK, it is nonetheless at the binding interface. Therefore, it is possible that an asparagine substitution at this position would also introduce favorable $p18^{INK4c}$-CDK contacts. Modeling studies suggests that an asparagine at position 71 would be in position to interact with a backbone NH of CDK (glycine 36 of CDK6). Interestingly, other INK4 proteins have a threonine in the corresponding position, which also does not participate in CDK interaction in the $p16^{INK4a}$ and $p19^{INK4d}$ complexes with CDK [31,37]. Taken together, it is therefore likely that an asparagine substitution in other INK4 proteins would increase both their thermodynamic stabilities, CDK interaction and most importantly, their cell cycle inhibitory activities.

F82Q The F82Q mutant protein also shows increased thermodynamic stability relative to the native protein. To determine the structural basis for this effect, we

determined the crystal structure of the F82Q mutant protein to 2.0 Å resolution (Table 4). In the native protein, phenyalanine 82 is located in the turn between the α5 and α6 helices of ankyrin repeat 3. The structure of the glutamine mutant reveals that the glutamine at position 82 hydrogen bonds to the backbone nitrogen of glycine 48 on the turn between the α3 and α4 helices of ankyrin repeat 2. The glutamine mutant also makes a water-mediated hydrogen bond to arginine 117 in helix α8 in ankyrin repeat 4 (Fig. 3*c* and *d*). Therefore, the glutamine mutation in position 82 of $p18^{INK4c}$ appears to increase the thermodynamic stability of the $p18^{INK4c}$ protein by increasing inter-ankyrin repeat interaction.

Our functional studies of the F82Q $p18^{INK4c}$ mutant protein reveals that, in addition to being thermodynamically more stable, it is also a better cell cycle inhibitor than the native protein when analyzed *in vivo* (Fig. 5). The structure of the INK4/CDK complexes reveals that phenylalanine 82 makes a van der Waals contact with the CDK. Indeed, the fact that this residue is strictly conserved within the INK4 family supports its functional importance. The observation that the F82Q $p18^{INK4c}$ mutant has increased cell cycle inhibitory activity relative to the native protein suggests that the glutamine residue must make somewhat compensatory interactions with CDK. Possibly, the aliphatic region of the glutamine may mimic the van der Waals interaction that is mediated by the native phenyalanine residue. In addition, modeling studies suggests that glutamine 82 is in position to make a backbone NH hydrogen bond with the CDK (Ser 155 of CDK6).

F92N The F92N mutant protein also shows increased thermodynamic stability relative to the native protein. The structure of the F92N mutant determined to 2.0 Å

resolution (Table 4) reveals the structural basis of the asparagine substitution on thermodynamic stability. Position 92 is located on the loop following the α6 helix of the third ankyrin repeat. Glutamine 92 participates in a water-mediated hydrogen bond with arginine 54 that is located on the α4 helix on the second ankyrin repeat (Fig. 3*e* and *f*). In addition, this new interaction might stabilize the turn from α-helix 6 to the adjacent loop. Therefore, the mutation F92N appears to facilitate new inter-ankyrin repeat interactions that stabilize a tight turn in the structure resulting in the increased stability of the mutant.

Functional analysis of the F92N $p18^{INK4c}$ mutant protein reveals that it is only marginally more active than the native proteins in cell-cycle inhibition *in vivo* (Fig. 5). Interestingly, unlike the F71N and F82Q $p18^{INK4c}$ mutants that were in position to mediate interactions with CDK, the F92N mutation is unlikely to effect interaction with CDK since it is located too far from the binding interface. Presumably, the increased thermodynamic stability alone of the F92N $p18^{INK4c}$ mutant is not sufficient for increased cell cycle inhibitory activity *in vivo*.

Analysis of the thermodynamically more stable $p18^{INK4c}$ mutants - F37H and R55V

The mutation of residues phenylalanine 37 and arginine 55 results in a modest increase in thermodynamic stability of the $p18^{INK4c}$ protein. Inspection of position 37 in the native $p18^{INK4c}$ structure suggests that the mutation of phenylalanine to a histidine replaces a solvent exposed hydrophobic with a hydrophilic residue. Therefore, it is not surprising that this mutation in $p18^{INK4c}$ increases its thermodynamic stability. However, since Phe 37 is present at the end of the loop between β1 and β2, a histidine at that position will not be able to form any new interactions with adjacent ankyrin repeats or

the α-helical regions of the structure. Therefore, although a histidine may be better tolerated than a phenylalanine at the protein surface, its relatively modest contribution to the increase in thermodynamic stability of the $p18^{INK4c}$ protein is correlated to the possibility that no new protein interactions are made by the substituted histidine residue.

Despite the increased thermodynamic stability of the F37H $p18^{INK4c}$ mutant protein, its cell cycle inhibitory properties *in vivo* are essentially unchanged compared to the native protein. The structure of the INK4 complexes with CDK reveal that Phe 37 makes a Van der Waals contact with the CDK, and a histidine substitution at this position may destabilize this interaction. Therefore, what is gained in thermodynamic stability of the $p18^{INK4c}$ protein, due to this mutation, may be counteracted by an unfavorable CDK interaction. Alternatively, the relatively modest increase in thermostability of the F37H mutant may not be significant enough to translate into increased *in vivo* activity.

The R55V mutant changes a buried hydrophilic residue to a hydrophobic residue. Therefore, it is also not surprising that this $p18^{INK4c}$ mutant shows increased thermodynamic stability, although the effect is more modest than anticipated (Fig. 1 and Table 3). Interestingly, the R55V mutant is a significantly better, cell cycle inhibitor *in vivo* than the native protein. The structure of the INK4/CDK complexes reveals that residue 55 is far from the INK4/CDK interface and is therefore unlikely to contribute to protein-protein interaction. The R55V mutation therefore appears to improve the cell cycle inhibitory effect of $p18^{INK4c}$, mainly through its increased thermodynamic stability. Arginine 55 is not conserved within the INK4 proteins and this position corresponds to a leucine residue in both $p16^{INK4a}$ and $p15^{INK4b}$. This suggests that a valine substitution in

these proteins may also increase their thermodynamic stabilities and cell cycle inhibitory activities.

Analysis of $p18^{INK4c}$ mutant proteins with decreased thermodynamic stability

Of the nine $p18^{INK4c}$ protein mutants that were prepared, we had designed 3 of them to be thermodynamically less stable than the native protein, H75F, T85F, and H108L. Not surprisingly, these proteins did indeed show reduced thermodynamic stability relative to the native protein. Surprisingly, however, one other $p18^{INK4c}$ protein mutant that was designed to have increased thermostability, W5R, was also found to be thermodynamically less stable. Interestingly, the thermodynamically less stable H75F and W5R mutants were more potent cell cycle inhibitors when analyzed *in vivo*. We discuss the analysis of these four $p18^{INK4c}$ protein mutants below.

Analysis of the native $p18^{INK4c}$ structure reveals that histidines 75 and 108 participate in interactions that stabilize the β-sheet regions of the ankyrin repeats. Specifically, in both cases these histidine residues are in a position to hydrogen bond with the backbone oxygen atoms of the residues in second β-sheet of the N-terminal repeat and first β-sheet of the C-terminal repeat. This leads to the stabilization of the tight turn between the β-sheet regions (Fig. 4*a*). The role played by these histidine residues (located on the first helix of the ankyrin-repeats) in ankyrin repeat stabilization appears to be conserved in each of the INK4 structures that have been reported to date. Based on the interactions described above, we proposed that mutation of these histidine residues to a hydrophobic residue would result in the loss of the hydrogen bond interaction mediated by the histidine and thus reduce the thermodynamic stability of the $p18^{INK4c}$ protein. Consistent with the importance of these histidine residues, they

are frequently targeted in mutations of $p16^{INK4a}$ found in primary tumors[10,38]. Based on the discussion above, the H75F and H108L $p18^{INK4c}$ mutants were predicted to be thermodynamically less stable than the native protein, and this was indeed the case as revealed by our thermal denaturation studies (Table 3).

As expected, an *in vivo* cell cycle inhibition assay of the H108L $p18^{INK4c}$ mutant showed that it was a relatively poor cell cycle inhibitor (indeed, the poorest of each of the mutants tested), consistent with its decreased thermodynamic stability relative the native $p18^{INK4c}$ protein. Surprisingly, however, the H75F mutant was a better cell cycle inhibitor than native $p18^{INK4c}$, despite its reduced thermodynamic stability. The structure of the INK4/CDK complexes suggests a possible explanation. His 75 is in the center of the third ankyrin repeat of the INK4 proteins and this repeat is the one that interacts most extensively with the CDK protein. Although, this histidine does not directly participate in CDK interaction in the co-crystal structure it is in position to do so if minor side chain rearrangements of either the INK4 of CDK proteins were to take place. Indeed, asparagine at position 76 (using $p18^{INK4c}$ numbering) interacts with CDK in the co-crystal structures. Based on these arguments, it is possible that a histidine to phenylalanine substitution in $p18^{INK4c}$ would introduce a new interaction between $p18^{INK4c}$ and the CDK. Therefore, although the nascent H75F structure may have reduced thermodynamic stability relative to the native protein, it may be more stable in the context of the CDK complex, thus allowing it to bind more tightly to CDK and to be a more potent cell cycle inhibitor. Burying a phenylalanine at the INK4/CDK interface would also be more energetically favorable than burying a histidine at the interface.

Thr 85 of $p18^{INK4c}$ also plays an important role in stabilizing the INK4 structure and thus the T85F mutation in $p18^{INK4c}$ was also predicted to reduce the thermodynamic stability of the protein. Specifically, threonine 85 of $p18^{INK4c}$ is in a position to hydrogen bond with proline 50 (Fig. 4*b*), is strictly conserved among the INK4 family members (residue 93 in $p16^{INK4a}$ and residue 89 in $p19^{INK4d}$), and is targeted in mutations of $p16^{INK4a}$ in primary tumors[10,38]. The residue in $p16^{INK4a}$ alanine 57 (corresponding to proline 50 of $p18^{INK4c}$) is also a site of $p16^{INK4a}$ mutation[10] suggesting that this interaction is important in the stability of the protein. Moreover, comparisons of the structures of INK4 family members reveals that threonine 85 in $p18^{INK4c}$ (ref.[17]), threonine 93 in $p16^{INK4a}$ (ref.[39]), and threonine 89 in $p19^{INK4d}$ (ref.[30]) are all in a position to hydrogen bond with proline 50, alanine 57, and threonine 54, respectively (Fig. 4*b*). Taken together, the mutation of the strictly conserved threonine to phenylalanine leads to loss of stability, probably due to a loss of an apparently very important hydrogen bonding interaction .

The W5R $p18^{INK4c}$ mutation was predicted to increase the thermodynamic stability of $p18^{INK4c}$, since the structure of the native protein revealed that tryptophan 5 was solvent exposed and we expected that an arginine would be a more favorable residue in this position. However, the thermal denaturation studies revealed that this mutation produces a thermodynamically slightly less stable protein (Fig. 2 and Table 3). The reduced thermodynamic stability of the W5R $p18^{INK4c}$ mutation is unclear at this point, although it could be that that arginine makes promiscuous interactions within the $p18^{INK4c}$ protein that may destabilize the first ankyrin repeat of the structure. Interestingly, the W5R $p18^{INK4c}$ mutation is a significantly better cell cycle inhibitor *in*

vivo. The structure of the INK4 proteins with CDK shows that position 5 is far from the INK4/CDK binding interface and is therefore unlikely to mediate direct interaction with the CDK. The tryptophan residue is also not conserved within the INK4 proteins, also suggesting that it does not play an important role. Nonetheless, it is possible that the arginine at position 5 may indirectly influence the interaction of neighboring ankyrin repeats of the INK4 proteins with CDK4. Alternatively position 5 of $p18^{INK4c}$ may be involved in other, yet unidentified $p18^{INK4c}$ interactions *in vivo*. Functional characterization of other mutations in this position may provide insights into the role played by residue 5 in the thermodynamic stability and the cell cycle inhibitory activity of the $p18^{INK4c}$ protein.

Conclusions

We have used a structure-based approach to successfully prepare $p18^{INK4c}$ proteins containing single site mutations that are thermodynamically more stable *in vitro,* and that are more potent cell cycle inhibitors *in vivo* than the native protein. In general, we find a strong correlation between increased thermostability *in vitro* and the increased potency of cell cycle inhibition *in vivo*. However, we have also identified some interesting exceptions. In this study, several factors that can play a role in the *in vivo* cell cycle activity of the $p18^{INK4c}$ mutant proteins are not controlled for, including their relative abilities to interact with CDK and their half-life *in vivo*. Therefore, it is not surprising that there is not a strict correlation between the degree of thermodynamic stability *in vitro* and the ability to inhibit cell cycle progression *in vivo* by the $p18^{INK4c}$ mutant proteins. Indeed, the importance of some of these other factors is supported by

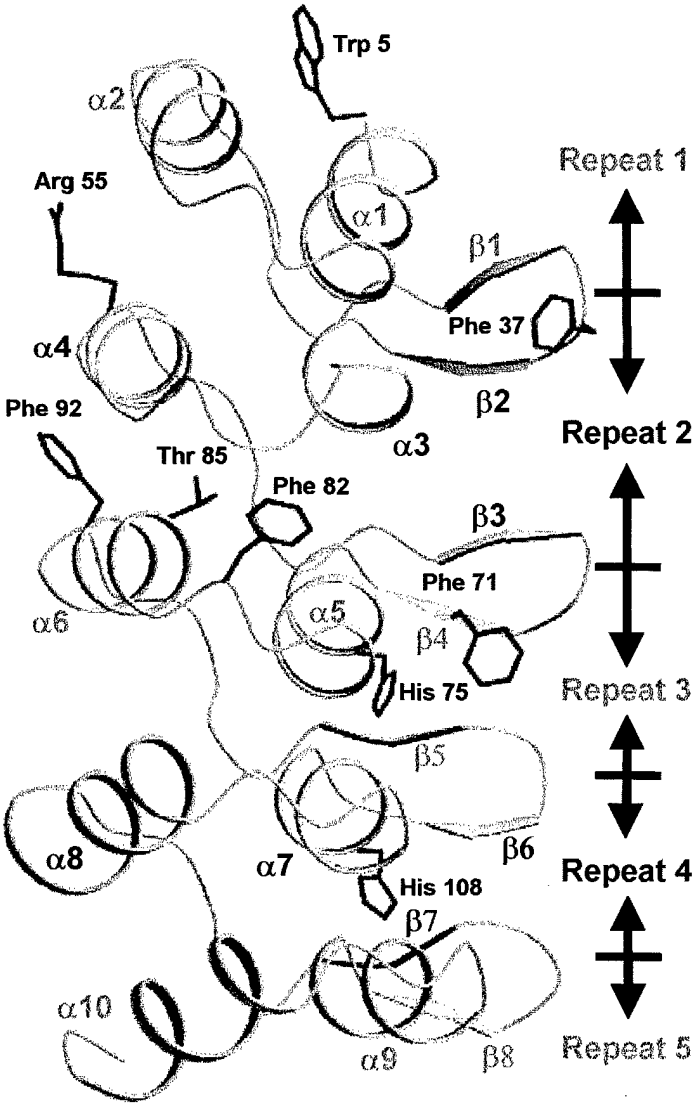
our finding that between thermodynamic stability *in vitro* and cell cycle inhibitory activity *in vivo* of two of the nine $p18^{INK4c}$ mutant proteins that we prepared (W5R and H75F) are not correlated. Nonetheless, this study does show that the thermodynamic stability of $p18^{INK4c}$ can be exploited to create more active cell cycle inhibitory proteins, and that a structure-based approach can be u sed to facilitate this process.

Of the single-site $p18^{INK4c}$ mutant proteins that we prepared, of particular interest is the F71N mutation. This protein was the most thermodynamically stable *in vitro* and the most potent cell cycle inhibitor *in vivo*. Since an asparagine substitution was chosen arbitrarily (within the constraints of a charged residue) there is a possibility that mutation to another residue may create an even more thermodynamically stable and more potent $p18^{INK4c}$ cell cycle inhibitor. Indeed, we are now in a position to combine $p18^{INK4c}$ mutants that show increased thermodynamic stability and cell cycle inhibitory activity to create even more stable and active molecules. Moreover, we would expect that some of the results obtained in this study can be extrapolated to the homologous $p16^{INK4a}$ protein, and that introduction of similar mutations into $p16^{INK4a}$ would result in a more potent cell cycle inhibitor and tumor suppressor. Such "enhanced" $p16^{INK4a}$ proteins may be ideal candidates for a gene therapy approach to help treat $p16^{INK4a}$ -mediated cancers. Moreover, the studies presented here provide a proof of principle that small molecule compounds may be prepared that can mimic the effect of some of the mutations described here and may function to reactivate tumor-derived $p16^{INK4a}$ mutations.

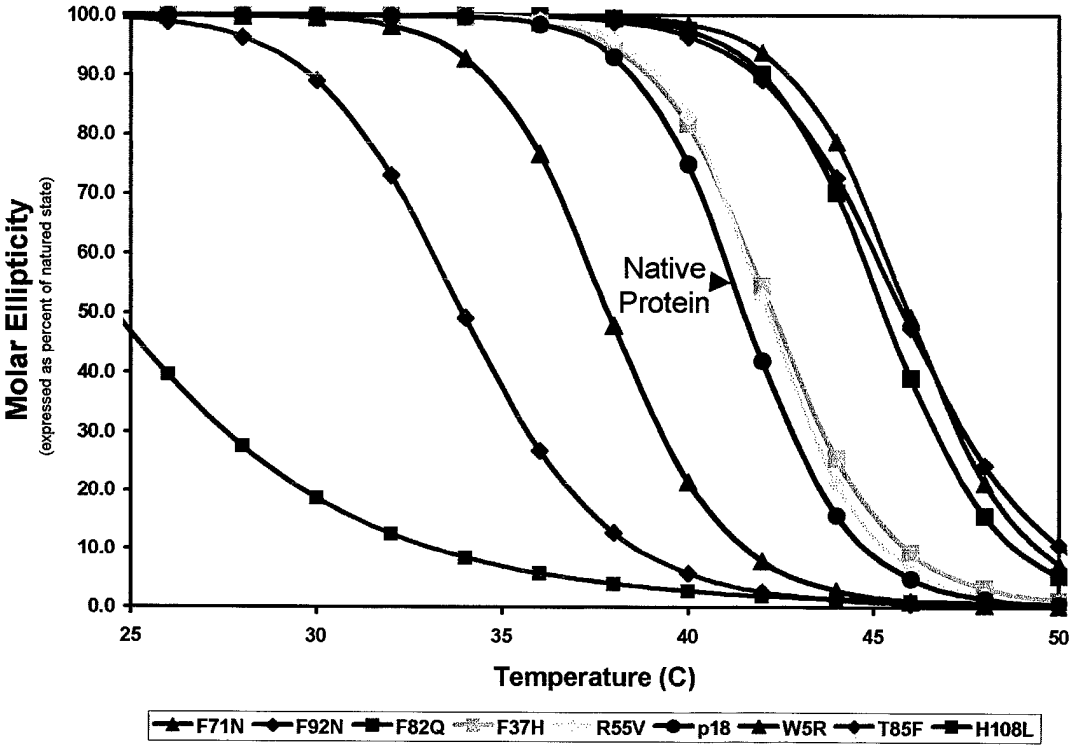
Acknowledgements

We would like to thank Yelena Shifman for technical assistance.

Venkataramani, et.al.
Figure 1

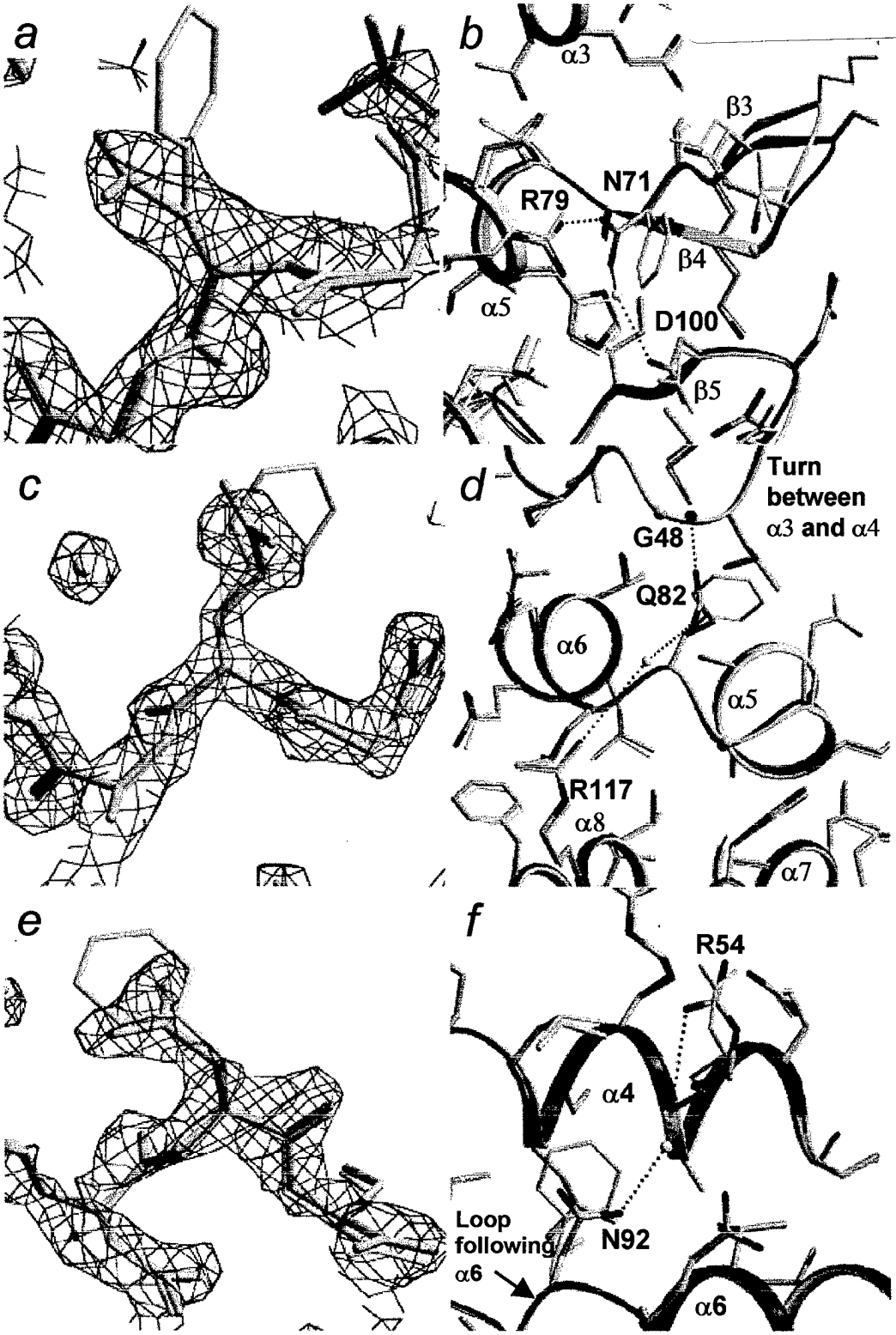


Venkataramani, et.al.
Figure 2

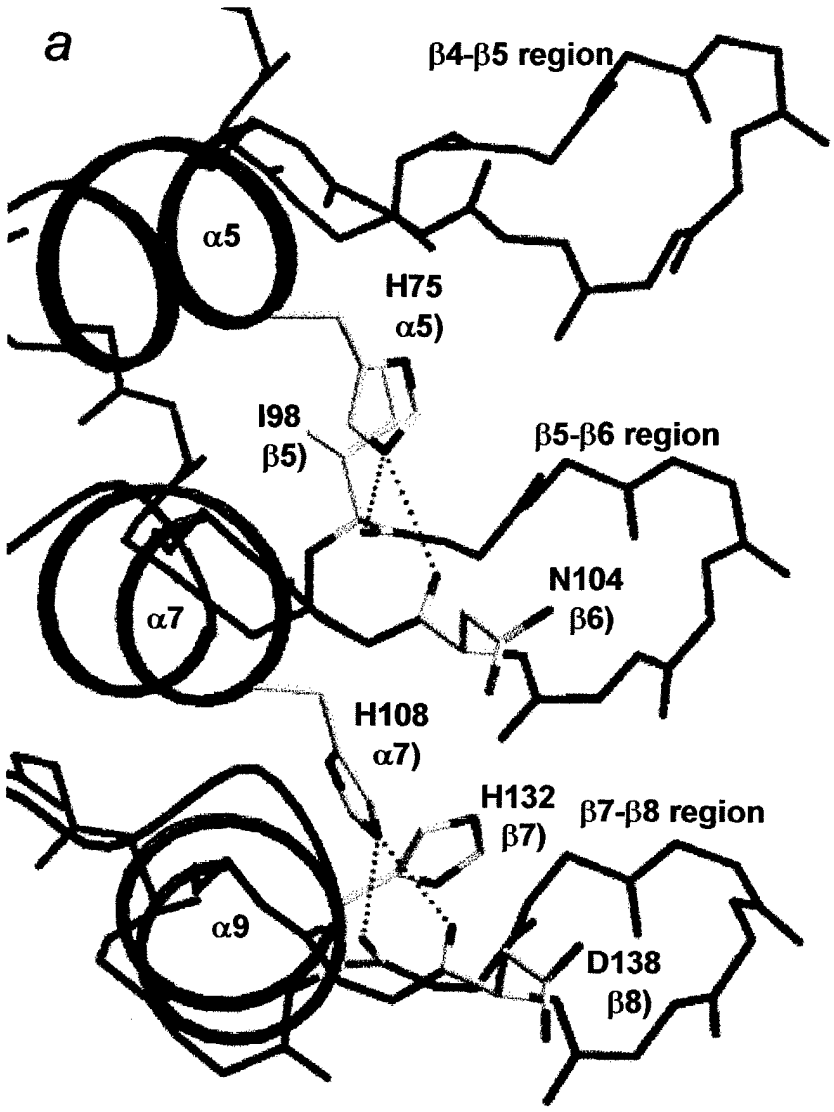


Venkataramani, et.al.

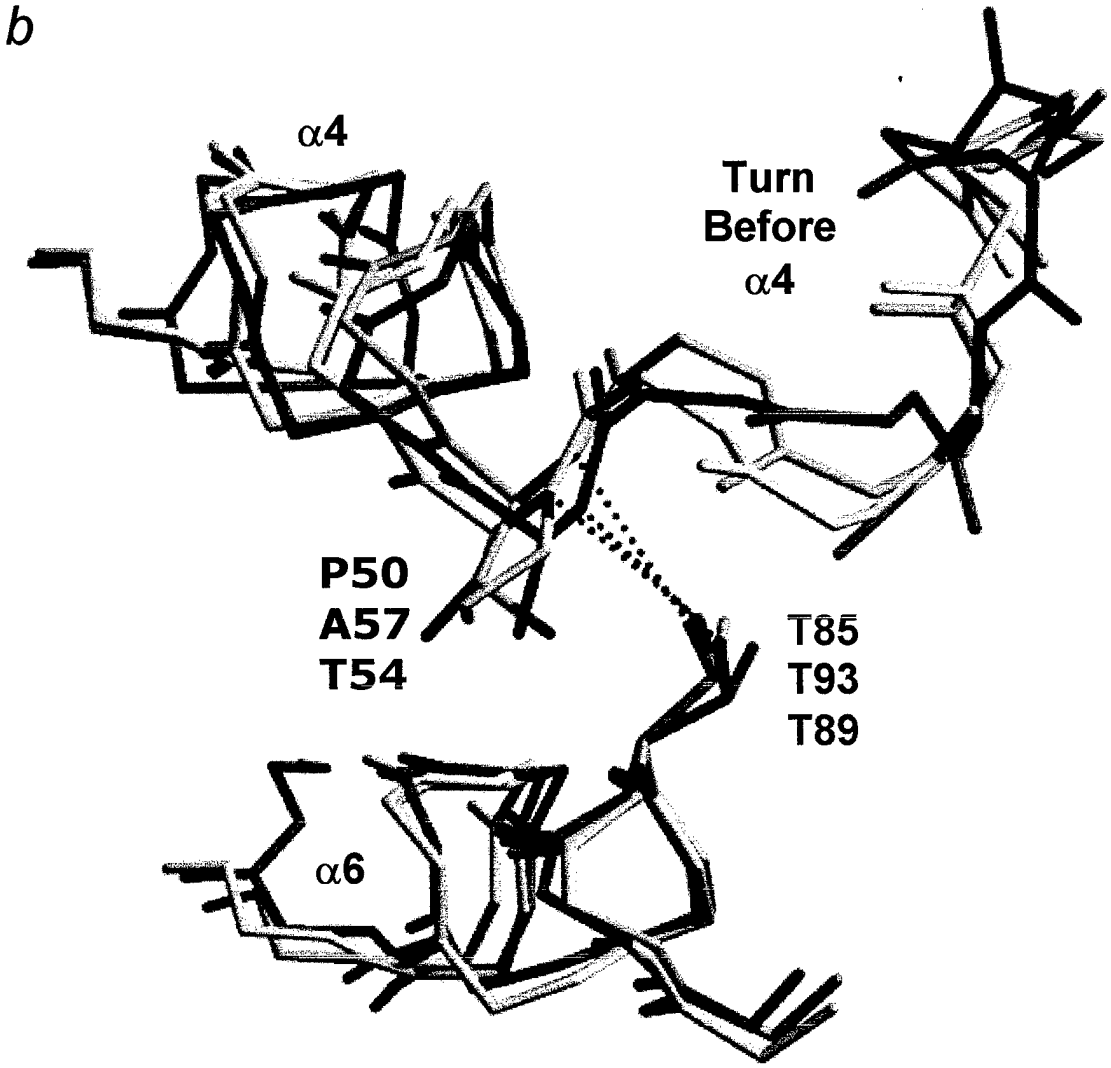
Figure 3



Venkataramani, et.al.
Figure 4a



Venkataramani, et.al.
Figure 4b



Venkataramani, et.al.
Figure 5

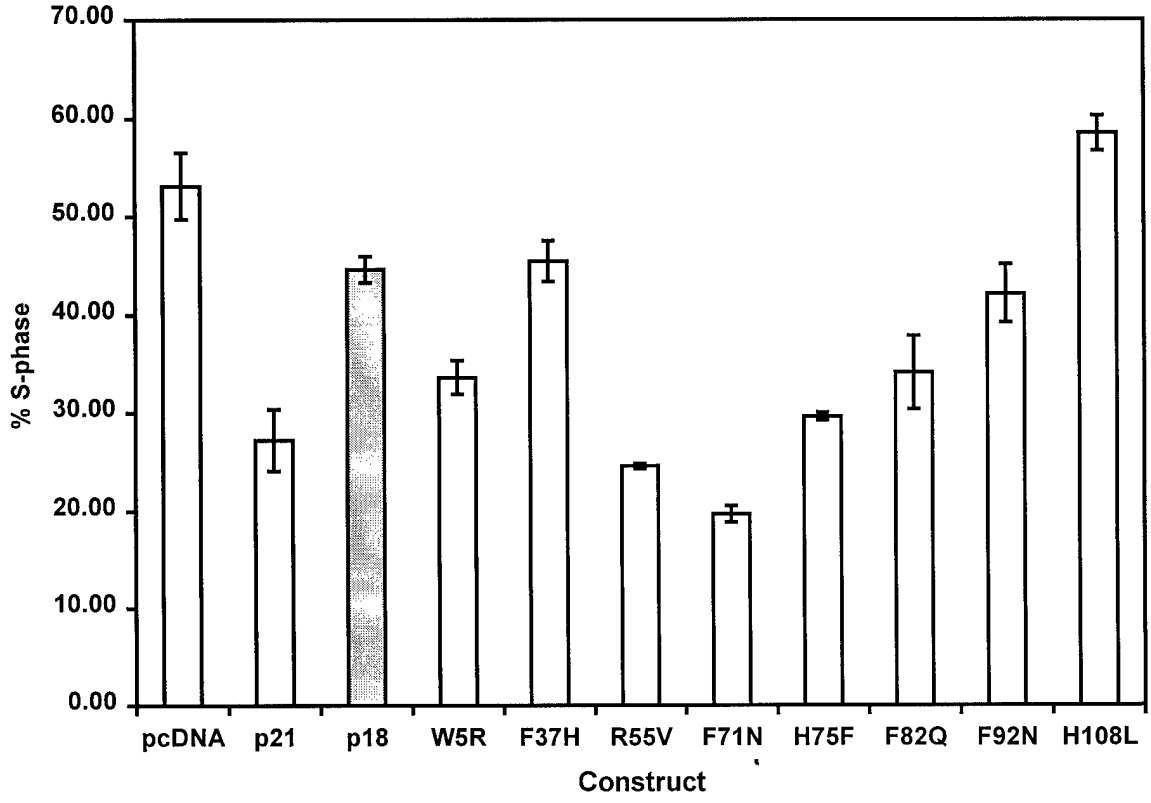


Figure Legends

Figure 1 Residues of $p18^{INK4c}$ targeted for mutations

The ankyrin repeat elements are shown in alternating olive green and red, and the residues targeted for mutations are shown in red.

Figure 2 Thermal denaturation profiles of native and mutant $p18^{INK4c}$ proteins

The overlay of thermal denaturation profiles of native and mutant $p18^{INK4c}$ proteins illustrates the three classes of mutant proteins classified according to stability. The most stable class of $p18^{INK4c}$ mutant proteins include F71N, F92N, and F82Q, and the stable class include F37H, and R55V. The least stable class includes W5R, T85F, and H108L. The H75F $p18^{INK4c}$ mutant is also included within the less stable class but was not purified or analyzed by thermal denaturation due to its poor aggregation properties. The mean at each temperature of duplicate experiments are shown here. The standard deviations of the data was very small due to the reproducibility of the data, therefore, error bars are not shown.

Figure 3 Structure of the F71N, F82Q, and F92N $p18^{INK4c}$ mutants in their comparison with the native $p18^{INK4c}$ protein

a. The superposition of the mutant, F71N (in green), and the native protein (in gray) along with the simulated annealing omit map around the site of mutation contoured at 1.5σ. The yellow sphere represents a water molecule.

b. Detailed interactions made in the F71N $p18^{INK4c}$ mutant are shown in green with CPK coloring, while the structure of the native protein, which lacks these, new

interactions are shown in gray. The mutation results in new hydrogen bonding interaction with arginine 79 and a water (shown in yellow) mediated hydrogen bond with aspartate 100. Glycines are shown as green or gray spheres.

c. Same as A except that the F82Q mutant is shown.

d. Detailed interactions made in the F82Q $p18^{INK4c}$ mutant are shown in green with CPK coloring, while the structure of the native protein, which lacks these, new interactions are shown in gray. The mutation results in new hydrogen bonding interaction with the backbone NH of glycine 48 and a water (shown in yellow) mediated hydrogen bond with arginine 117. Glycines are shown as green or gray spheres.

e. Same as A except that the F92N mutant is shown.

f. Detailed interactions made in the F92N $p18^{INK4c}$ mutant are shown in green with CPK coloring, while the structure of the native protein, which lacks these, new interactions are shown in gray. The mutation results in a new water mediated hydrogen bond with arginine 54. Glycines are shown as green or gray spheres.

Molscript objects for the electron density was created using CONSCRIPT[40] and the figures were prepared with the programs MOLSCRIPT[41] and RASTER3D[42].

Figure 4 Interactions of conserved histidine and threonine residues

a. The conserved interactions of histidine residues 75 and 108 are highlighted in the β-sheet region of the ankyrin repeats 3,4 and 5. The β-sheet regions are shown as a Cα trace while the helical regions are shown as coils. The secondary structural

element in which the residues are found is shown in parentheses under the residue label.

b. The conserved interactions of threonine 85 in $p18^{INK4c}$ (shown in aqua), is compared to interactions made by the corresponding threonine of other INK4 proteins; threonine 93 in $p16^{INK4a}$ (shown in brown) and threonine 89 in $p19^{INK4d}$ (shown in green).

Figure 5 *In vivo* cell cycle assays for $p18^{INK4c}$ and mutants

The comparison of $p18^{INK4c}$ with the mutants showing the percent of U2OS cells in S-phase after 24 hours of transient transfections (along with 95% Confidence Intervals). The vector pcDNA and CDK inhibitor, p21 are shown as negative and positive controls respectively. The native $p18^{INK4c}$ is shaded in gray while the mutants are arranged in the order of their position in the $p18^{INK4c}$ protein sequence. The transfection experiments were performed four times, and in each experiment native or mutant expression vectors was transfected in triplicate flasks. For each cell cycle analysis from each individual transfection, 10,000-20,000 GFP-positive cells were analyzed. The data presented represents the aggregate analysis using all experiments performed, including the confidence intervals (calculated from standard deviation estimates) of the mean.

Table 1 Properties of Residues Targeted for Mutations

Residue	Rationale for mutation	Mutation	Predicted change in thermostability	Interaction with residues in CDK6[1]	Conservation among INK4 proteins
Trp 5	Exposed Hydrophobic	Arg	Increase	No	No
Phe 37	Exposed Hydrophobic	His	Increase	Yes, non-polar side chain contacts	Yes
Arg 55	Buried Hydrophilic	Val	Increase	No	No
Phe 71	Exposed Hydrophobic, Possible inter-repeat contacts	Asn	Increase	Yes, non-polar side chain contacts	No
His 75	Buried polar	Phe	Decrease	No	Yes
Phe 82	Exposed Hydrophobic, Possible inter-repeat contacts	Gln	Increase	Yes, non-polar side chain contacts	Yes
Thr 85	Buried polar	Phe	Decrease	No	Yes
Phe 92	Exposed Hydrophobic, Possible inter-repeat contacts	Asn	Increase	No	No
His 108	Buried polar	Leu	Decrease	No	Yes

[1] Based on $p18^{INK4c}$ interactions with CDK6 as reported in the ternary complex with viral cyclin[32]

Table 2 Biochemical Characterization of $p18^{INK4c}$ Mutants

Construct	Solubility of protein from bacterial lysates[1]	Ability to renature protein[2]	Oligomerization State[3]	Crystallization [4]
Native	Insoluble	Yes	Monomer	Yes at 4 °C
W5R	Insoluble	Yes	Monomer	NT
F37H	Insoluble	Yes	Monomer	NT
R55V	Insoluble	Yes	Monomer	NT
F71N	100% soluble	NT	Monomer	Yes at 20 °C
H75F	Insoluble	Yes	Aggregated	NT
F82Q	~ 25% soluble	Yes	Monomer	Yes at 20 °C
T85F	Insoluble	Yes	Mixture of aggregates and monomer	NT
F92N	~ 25% soluble	Yes	Monomer	Yes at 20 °C
H108L	Insoluble	Yes	Mixture of aggregates and monomer	NT

NT: Not Tested

[1] The solubility of the protein refers to the presence of the protein in either the soluble or insoluble fractions of the bacterial lysate when the protein is over-expressed at 37°C

[2] The ability to renature the protein refers to whether the protein can be isolated in the soluble fraction after denaturation by 6M urea followed by stepwise dialysis to remove urea

[3] Molecular size is judged by size exclusion chromatography of the renatured protein

[4] Crystallization was only carried out with a few mutants; the crystals so obtained were isomorphous to the crystals of the native protein.

Table 3 Thermodynamic Parameters of $p18^{INK4c}$ and Mutants

Protein	T_m [1] (°C)	p-value [2]	ΔG_{UF} [3] (kcal mol^{-1})	$\Delta\Delta G_{UF}$ [4] (kcal mol^{-1})
F71N	45.96 ± 0.03	0.001	9.32 ± 0.06	3.45
F92N	45.81 ± 0.22	0.001	8.94 ± 0.10	3.08
F82Q	45.31 ± 0.14	0.001	8.52 ± 0.07	2.66
F37H	42.28 ± 0.11	0.1	7.20 ± 0.06	1.34
R55V	42.15 ± 0.01	0.1	7.09 ± 0.07	1.23
Native $p18^{INK4c}$	**41.53 ± 0.56**	**-**	**5.86 ± 0.11**	**-**
W5R	37.86 ± 0.05	0.001	3.52 ± 0.03	-2.35
T85F	33.93 ± 0.08	0.001	2.48 ± 0.07	-3.38
H108L	24.55 ± 0.23	0.001	-0.46 ± 0.07	-6.32
H75F	-	-	-	-

[1] Midpoint of thermal unfolding transition
[2] The probability that the difference of T_m between native and mutant is significant (99.9% corresponds to p-value of 0.001 and 90.0% corresponds to p-value of 0.1).
[3] Free energy of unfolding at 27 °C (300 K) with estimated error
[4] Change in free energy of unfolding with respect to Native $p18^{INK4c}$

Table 4 Summary of Crystallographic Data and Refinement Parameters

	F71N	F82Q	F92N
Data Statistics			
Resolution Range	50.0-2.25	50.0-2.0	50-2.0
Unique Reflections	22,638	22,818	22,869
R_{sym} (%)	8.5 %	3.9 %	6.2 %
I/sigma	15.19	23.87	10.10
Completeness	92.7 %	97.7 %	96.6 %
Refinement Statistics			
Protein atoms	2376	2404	2376
Water atoms	192	179	174
R-factors			
R_{free}	25.4	26.0	26.3
$R_{working}$	21.4	23.1	23.7
Model Statistics			
rms values			
Bond length	0.004	0.005	0.006
Bond Angles	1.19	1.24	1.34
NCS molecules	0.819	0.756	0.745
Average B-factors	19.51	25.69	26.61

Materials and Methods

Cloning and protein purification

Single amino acid substitutions in $p18^{INK4c}$ were produced using the QuickChange mutagenesis kit (Stratagene) using the pRSETA-$p18^{INK4c}$ vector[17] as a template, and sequence changes were confirmed by DNA sequencing. Recombinant native and mutant $p18^{INK4c}$ proteins were prepared by growing pRSETA-$p18^{INK4c}$ (native or mutant) transformed BL21 (DE3) cells grown and inducing at 37 °C. Cells were lysed and purified in a low salt buffer (LSB) containing 50 mM Tris (pH 8.5), 50 mM NaCl, 5 mM βME, and 0.1 mg ml^{-1} PMSF (phenylmethylsulfonyl fluoride). Except for the F71N $p18^{INK4c}$ mutant, each of the recombinant proteins were isolated by renaturing the protein found in the inclusion bodies with 6M Urea denaturation and renaturation by stepwise dialysis to remove urea. The soluble fraction from renaturation was further purified by cation exchange with a Q-sepharose fast flow column (Pharmacia) and gel filtration using Superdex-75 column (Pharmacia)[17]. The F71N $p18^{INK4c}$ mutant was the only protein that was found exclusively in the soluble protein fraction and the supernatant was directly purified using cation exchange and gel filtration chromatography[17]. Protein aliquots were flash-frozen at -70 °C and thawed as needed for biophysical studies or crystallization.

Circular Dichroism monitored thermal denaturation

Circular Dichroism data was collected on an AVIV Circular Dichroism spectrophotometer (Model 62A-DS) equipped with a thermoelectric unit and using a 1-

mm path length cell. Thermal protein denaturation was monitored at 222 nm. Data was collected every 2 °C with an equilibrium time of 4 min and an averaging time of 10 seconds. Protein samples were analyzed at a concentration of 2 mg mL^{-1} (approx. 20 mM) in 50 mM HEPES pH 7.5. Protein concentrations were determined using UV spectrophotometry. However, concentrations of the W5R mutant which had low extinction coefficients was determined using a Bradford assay with native protein as a standard was used. The optimal pH for the studies was chosen based on thermal denaturation data collected at different pH values - Borate pH 9.0, Tris pH 8.5, HEPES pH 7.5, and Cacodylate pH 6.0 (data not shown). Among the conditions tested, HEPES pH 7.5 gave the best sigmoidal transition. The melting temperature, T_m and error estimates were calculated by fitting the denaturation data (molar ellipticities) to a non-linear dose-response logistical transition model [Calculated molar ellipticity, $\theta_{calc} = \theta_d + \{(\theta_n-\theta_d) / 1 + (T/T_m)^{slope}\}$] using the Microsoft Excel solver module where slope refers to the slope of the transition from the native to the denatured state and θ_d and θ_n refers to the mollar ellipticities of the denatured state and the native state respectively.

The quality of the fit of model to data was evaluated using R^2 with all fits having estimates of greater than 99 %. F-tests were used to analyze the variances of the model and data which were found to agree with each other >99% of the time for all models. Using these models, T_m's and standard deviation estimates for native and mutant $p18^{INK4c}$ proteins were calculated from duplicate experiments. The p-level of significance of T_m values was calculated using the confidence intervals associated with the mean T_m measurements. ΔG's and error estimates were calculated by fitting the denaturation data to a two-state transition model[43]. For each mutant, ΔG's at different

temperatures within the transition zone were fitted to a straight-line (all R^2 estimates > 99 %). The resulting straight-line equation was used to extrapolate and calculate the change in free energy at room temperature, ΔG_{UF} at 300 K (27 °C).

Crystallization and structure determination of the F71N, F82Q, and F92N p18^{INK4c} mutants

Crystallization was carried out by vapor diffusion at room temperature, using a 2 μL hanging drop containing 5 mg ml^{-1} protein, 20 mM Tris (pH 8.5), 0.5 mM DTT, 7 % PEG 6000 (polyethylene glycol, average molecular mass 6000 M_r), and 1M NaCl equilibrated over a reservoir containing 14 % PEG 6000 and 2M NaCl. These conditions are similar to those used to crystallize the native protein. Crystals grew to a typical size of 100-200 μm x 75-150 μm x 50-100 μm in the space group $P2_12_12$, with cell dimensions that are isomorphous with the native p18^{INK4c} crystals (a= 55.6 Å, b=151.6 Å, c=40.5 Å and $\alpha=\beta=\gamma=90°$). Crystals were transferred to a harvest solution (HS) containing 20 mM Tris (pH 8.5), 15% PEG 6000 and 2M NaCl, transferred stepwise to HS supplemented with 25 % glycerol, and frozen in liquid nitrogen-cooled liquid propane prior to data collection.

Diffraction data was collected at 110 K using a Rigaku Raxis IV image plate detector with CuKα radiation from a Rigaku RU-300 generator, and the data was processed and scaled using DENZO and SCALEPACK[44]. The coordinates of native p18^{INK4c} was used as a starting model for the mutant proteins. The position and orientation of these coordinates was initially adjusted using rigid body refinement with the program CNS[45]. Model building was carried out using the program O[46] using omit maps and sigma A weighted 2 F_o-F_c and F_o - F_c Difference Fourier maps created with

the program CNS[45]. Iterative rounds of model building followed by positional refinement, simulated annealing[47], and torsion angle dynamic[48] refinement protocols were carried out at successively higher resolution shells. At the final resolution shell, solvent molecules were added and the final model was checked for errors with simulated annealing omit maps omitting 10 residues at a time[49]. The resulting protein structures have excellent refinement statistics and geometry (Table 4).

In vivo cell cycle inhibition assay

DNA encoding native and mutant $p18^{INK4c}$ proteins were subcloned into the pCDNA3.1 vector creating a set of pCDNA-p18 vectors (T85F was not tested due to complications in the sub-cloning). The cell culture conditions of the U2OS cells were essentially as described[50]. A total of 1×10^6 U2OS cells were transfected using Superfect reagent (QIAGEN Inc., Valencia, CA) with 0.5 µg of green fluorescent protein reporter and 1.5 µg of pCDNA-p18 (native or mutant), pCDNA-p21, or vector alone. Western blotting was carried out essentially as described[50] using mouse anti-human p18 polyclonal (kind gift of Dr. Yue Xiong, University of North Carolina at Chapel Hill) to confirm comparable expression levels of the native and mutant proteins (data not shown). The cells were prepared in 1% fetal bovine serum/phosphate-buffered saline for Fluorescence-activated Cell Sorting (FACS) analysis after 24 hours. Preparation of cells for fluorescence-activated cell sorting was performed essentially as described[51]. Cell sorting was performed on a Coulter Epics Elite counter. DNA content analysis was performed using MacCycle software (Phoenix Flow Systems, San Diego, CA).

Bibliography

1. Guan, K.-L. *et al.* Growth suppression by p18, a $p16^{INK4/MTS1}$- and $p14^{INK4B/MTS2}$-related CDK6 inhibitor, correlates with wild-type pRb function. *Genes Dev.* **8**, 2939-2952 (1994).
2. Chan, F.K.M., Zhang, J., Cheng, L., Shapiro, D.N. & Winoto, A. Identification of human and mouse p19, a novel CDK4 and CDK6 inhibitor with homology to $p16^{INK4}$. *Mol. Cell. Biol.* **15**, 2682-2688 (1995).
3. Hannon, G.J. & Beach, D. $p15^{INK4B}$ is a potential effector of TGF-β-induced cell cycle arrest. *Nature* **371**, 257-261 (1994).
4. Kamb, A. *et al.* A cell cycle regulator potentially involved in the genesis of many tumor types. *Science* **264**, 436-440 (1994).
5. Serrano, M., Hannon, G.J. & Beach, D. A new regulatory motif in cell cycle control causing specific inhibition of cyclin D/CDK4. *Nature* **366**, 704-707 (1993).
6. Hirai, H., Roussel, M.F., Kato, H.-Y., Ashmun, R.A. & Sherr, C.J. Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. *Mol. Cell. Biol.* **15**, 2672-2681 (1995).
7. Della Ragione, F., Borriello, A., Giordani, L. & Iolascon, A. Cell division cycle alterations in human malignancies. *Cancer J.* **10**, 151-156 (1997).
8. Okuda, T. *et al.* Frequent deletion of $p16^{INK4a/MTS1}$ and $p15^{INK4b/MTS2}$ in pediatric acute lymphoblastic leukemia. *Blood* **85**, 2321-2330 (1995).
9. Pollock, P.M., Pearson, J.V. & Hayward, N.K. Compilation of somatic mutations of the CDKN2 gene in human cancers: non-random distribution of base substitutions. *Genes Chromosomes Cancer* **15**, 77-88 (1996).
10. Smith-Sorensen, B. & Hovig, E. CDK2A ($p16^{INK4a}$) somatic and germline mutations. *Hum. Mutat.* **7**, 294-303 (1996).
11. Gemma, A. *et al.* Molecular analysis of the cyclin-dependent kinase inhibitor genes $p15^{INK4b/MTS2}$, $p16^{INK4/MTS1}$, p18 and p19 in human cancer cell lines. *Int. J. Cancer* **68**, 605-611 (1996).
12. Lapointe, J., Lachance, Y., Labrie, Y. & Labrie, C. A p18 mutant defective in CDK6 binding in human Breast Cancer cells. *Cancer Res.* **56**, 4586-4589 (1996).
13. Pohl, U., Cairncross, J.G. & Louis, D.N. Homozygous deletions of the CDKN2C/$p18^{INK4c}$ gene on the short arm of chromosome 1 in anaplastic oligodendrogliomas. *Brain Pathol.* **9**, 639-643 (1999).
14. Miller, C.W. *et al.* The $p19^{INK4D}$ cyclin dependent kinase inhibitor gene is altered in osteosarcoma. *Oncogene* **15**, 231-235 (1997).
15. Roussel, M.F. The INK4 family of cell cycle inhibitors in cancer. *Oncogene* **18**, 5311-5317 (1999).
16. Luh, F.Y. *et al.* Structure of the cyclin-dependent kinase inhibitor $p19^{Ink4d}$. *Nature* **389**, 999-1003 (1997).
17. Venkataramani, R., Swaminathan, K. & Marmorstein, R. Crystal structure of the CDK4/6 inhibitory protein $p18^{INK4c}$ provides insights into ankyrin-like repeat structure/function and tumor derived $p16^{INK4}$ mutations. *Nat. Struct. Biol.* **5**, 74-81 (1998).

18. Arap, W., Knudsen, E.S., Wang, J.Y.J., Cavenee, W.K. & Huang, H.-J.S. Point mutations can inactivate *in vitro* and *in vivo* activities of $p16^{INK4a}$/CDKN2A in human glioma. *Oncogene* **14**, 603-609 (1997).
19. Harland, M. *et al.* Germline mutations of the CDKN2 gene in UK melanoma families. *Hum. Mol. Genet.* **6**, 2061-2067 (1997).
20. Koh, J., Enders, G.H., Dynlacht, B.D. & Harlow, E. Tumour-derived p16 alleles encoding proteins defective in cell-cycle inhibition. *Nature* **375**, 506-510. (1995).
21. Lilischkis, R., Sarcevic, B., Kennedy, C., Warlters, A. & Sutherland, R.L. Cancer-associated mis-sense and deletion mutations impair $p16^{INK4}$ CDK inhibitory activity. *Int. J. Cancer* **66**, 249-254 (1996).
22. Parry, D. & Peters, G. Temperature-sensitive mutants of $p16^{CDKN2}$ assocated with familial melanoma. *Mol. Cell. Biol.* **16**, 3844-3852 (1996).
23. Ranade, K. *et al.* Mutations associated with familial melanoma impair $p16^{INK4}$ function. *Nat. Genet.* **10**, 114-116 (1995).
24. Reymond, A. & Brent, R. p16 proteins from melanoma-prone families are deficient in binding to CDK4. *Oncogene* **11**, 1173-1178 (1996).
25. Tang, K.S., Guralnick, B.J., Wang, W.K., Fersht, A.R. & Itzhaki, L.S. Stability and folding of the tumour suppressor protein p16. *J. Mol. Biol.* **285**, 1869-1886 (1999).
26. Tevelev, A. *et al.* Tumor suppressor $p16^{INK4}$: structural charcterization of wild-type and mutant proteins by NMR and circular dichroism. *Biochemistry* **35**, 9475-9487 (1996).
27. Yang, R., Gombart, A.F., Serrano, M. & Koeffler, P.H. Mutational effects on the $p16^{INK4a}$ tumor suppressor protein. *Cancer Res.* **55**, 2503-2506 (1995).
28. Zhang, B. & Peng, Z.-y. Defective folding of mutant $p16^{INK4}$ proteins encoded by tumor-derived alleles. *J. Biol. Chem.* **271**, 28734-28737 (1996).
29. Ruas, M., Brookes, S., McDonald, N.Q. & Peters, G. Functional evaluation of tumour-specific variants of p16INK4a/CDKN2A: correlation with protein structure information. *Oncogene* **18**, 5423-34. (1999).
30. Baumgartner, R. *et al.* Structure of human cyclin-dependent kinase inhibitor $p19^{INK4d}$: comparison to known ankyrin-repeat-containing structures and implications for the dysfunction of tumor suppressor $p16^{INK4a}$. *Structure* **6**, 1279-90. (1998).
31. Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D. & Pavletich, N.P. Structural basis for inhibition of the cyclin-dependent kinase CDK6 by the tumour suppressor $p16^{INK4a}$. *Nature* **395**, 237-243 (1998).
32. Jeffrey, P.D., Tong, L. & Pavletich, N.P. Strcutural basis of inhibition of CDK-cyclin complexes by INK4 inhibitors. *Genes. Dev.* **14**, 3115-3125 (2000).
33. Pavletich, N.P. Mechanisms of Cyclin Dependent Kinase regulation: structures of CDKs, their cyclin activators, and Cip and INK4 Inhibitors. *J. Mol. Biol.* **287**, 821-828 (1999).
34. Foster, B.A., Coffey, H.A., Morin, M.J. & Rastinejad, F. Pharmacological rescue of mutant p53 conformation and function. *Science* **286**, 2507-10. (1999).
35. Jones, T.A. & Kjeldguard, M. *Manual for O, Version 6.1*, (Aarhus University Press, Uppsula, Sweden., 1996).

36. Creighton, T.E. *Proteins*, 147-148 (W.H. Freeman and Company, New York, 1993).
37. Botherton, D.H. *et al.* Crystal structure of the complex of the cyclin D-dependent kinase CDK6 bound to the cell-cycle inhibitor $p19^{INK4d}$. *Nature* **395**, 244-250 (1998).
38. Foulkes, W.D., Flanders, T.Y., Pollock, P.M. & Hayward, N.K. The CDKN2A (p16) gene and human cancer. *Mol. Med.* **3**, 5-20 (1997).
39. Byeon, I.J. *et al.* Tumor suppressor $p16^{INK4A}$: determination of solution structure and analyses of its interaction with cyclin-dependent kinase 4. *Mol. Cell.* **1**, 421-31. (1998).
40. Lawrence, M.C. & Bourke, P. A program for generating electron density isosurfaces from Fourier syntheses in protein crystallography. *J. App. Cryst.* **33**, 990-991 (2000).
41. Kraulis, P.J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950 (1991).
42. Merritt, E.A. & Murphy, M.E.P. RASTER3D version 2.0: a program for photorealistic molecular graphics. *Acta Crystallogr.* **D50**, 869-873 (1994).
43. Pace, C.N. Determination and analysis of Urea and Guanidine Hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266-280 (1986).
44. Otwinowski, Z. Oscillation data reduction program. in *Proceedings of the CCP4 Study Weekend: Data collection and processing.* (eds. Sawyer, L., Isaacs, N. & Bailey, S.) 56-62 (SERC Daresbury Laboratory, Warrington, U.K., 1993).
45. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905-921 (1998).
46. Jones, T.A., Zou, J.Y. & Cowen, S.W. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**, 110-119 (1991).
47. Brunger, A.T. & Krukowski, A. Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr.* **A46**, 585-593 (1990).
48. Rice, L.M. & Brunger, A.T. Torsion angle dynamics: Reduced variable conformational sampling enhances crystallographic structure refinement. *Proteins* **19**, 277-290 (1994).
49. Brunger, A.T., Kuriyan, J. & Karplus, M. Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458-460 (1987).
50. Somasundaram, K. *et al.* Arrest of the cell cycle by the tumour-suppressor BRCA1 requires the CDK-inhibitor $p21^{WAF1/CIP1}$. *Nature* **389**, 187-190 (1997).
51. van den Heuvel, S. & Harlow, E. Distinct roles for cyclin-dependent kinases in cell cycle control. *Science* **262**, 2050-2054 (1993).