Solution Structure of the p53 Regulatory Domain of the p19^{Arf} Tumor Suppressor Protein†

Enrico L. DiGiammarino,‡ Igor Filippov,‡ Jason D. Weber,§,|| Brian Bothner,⊥ and Richard W. Kriwacki*,‡,⊥

Departments of Structural Biology and Tumor Cell Biology and Howard Hughes Medical Institute, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, and Department of Molecular Sciences, University of Tennessee Health Sciences Center, Memphis, Tennessee 38163

Received October 16, 2000; Revised Manuscript Received December 11, 2000

ABSTRACT: Arf is a tumor suppressor that regulates p53 function and is a frequent target for loss in human cancers. Through two novel mechanisms, Arf inhibits the oncoprotein Hdm2, a negative regulator of p53. (1) Arf inhibits the E3 ubiquitin ligase activity of Hdm2 that leads to p53 degradation, and (2) Arf sequesters Hdm2 within nucleoli. These activities of Arf promote p53-mediated cell cycle arrest and apoptosis. Fundamental to these processes are interactions between Arf and Hdm2. Here we show that a peptide containing the 37 N-terminal amino acids of mouse Arf (mArfN37) localizes to nucleoli, sequesters Hdm2 within nucleoli, and causes cell cycle arrest. Circular dichroism and NMR spectroscopy show that mArfN37 is largely unstructured under aqueous conditions; however, the peptide adopts two α-helices (helix 1, residues 4−14; and helix 2, residues 20−29) in 2,2,2-trifluoroethanol (TFE). Each helix contains an amino acid motif that is repeated twice in mArfN37, once in each helix. The two helices, however, do not interact but are connected by an apparently flexible linker. The repeated motif contains Arg residues spaced by a hydrophobic segment that may be involved in Hdm2 recognition and binding. The RRPR nucleolar localization signal, contained within residues 31−34, appears to be disordered under all conditions. The identification of two Arf structural modules suggests that short peptides containing the repeated motif may function as Arf mimics and may allow the design of small molecule Arf mimics in the future.

Disruption of cell cycle control mechanisms contributes significantly to the development of cancer in humans (*1*). The *INK4a/Arf* gene locus has been shown to encode two unrelated proteins, p16^{INK4A} and Arf,[1] from alternative but partially overlapping reading frames (*2*). These proteins independently target two cell cycle control pathways; in humans, p16^{INK4a} inhibits Cdk4 and Cdk6 within the Rb pathway (*3*, *4*) and p14^{Arf} targets the oncoprotein Hdm2 within the p53 pathway (*5*). Disruption in mice of either the entire *INK4a/Arf* locus (*6*) or only exon 1β that gives rise to p19^{Arf} (*7*), the mouse Arf ortholog, leads to multitype tumor growth and early death, identifying Arf as a *bona fide* tumor suppressor. Further, the *INK4a/Arf* locus is frequently disrupted in human cancers (*8*); however, a clear demarcation of effects from the *INK4a* and/or *Arf* portion is not yet available. The molecular details of p16^{INK4a} action are well understood as a result of cellular, biochemical, and structural studies (as reviewed in refs *8*−*10*), while the mechanistic details for Arf are not yet complete.

The current understanding is that Arf stabilizes p53 by inhibiting Hdm2 which, in the absence of Arf, mediates p53 degradation. Hdm2 is the human ortholog of mouse Mdm2 (*11*). The two protein sequences are ∼71% identical and 76% similar and have been shown to functionally substitute for each other (*12*). Mdm2 and Hdm2 regulate 26S proteosome-dependent degradation of p53 through several mechanisms. First, Mdm2 has been shown to possess E3 ubiquitin ligase activity toward p53 (*13*, *14*). Second, Mdm2 has been shown to be required for nucleoplasmic−cytoplasmic shuttling of p53 that leads, in turn, to degradation of p53 in the cytoplasm (*15*, *16*). Importantly, Arf inhibits these two negative regulatory mechanisms by binding directly to Hdm2 (*15*−*17*). In addition, Arf-dependent relocalization of Hdm2 also plays a role in p53 stabilization. It was first shown by Sherr and co-workers that alone Arf is localized to nucleoli and Hdm2 is in the nucleoplasm. When expressed together, Hdm2 is sequestered by Arf within nucleoli (*18*). Nucleolar colocalization, which requires the interaction of the N-

† This work was supported by the American Lebanese Syrian Associated Charities, the American Cancer Society, and Cancer Center (CORE) Support Grant CA 21765. E.L.D. is a Special Postdoctoral Fellow at St. Jude Children's Research Hospital. J.D.W. is an Associate of the Howard Hughes Medical Institute. B.B. acknowledges support from the University of Tennessee Health Sciences Center through the Hal and Alma Regean Fellowship in Cancer Research.

* To whom correspondence should be addressed. Phone: (901) 495-3290. Fax: (901) 495-3032. E-mail: richard.kriwacki@stjude.org.

‡ Department of Structural Biology, St. Jude Children's Research Hospital.

§ Department of Tumor Cell Biology, St. Jude Children's Research Hospital.

|| Howard Hughes Medical Institute, St. Jude Children's Research Hospital.

⊥ University of Tennessee Health Sciences Center.

[1] Abbreviations: 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; Arf, general reference to either full-length or functional fragments of p19^{Arf} or p14^{Arf}; CNS, crystallography and NMR system; MEF, mouse embryo fibroblast; mArfN37, 37 N-terminal amino acids of p19^{Arf}; NrLS, nucleolar localization signal; rmsd, root-mean-square deviation; TALOS, torsion angle likelihood obtained from shift and sequence similarity; TFA, trifluoroacetic acid; TFE, 2,2,2-trifluoroethanol; TKO, *Mdm2/p53/Arf* triple-null mouse embryo fibroblast cells.

10.1021/bi0024005 CCC: $20.00
Published on Web 01/30/2001

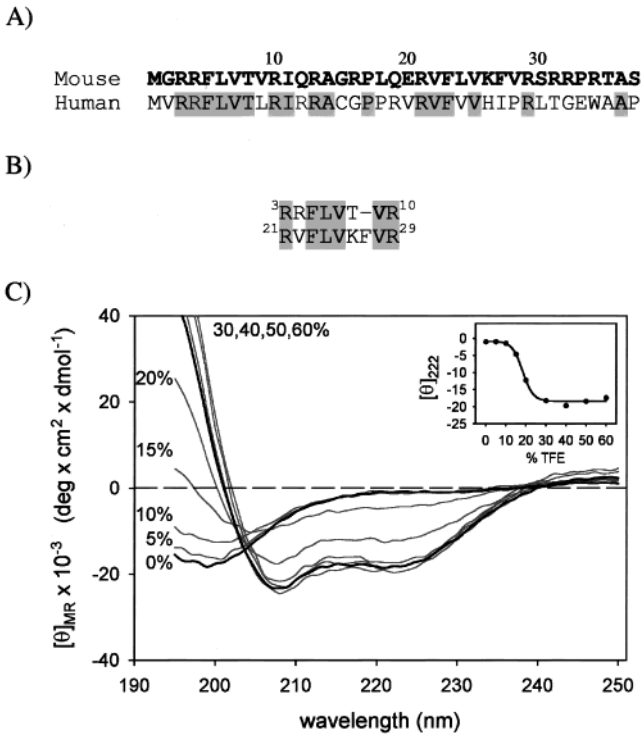

FIGURE 1: Comparison of Arf amino acid sequences and structural properties based on circular dichroism (CD). (A) Mouse (GenBank accession number AAC42080) and human (U26727) N-terminal 37-amino acid sequences are aligned with identical residues highlighted in gray (46% identical). (B) Alignment of the repeated amino acid motifs found within mArfN37 with identical residues highlighted in gray. (C) CD spectra for mArfN37 at different TFE concentrations. $[\theta]_{MRW}$ is the molar ellipticity per residue. Solution conditions were as follows: 10 mM KP_i (pH 5.0), 50 mM NaCl at 25 °C, and 0% TFE (black line) and the same conditions with 5, 10, 15, and 20% TFE (gray lines), 30% TFE (black line), and 40, 50, and 60% TFE (additional gray lines). The inset shows $[\theta]_{MRW}$ at 222 nm ($[\theta]_{222}$) vs % TFE.

terminal domain of Arf and a central acidic domain of Hdm2, is thought to inhibit Hdm2-dependent degradation of p53 by sequestering Hdm2 in a nuclear suborganelle that is physically isolated from p53 (*19*). Therefore, Arf appears to prevent the inhibitory effects of Hdm2 toward p53 directly by binding Hdm2 and indirectly by sequestering Hdm2 in nucleoli. Despite our growing knowledge of the mechanisms of Arf function, data have not previously been available to relate the three-dimensional structure of this domain and its biological function.

The mouse protein, $p19^{Arf}$, is composed of 169 amino acids, is both highly basic and hydrophobic, and localizes to nucleoli within cells (*18*). The human protein, $p14^{Arf}$, is composed of 132 amino acids, is also basic and hydrophobic, exhibits nucleolar localization (*20*), and shows the greatest degree of sequence similarity with $p19^{Arf}$ within an N-terminal domain of 14 amino acids (Figure 1A; 11 of 14 are identical); beyond this domain, the two protein sequences diverge. Two segments of $p19^{Arf}$, residues 1−14 and 26−37, have been shown to be important for nucleolar localization, Mdm2 nucleolar sequestration, and p53-dependent cell cycle arrest (*19*). Two segments of $p14^{Arf}$ mediate interactions with Hdm2, but one of these is found outside the segment of residues 1−37; these are residues 1−14 and 82−101 (*19*, *21*). Both Arf proteins are believed to act through similar mechanisms involving direct interactions with Mdm2 (Hdm2 in humans), leading to inhibition of p53 ubiquitination and degradation, and colocalization of Arf and Mdm2 in nucleoli. Arf proteins have distinct nucleolar localization signals (NrLSs) with the sequence RRPR (residues 31−34 in $p19^{Arf}$ and residues 87−90 in $p14^{Arf}$); however, when Arf interacts with Mdm2 (or Hdm2), the Arf NrLS is masked and nucleolar colocalization of the Arf−Mdm2 complex relies on the exposure of a cryptic NrLS within the RING domain of Mdm2. Prior to Arf binding, this NrLS is not functional (*19*, *20*). Importantly, the 37-amino acid N-terminal segment of $p19^{Arf}$ has been shown to be necessary and sufficient for nucleolar localization, Mdm2 binding and nucleolar colocalization, and activation of p53 followed by cell cycle arrest. In addition, a peptide comprised of human Arf amino acids 1−22 binds the central domain of Hdm2, inhibits Hdm2-mediated E3 activity toward p53 in vitro, and activates p53 in vivo (*17*), further emphasizing the importance of the N-terminus in Arf function. Here we show that a recombinant peptide containing amino acids 1−37 of mouse $p19^{Arf}$ (mArfN37) recapitulates previously observed Arf functions, as described above, and report its structural properties in aqueous solution and in the presence of the cosolvent 2,2,2-trifluoroethanol (TFE) based on CD and NMR analysis.

EXPERIMENTAL PROCEDURES

mArfN37 Protein Purification. mArfN37 was expressed in *Escherichia coli* with an N-terminal His tag using the expression vector pET28a-polyHIS-syn-ARF N37 (*19*) and purified as previously described with only minor changes. In brief, mArfN37 was purified from the soluble cellular fraction using Ni^{2+}-affinity chromatography in the presence of 6 M urea. mArfN37-containing fractions were further purified using reverse-phase C_4 high-performance liquid chromatography (HPLC) using a 0.1% TFA/2-propanol buffer system. Lyophilized mArfN37 was dissolved in 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 2.5 mM $CaCl_2$ and treated with thrombin to cleave the His tag followed by further purification using HPLC with a 0.1% TFA/acetonitrile buffer system. This step was necessary because secondary thrombin cleavage products are observed. The sequence of the recombinant mArfN37 peptide, corresponding to the 37 N-terminal amino acids of mouse Arf along with the sequence GSH at the N-terminus from pET28a, was confirmed by MALDI-TOF mass spectrometry.

Cell Culture and Virus Infection. NIH 3T3 cells (wild-type p53, *Arf*-null) were maintained in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum, 2 mM glutamine, and 100 units/mL penicillin and streptomycin. Primary mouse embryo fibroblasts (MEFs) explanted from TKO mice (*Arf/p53/mdm2*-null) were cultured in DMEM supplemented with 10% fetal bovine serum, 2 mM glutamine, 0.1 mM nonessential amino acids, 55 μM β-mercaptoethanol, and 10 μg/mL gentamycin (all from Gibco). Virus production and infection of cells were carried out using helper and vector plasmids (*22*) provided by C. Sawyers (University of California, Los Angeles, CA). Hdm2 cDNA was subcloned into the pSRα-MSV-tkCD8 retroviral vector.

Microinjection and Immunofluorescence. Primary MEFs from TKO mice (3×10^4) seeded onto glass coverslips were infected with retroviruses encoding CD8 or Hdm2. The

purified mArfN37 protein (2 mg/mL) was microinjected 48 h later into nuclei of infected cells supplemented with 10 mM HEPES and allowed to recover for an additional 4 h. Cells were fixed in a methanol/acetone mixture (1:1, v/v) and stained for 1 h with affinity-purified rabbit anti-p19^{Arf} antibody (raised against an Arf N-terminal epitope) followed by a 30 min exposure to biotinylated anti-rabbit immunoglobulin and streptavidin Texas red (both from Amersham). Cells were then stained for 1 h with affinity-purified mouse monoclonal anti-Hdm2 antibody (2A10) followed by a 30 min incubation with fluorescein isothiocyanate (FITC)-conjugated anti-mouse immunoglobulin (Amersham). DNA was visualized with Hoechst dye (Molecular Probes). For assessment of DNA replication, 5-bromodeoxyuridine (BrdU) (10 μM) was added to the culture medium 30 min after microinjection of NIH 3T3 cells with mArfN37 or dextran conjugated to Texas red. Cells were fixed 8 h later in a methanol/acetone mixture, treated for 10 min with 1.5 N HCl, and stained for 1 h with mouse monoclonal anti-BrdU antibody (Amersham) followed by FITC-conjugated anti-mouse immunoglobulin. Fluorescence signals were detected using a BX50 microscope (Olympus) fitted with a Sensys 1400 charge-coupled device camera (Photometrics).

CD Spectroscopy. Circular dichroism (CD) spectra were recorded at 25 °C using an AVIV model 62A DS circular dichroism spectrometer. Samples of mArfN37 were prepared in 10 mM KP_i (pH 5.0) and 50 mM NaCl with various amounts of TFE. Protein concentrations were determined by quantitative amino acid analysis.

NMR Spectroscopy. Due to the high Arg content of mArfN37 (10 of 37 residues), we prepared isotope-labeled protein samples for NMR studies. Uniformly ^{15}N- or ^{15}N/^{13}C-labeled mArfN37 was obtained by growing BL21(DE3) cells transformed with pET28a-polyHIS-syn-ARF N37 (*19*) in a defined medium (*23*) containing ^{13}C-glucose and/or ^{15}N-ammonium chloride and purifying the protein as described above. The mArfN37 samples for NMR spectroscopy were dissolved at 1−2 mM in 10 mM KP_i (pH 5.0), 50 mM NaCl, 10% D_2O (v/v), and 0.02% sodium azide and placed in Shigemi micro-cell NMR tubes (Shigemi, Inc.).

All NMR spectra were acquired using a Varian Inova 600 MHz spectrometer (Varian, Inc.) with a 5 mm triple-resonance probe equipped with *x*-, *y*-, and *z*-axis pulsed magnetic field gradients (PFGs). In general, pulse sequences provided by Varian in the Protein Pack suite were used. Additional experiments were implemented as described by Cavanagh et al. (*24*), with minor modifications to include the use of PFGs for artifact suppression and coherence order selection and sensitivity enhancement (*25*−*27*).

NMR Data Processing and Analysis. 2D and 3D NMR spectra were processed using the program Felix 98 (Molecular Simulations, Inc.). In general, time domain data were apodized using sine-bell or shifted sine-bell functions, followed by zero-filling, and Fourier transformation. TSP (10 mM) in 10 mM KP_i (pH 5.0) was used as an external standard. ^{13}C and ^{15}N chemical shifts were indirectly referenced using the appropriate gyromagnetic ratios (*24*).

Sequence-specific assignments of ^{1}HN, ^{15}NH, and ^{13}Cα nuclei were established through the analysis of 3D constant time (CT)-HN(CO)CA (*28*) and CT-HNCA (*29*) spectra. Side chain assignments were made through the analysis of 3D (H)C(CO)NH-TOCSY, H(CCO)NH-TOCSY, HCCH-COSY, and HCCH-TOCSY spectra. Resonance assignments are reported as Supporting Information. Distance restraints were obtained through the analysis of 3D ^{1}H−^{15}N and 3D ^{1}H−^{13}C NOESY-HSQC spectra obtained with a mixing time of 400 ms.

Structure Calculation and Refinement. mArfN37 structures were calculated using NMR data collected in the presence of 30% TFE at 40 °C. Interproton distances were estimated from the volumes of cross-peaks in 3D NOESY spectra. The restraint lower bounds were set to the van der Waals distance (1.8 Å), and upper bounds were set to 2.5, 3.5, and 6.0 Å for strong, medium, and weak resonances, respectively. Estimates of backbone dihedral angles ϕ and φ were obtained using the program TALOS (*30*) and ^{1}Hα, ^{15}N, ^{13}Cα, ^{13}Cβ, and ^{13}CO chemical shift values. ^{13}CO chemical shifts were obtained from a 3D HNCO spectrum. TALOS-based ϕ and φ restraints were used only when a well-defined prediction was obtained, and this occurred within the helical segments of mArfN37.

A total of 553 NOE interproton distance restraints and 21 ϕ and φ backbone dihedral angle restraints were used to calculate an ensemble of structures using the Torsion Angle Dynamics (TAD) routine (*31*) within CNS (*32*). The TAD protocol was as follows: 15 ps high-temperature TAD (50 000 K) followed by cooling to 1000 K over the course of 15 ps and ramping of the van der Waals scaling term from 0.1 to 1.0. The molecules were further cooled to 300 K over the course of 6 ps using conventional Cartesian dynamics followed finally by 1000 steps of conjugate gradient energy minimization. The NOE energy term was 150 kcal/mol for the first three steps and 100 kcal/mol for the last. The dihedral restraint term was 100 kcal/mol for the first three steps and 300 kcal/mol for the last. Two hundred structures were calculated, and the final structure statistics for the 20 lowest-energy structures are given in Table 1.

RESULTS

Functional Analysis of mArfN37. Nucleolar localization, Mdm2 colocalization, and the cell cycle arrest potential of recombinant mArfN37 were assessed by nuclear microinjection of Mdm2/p53/Arf triple-null (TKO) and Arf-null (NIH 3T3) MEFs. Microinjected mArfN37 localizes to nucleoli in TKO cells (Figure 2E−H). DAPI staining demarcates the nucleus (Figure 2, first row), while mArfN37 localizes to dense subnuclear bodies (Figure 2G). When TKO cells are infected with a retrovirus encoding Hdm2, the ectopically expressed Hdm2 is nuclear but not nucleolar (Figure 2A−D). In contrast, mArfN37 and Hdm2 colocalize to nucleoli (Figure 2I−L) in TKO cells that were first infected with retrovirus encoding Hdm2 and subsequently microinjected with mArfN37. BrdU incorporation was used to assay DNA replication in NIH 3T3 cells either with or without mArfN37 nuclear microinjection. Sixty-two percent (48 of 78) of dextran-injected control cells exhibited BrdU incorporation into replicating DNA, while 35% (24 of 68) of the mArfN37-injected cells exhibited BrdU incorporation. The decreased level of BrdU incorporation observed for mArfN37-injected cells is consistent with the ability of full-length mouse Arf to cause G1 cell cycle arrest (*18*). These results, collectively, show that recombinant mArfN37 used

Table 1: Structural Statistics[a]

total no. of NOEs	553
intraresidue	304
interresidue	249
sequential	169
medium-range	80
long-range	0
no. of dihedral restraints	
ψ	21
ϕ	21
rmsd from mean structure (Å)	
all residues	
backbone	7.57 ± 2.34
heavy atoms	8.92 ± 2.26
helix 1 (residues 4−14)	
backbone	0.38 ± 0.09
heavy atoms	1.90 ± 0.22
helix 2 (residues 20−29)	
backbone	0.22 ± 0.06
heavy atoms	1.80 ± 0.21
violations	
distance restraint violations (average no. per structure)	
no. >0.50 Å	4.25
torsion angle violations	
no. of ψ values >5°	0
no. of ϕ values >5°	0
maximum distance restraint violation (Å)	1.23
CNS energies (kcal/mol)	
total energy	25.74 ± 1.12
bond length	1.38 ± 0.15
bond angle	12.77 ± 0.21
improper	0.50 ± 0.05
van der Waals	5.10 ± 1.33
NOE	5.91 ± 1.39
dihedral	0.09 ± 0.06
Ramachandran statistics	
most favored region	78.7%
allowed region	15.3%
generously allowed region	3.8%
disallowed region	2.2%

[a] Statistics for the 20 lowest-energy structures from an original ensemble of 200 structures.

for structural studies exhibits key functional properties that are hallmarks of Arf activity, namely, nucleolar localization, sequestration of Hdm2 within nucleoli, and cell cycle arrest.

Circular Dichroism. Figure 1C shows CD spectra obtained for mArfN37 under aqueous conditions and with different TFE concentrations (5−60%, v/v). The CD spectrum for mArfN37 without TFE is characteristic of random coil conformations. A conformational change occurs upon TFE addition, and the CD spectra become increasingly characteristic of α-helical secondary structure. The spectral minimum shifts from 199 to 208 nm, and the molar ellipticity at 222 nm becomes increasingly negative as the TFE concentration approaches 30%. The inset of Figure 1C shows that the coil to α-helix transition is cooperative and is complete at 30% TFE. The α-helical content for mArfN37 in 30% TFE is estimated to be ∼50% (*33*).

NMR Spectroscopy. NMR spectra for mArfN37 were recorded under aqueous conditions and in the presence of various concentrations of TFE (5−40%, v/v). Quite unusually, the mArfN37 protein is both arginine rich (10 of 37) and highly hydrophobic (14 of 37). The high Arg content produced a great deal of spectral overlap in ^{1}H NMR spectra. To overcome this limitation, we prepared both ^{15}N and ^{13}C/^{15}N isotope-labeled forms of mArfN37 and used 2D and 3D heteronuclear NMR methods.

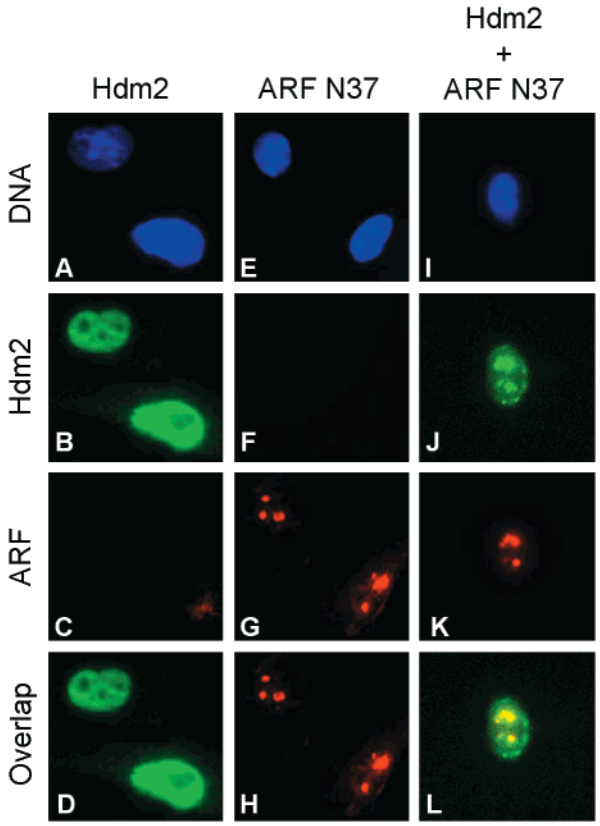

FIGURE 2: Functional analysis of recombinant mArfN37; nucleolar localization and sequestration of Hdm2. TKO MEFs were infected with Hdm2 (A−D and I−L) or CD8 (E−H) and were microinjected with PBS (A−D) or purified mArfN37 (E−L). The lettered squares show nuclear DNA staining (first row, blue), Hdm2 fluorescence (second row, green), ARF fluorescence (third row, red), and Hdm2 and ARF overlapping fluorescence (fourth row, yellow). Data are representative of more than 50 stained cells.

The 2D ^{1}H−^{15}N HSQC spectra (*24*) for ^{15}N-mArfN37 under aqueous conditions (4−40 °C) are characteristic of an unfolded peptide (Figure 3A,B), consistent with the CD results. However, addition of TFE causes a gradual increase in ^{1}H and ^{15}N resonance dispersion, with the maximum effect produced at 30% TFE (Figure 3C). To obtain detailed information about the structure, resonance assignments for ^{13}C-, ^{15}N-mArfN37 with and without 30% TFE were obtained (reported as Supporting Information). To analyze secondary structure in an amino acid-specific manner, we examined the deviation of ^{13}Cα chemical shifts from random coil values (termed the secondary ^{13}Cα chemical shift, $\Delta\delta^{13}$Cα) (Figure 4A,B); $\Delta\delta^{13}$Cα values greater than 2 ppm are indicative of α-helix secondary structure, while values less than −1.5 ppm are indicative of *β*-strand secondary structure (*34*). In aqueous solution, mArfN37 clearly lacks secondary structure (Figure 4A) while two segments of α-helix, between residues 3−12 and 18−29, are observed in the presence of 30% TFE (Figure 4B). The first helix is broken by the GRP motif (residues 15−17), and the second helix begins following P17 and is broken by the SRRPR segment that contains the nucleolar localization signal. Patterns of NOEs derived from heteronuclear 3D NOESY spectra are consistent with the presence of two α-helices (Figure 4C). The NOE patterns do not extend as far toward the N-terminus as the $\Delta\delta^{13}$Cα data

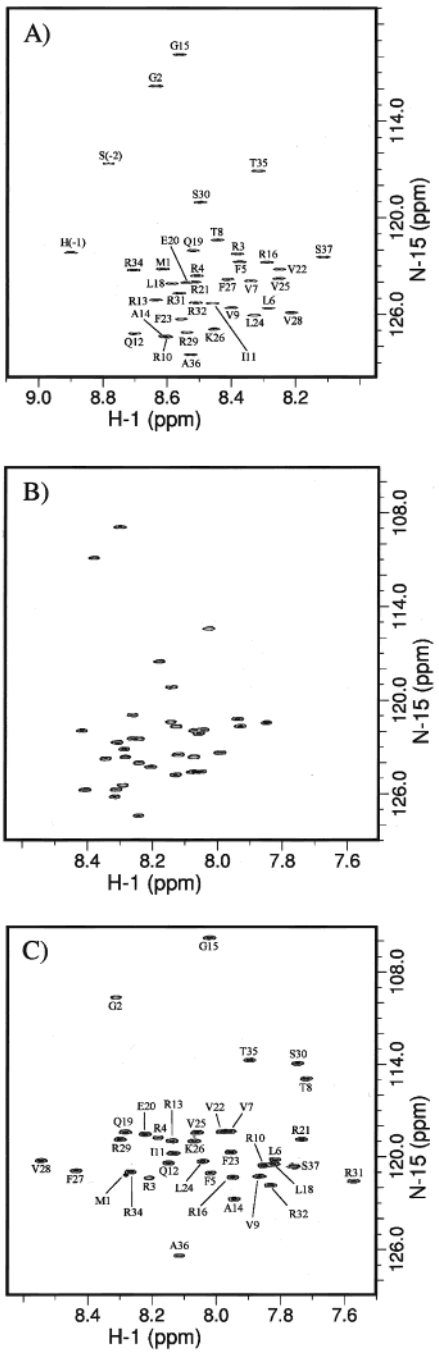

FIGURE 3: $^{1}H-^{15}N$ 2D HSQC spectra of ^{15}N-mArfN37 in 10 mM KP_i (pH 5.0) and 50 mM NaCl with and without TFE. (A) $^{1}H-^{15}N$ 2D HSQC spectrum of ^{15}N-mArfN37 at 4 °C in 0% TFE. Backbone assignments (^{1}HN, ^{15}N, and $^{13}C\alpha$) for mArfN37 under aqueous conditions (0% TFE) were made at 4 °C, and ^{1}HN and ^{15}N assignments are shown. (B) $^{1}H-^{15}N$ 2D HSQC spectrum of ^{15}N-mArfN37 at 40 °C in 0% TFE. (C) $^{1}H-^{15}N$ 2D HSQC spectrum of ^{15}N-mArfN37 at 40 °C in 30% TFE. Complete assignments (^{1}H, ^{15}N, and ^{13}C) for mArfN37 at 30% TFE were made at 40 °C, and ^{1}HN and ^{15}N assignments are shown.

would suggest because ^{1}H, ^{13}C, and ^{15}N resonances for Arg residues 3 and 4 are highly degenerate and precluded detailed analysis of NOEs in this region.

mArfN37 Solution Structure. The results of restrained molecular dynamics calculations using CNS (*32*) for mArfN37 show two well-defined α-helices; helix 1 is between residues 4 and 14, and helix 2 is between residues 20 and 29 (Figure 5A,B). The two helices, however, do not interact with one

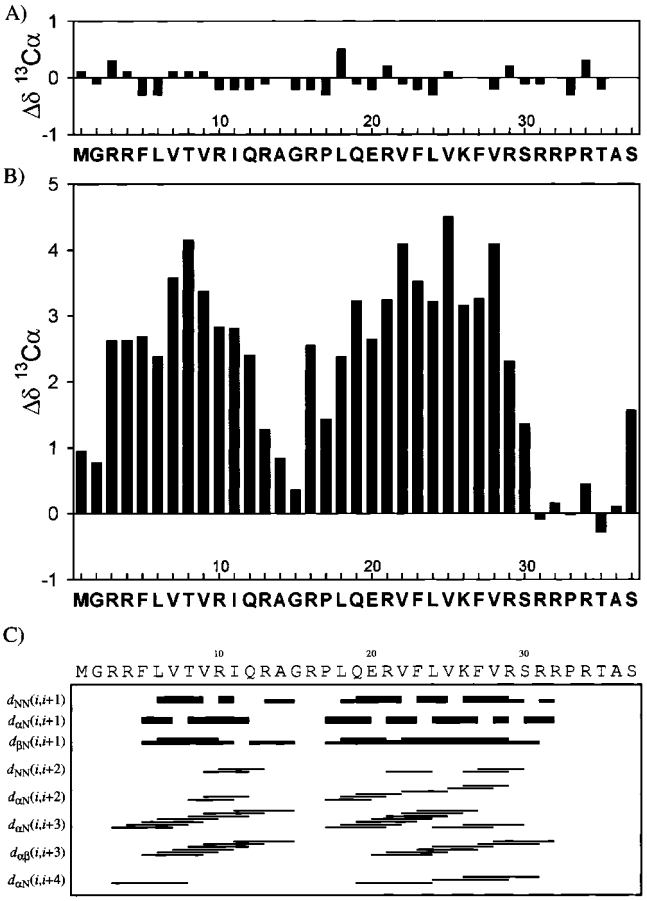

FIGURE 4: Analysis of mArfN37 secondary structure. Histograms of secondary $^{13}C\alpha$ chemical shifts ($\Delta\delta^{13}C\alpha$) for mArfN37 (A) at 4 °C with buffer and 0% TFE and (B) at 40 °C with buffer and 30% TFE. (C) Patterns of NOEs observed for mArfN37 in the presence of buffer and 30% TFE. Bar thickness indicates the relative strength of the NOEs observed between the indicated protons (thick, strong; thin, weak).

another to form a compact structure but rather are connected by an apparently flexible linker consisting of residues 15−17. We considered the possibility of a turn stabilized by a salt bridge between Arg16 and Glu20 which might allow the two helices to interact in an antiparallel manner. However, we found no evidence of this conformation. A comparison of panels A and B of Figure 5 indicates that the two helices are structurally independent. We have identified a repeated peptide motif within mArfN37, as illustrated in Figure 1B, and observe that both occurrences of this motif are within the helical segments shown in Figure 5C. Interestingly, the RRPR nucleolar localization signal seems to be in a flexible region of the peptide. Coordinates for the 20 lowest-energy structures reported here have been deposited with the PDB as entry 1HN3.

DISCUSSION

Arf has been identified as a tumor suppressor that protects cells against hyperproliferative signals (*5*). Oncogenic signals such as Myc (*35*) induce Arf which, in turn, stabilizes p53 to produce either cell cycle arrest or apoptosis (*5*). Moreover, Arf-null mice are prone to developing tumors, and Arf-null MEFs do not undergo replicative senescence (*7*). It has been suggested that virtually all human cancers possess genetic lesions within either p53 or Arf (*1*), placing Arf among the

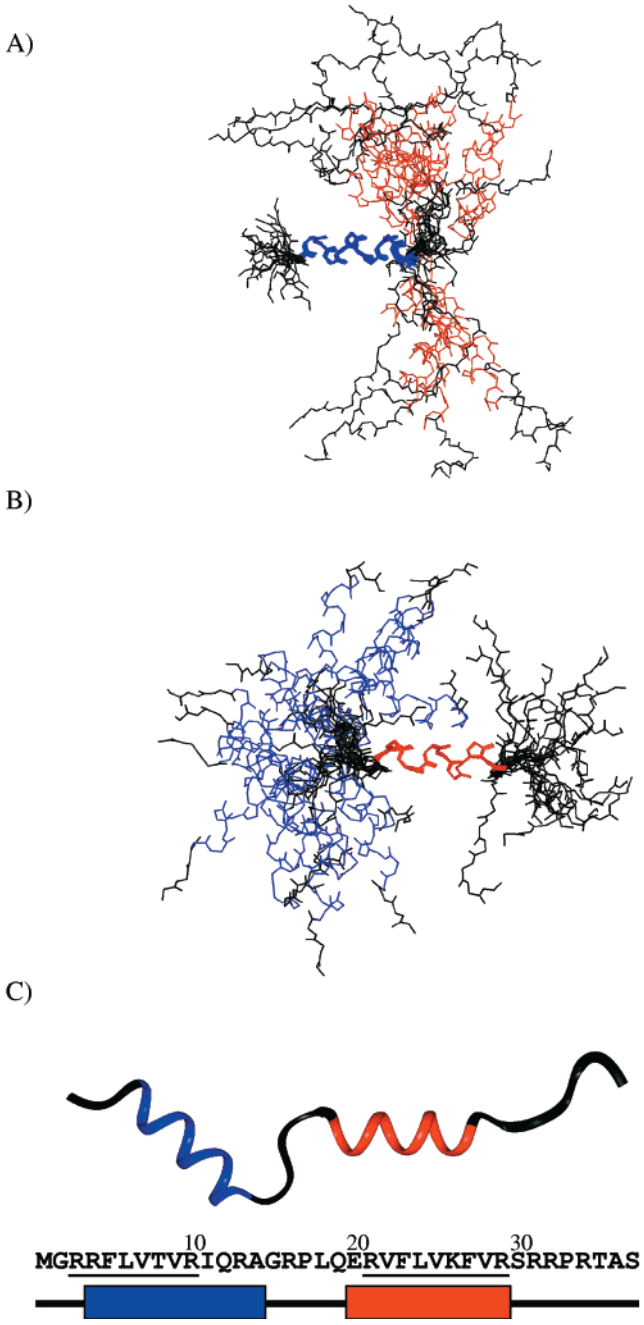

FIGURE 5: Structure of mArfN37 in buffer and 30% TFE. (A and B) Backbone renderings of the 20 lowest-energy structures superimposed for helix 1 (blue, residues 4−14) and for helix 2 (red, residues 20−29), respectively. (C) Ribbon diagram of one structure from the ensemble (colored as described above) with the amino acid sequence shown below. The blue and red rectangles below the sequence identify helices 1 and 2, respectively. The repeated amino acid motifs are underlined.

most important contributors to human neoplastic disease. Importantly, as discussed earlier, peptides comprised of either the 37 N-terminal amino acids of mouse Arf (this work) or the 22 N-terminal amino acids of human Arf (*17*) recapitulate Arf function in both in vitro and in vivo assays. Further, recent results from Zambetti and co-workers (*36*) reveal a p53-independent component of Arf-dependent cell cycle arrest and identify that this activity requires amino acids 1−14. While our understanding of Arf's genetic, cellular, and biochemical properties is quite extensive, very little is known about the molecular properties of the Arf polypeptide. Here we describe the structural properties of mArfN37 in solution and provide insights into Arf structure−function relationships.

mArfN37 is unstructured, or disordered, in aqueous solution. Despite this, we have shown that this polypeptide possesses functional properties that are hallmarks of Arf function, namely, the ability to localize within nucleoli, to sequester Mdm2 within nucleoli, and to produce cell cycle arrest. As has now been observed for a large number of proteins, conformational disorder can be associated with biological function (*33*, *37*); mArfN37 is clearly another example, exhibiting $\Delta\delta^{13}C\alpha$ values virtually indistinguishable from random coil values (Figure 4A). This finding indicates that, despite the existence of multiple conformational states in solution, mArfN37 is able to recognize and bind its biological target(s). The C-terminal portion of mouse Arf (amino acids 38−169) is dispensable with regard to known Arf function (*18*, *19*), highlighting the importance of the 37-amino acid N-terminal segment. One biological target of Arf is Mdm2 (Hdm2 in humans), and we have previously shown that mArfN37 can specifically bind Hdm2 and have mapped the interaction site to an ∼100-amino acid central segment of Hdm2 (residues 210−304) (*19*). Between mice and humans, this segment is 92% similar and, in striking contrast to Arg-rich Arf, is highly acidic (for Hdm2, 32% Asp/Glu, predicted pI ∼ 3.2; for Mdm2, 33% Asp/Glu, predicted pI ∼ 3.5). The predicted pI for mArfN37 is 12.6 due to its high Arg content (10 of 37, or 27%). On the basis of these unusual but coordinated charge characteristics for Hdm2 and Arf, we propose that charge−charge interactions mediate Hdm2−Arf interactions. However, due to charge repulsion, the free proteins may be unstructured, as we observe for mArfN37.

In the presence of 30% TFE, mArfN37 adopts bihelical structure. The use of TFE in structural studies and the extrapolation of TFE-induced structural results to biological situations can be problematic because the energetic factors that govern polypeptide conformations in TFE are different from those active in biological systems. However, extensive study of peptides and proteins in the presence of TFE and other alcohol cosolvents reveals a correlation between TFE-induced structure and native structure (*38*). By enhancing the energy contribution from local, intrapeptide hydrogen bonds, TFE preferentially stabilizes peptides in helical conformations. It must be appreciated that TFE seems to stabilize structures that have an inherent tendency to form on the basis of the local chemical properties of the polypeptide chain rather than to indiscriminately drive all polypeptides to form α-helices (*38*). Notable exceptions are peptides excised from proteins that in the native state form β-strands but, when stabilization due to tertiary interactions is eliminated, form α-helices in the presence of TFE (*39*−*42*). Importantly, however, in these examples the inherent preference of the peptide sequences is toward α-helix but the tertiary context of the folded proteins overrides this to give β-structure. In contrast, several studies of peptides have shown that alcohol cosolvents stabilize β-hairpin and β-strand conformations that have been observed for the same sequences in the context of native proteins (*38*). TFE can be thought to accentuate natural conformational preferences that stem from local sequence context. Therefore, our structural results for mArfN37 in the presence of TFE provide insight

into local structural propensities for this peptide. The fact that two regions of mArfN37 possess similar sequences and adopt similar structures in TFE is important and suggests that the biologically relevant structures for these regions, whether they are α-helical *in vivo*, are likely to be similar to each other.

The structural propensities for amino acids of mArfN37, based on statistical relationships established by Chou and Fasman (*43*), fail to provide clear insights into secondary structure, with α-helix and *β*-strand conformations predicted to be equally likely for residues 5−13 and 18−28. The GRP motif between residues 15 and 17, however, is predicted to break secondary structure. The thermodynamic α-helix propensity scale established by O'Niel and DeGrado (*44*) provides a more clear prediction of two α-helices between residues 3−14 and 18−29 because Arg is thermodynamically favored to adopt helical conformations. The abundance of Arg residues in mArfN37 thus favors two α-helices. In the presence of 30% TFE, then, mArfN37 adopts conformations that can be predicted on thermodynamic grounds. The difference in structure under aqueous conditions versus in the presence of TFE is striking, however, and indicates that factors beyond local conformational preferences govern the structure of mArfN37 in solution. Charge repulsion between the multiple Arg residues could favor an extended structure, and TFE may override these energetic factors by stabilizing local hydrogen bonds.

As we have previously shown (*19*), two peptide segments within mArfN37 (residues 1−14 and 26−37) independently participate in interactions between mouse Arf and Hdm2. The first of these segments brackets the first helix that we observe, while the second segment approximately bisects the second α-helix. Further, both of the repeated peptide motifs discussed earlier (Figure 1B) are contained within the TFE-induced helices; the previously identified segment of residues 1−14 falls within the first of these, while the segment of residues 26−37 approximately bisects the second repeated motif. We suggest that the common structural properties associated with this repeated motif (helicity in the presence TFE) are related to a common Hdm2 binding function and that this motif, repeated twice, functions in Hdm2 recognition. Our results also suggest that the RRPR NrLS possesses structural propensities different from the two helical segments and may function in nucleolar localization in a structurally different manner.

We have shown that mArfN37 sequesters Hdm2 within nucleoli, and it is well-known that this sequestration is associated with stabilization and activation of p53 that subsequently leads to cell cycle arrest; mArfN37, therefore, constitutes the p53 regulatory domain of Arf. What, then, is the biologically relevant conformation of mArfN37? Our solution NMR data clearly show that mArfN37, under aqueous conditions, is disordered in the absence of a biological target. The fact that mArfN37 adopts two α-helices in the presence of TFE indicates that these structures are energetically accessible. However, since the amino acid composition of mArfN37 is also consistent with the formation of *β*-strands, it is not possible to predict with confidence the biologically relevant, bound-state Arf conformation. An emerging trend, however, is that disordered proteins generally adopt highly extended conformations, with or without elements of regular secondary structure, upon binding their biological targets. Examples are the cyclin-dependent kinase inhibitors; p21$^{Waf1/Cip1}$ and p27^{Kip1} are disordered in solution (*33*), and p27 adopts a highly extended conformation when bound to cyclin A/Cdk2 (*45*). The association of p21 or p27 with cyclin A/Cdk2 is driven by the burial of an extensive hydrophobic interface. Our structural analysis of mArfN37 is consistent with the idea that Arf contains two Mdm2 contact sites and that these two regions adopt structure upon interaction with Mdm2. Due to the similarity of the amino acid sequences within these domains, they are likely to adopt similar bound structures. To maximize the burial of surface area and to maximize the opportunity for electrostatic interactions between Arg residues of Arf and Asp or Glu residues of Mdm2, the two proteins may adopt extended conformations when they interact. The disorder observed for mArfN37 in aqueous solution is consistent with this suggestion. Our results showing that mArfN37 forms two α-helices in 30% TFE indicate that, under certain energetic circumstances, these conformations are stable and, therefore, cannot be ruled out as a possibility for the biological conformation of the Arf N-terminus in vivo. The finding that the segments that contain the repeated Arf motif adopt similar helical conformations strengthens the view that the bound-state conformations for these two segments will be similar.

The identification of two Arf structural modules suggests that short peptides containing the repeated motif may function as Arf mimics and may provide opportunities for the design of small molecule Arf mimics in the future. Alternatively, small molecules that bind these Arf modules may disrupt normal Arf function and produce a favorable therapeutic effect by preventing cell cycle arrest normally associated with aberrant mitogenic (oncogenic) signals. Our studies provide clear directions for drug discovery. First, the relevance and potential therapeutic impact of helical Arf conformations can be tested through the design of conformationally restricted Arf peptide analogues. Similarly, rigid and extended analogues will test the biological role of these alternative Arf conformers. These restricted analogues could then serve as leads for the discovery of small molecule Arf mimics. It is already known that a 22-amino acid human Arf peptide produces cell cycle arrest in cells in culture (*17*). This peptide contains only one of the structural modules we have identified, suggesting that only one module is required for biological function. However, if small molecule Arf mimics can be developed, perhaps tethering them as the two modules are tethered in mouse Arf will enhance biological effects. We are pursuing these ideas using both experimental and theoretical approaches.

ACKNOWLEDGMENT

We thank Dr. C. Sherr for encouragement and insightful discussions, Dr. W. Zhang for management of the St. Jude NMR facility, Mr. Limin Xiao for assistance with preparation of Arf samples, Dr. Jill Lahti and Ms. Sarah Bothner for assistance with fluorescence imaging, and other members of the Kriwacki laboratory for comments and discussion.

SUPPORTING INFORMATION AVAILABLE

^{15}NH, HN, and ^{13}Cα backbone resonance assignments for mArfN37 in aqueous solution and complete assignments in 30% TFE. This material is available free of charge via the

Internet at http://pubs.acs.org.

REFERENCES

1. Sherr, C. J. (2000) *Cancer Res. 60*, 3689−3695.
2. Quelle, D. E., Zindy, F., Ashmun, R. A., and Sherr, C. J. (1995) *Cell 83*, 993−1000.
3. Dyson, N. (1998) *Genes Dev. 12*, 2245−2262.
4. Serrano, M., Hannon, G. J., and Beach, D. (1993) *Nature 366*, 704−707.
5. Sherr, C. J. (1998) *Genes Dev. 12*, 2984−2991.
6. Serrano, M., Lee, H., Chin, L., Cordon-Cardo, C., Beach, D., and DePinho, R. A. (1996) *Cell 85*, 27−37.
7. Kamijo, T., Zindy, F., Roussel, M. F., Quelle, D. E., Downing, J. R., Ashmun, R. A., Grosveld, G., and Sherr, C. J. (1997) *Cell 91*, 649−659.
8. Raus, M., and Peters, G. (1998) *Biochim. Biophys. Acta 1378*, F115−F177.
9. Pavletich, N. P. (1999) *J. Mol. Biol. 287*, 821−828.
10. Sherr, C. J., and Roberts, J. M. (1999) *Genes Dev. 13*, 1501−1512.
11. Fakharzadeh, S. S., Trusko, S. P., and George, D. L. (1991) *EMBO J. 10*, 1565−1569.
12. Momand, J., Wu, H. H., and Dasgupta, G. (2000) *Gene 242*, 15−29.
13. Honda, R., Tanaka, H., and Yasuda, H. (1997) *FEBS Lett. 420*, 25−27.
14. Honda, R., and Yasuda, H. (1999) *EMBO J. 18*, 22−27.
15. Freedman, D. A., and Levine, A. J. (1998) *Mol. Cell. Biol. 18*, 7288−7293.
16. Tao, W., and Levine, A. J. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 6937−6941.
17. Midgley, C. A., Desterro, J. M. P., Saville, M. K., Howard, S., Sparks, A., Hay, R. T., and Lane, D. P. (2000) *Oncogene 19*, 2312−2323.
18. Weber, J. D., Taylor, L. J., Roussel, M. F., Sherr, C. J., and Bar-Sagi, D. (1999) *Nat. Cell Biol. 1*, 20−26.
19. Weber, J. D., Kuo, M. L., Bothner, B., DiGiammarino, E. L., Kriwacki, R. W., Roussel, M. F., and Sherr, C. J. (2000) *Mol. Cell. Biol. 20*, 2517−2528.
20. Lohrum, M. A., Ashcroft, M., Kubbutat, M. H., and Vousden, K. H. (2000) *Nat. Cell Biol. 2*, 179−181.
21. Zhang, Y., and Xiong, Y. (1999) *Mol. Cell 3*, 579−591.
22. Roussel, M. F., Theodoras, A. M., Pagano, M., and Sherr, C. J. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6837−6841.
23. Neidhardt, F. C., Bloch, P. F., and Smith, D. F. (1974) *J. Bacteriol. 119*, 736−747.
24. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996) *Protein NMR Spectroscopy*, Academic Press, New York.
25. Cavanagh, J., and Rance, M. (1992) *J. Magn. Reson. 96*, 670−678.
26. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
27. Palmer, A. G., III, Rance, M., and Wright, P. E. (1991) *J. Am. Chem. Soc. 113*, 4371−4380.
28. Yamazaki, T., Lee, W., Arrowsmith, C. H., Muhandiram, D. R., and Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 11655−11666.
29. Yamazaki, T., Lee, W., Revington, M., Mattiello, D. L., Dahlquist, F. W., Arrowsmith, C. H., and Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 6464−6465.
30. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR 13*, 289−302.
31. Stein, E. G., Rice, L. M., and Brünger, A. T. (1997) *J. Magn. Reson., Ser. B 124*, 154−164.
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
33. Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I., and Wright, P. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 11504−11509.
34. Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.
35. Eischen, C. M., Weber, J. D., Roussel, M. F., Sherr, C. J., and Cleveland, J. L. (1999) *Genes Dev. 13*, 2658−2669.
36. Weber, J. D., Jeffers, J. R., Rehg, J. E., Randle, D. H., Lozano, G., Roussel, M. F., Sherr, C. J., and Zambetti, G. P. (2000) *Genes Dev. 14*, 2358−2365.
37. Dyson, H. J., and Wright, P. E. (1998) *Nat. Struct. Biol. 5*, 499−503.
38. Buck, M. (1998) *Q. Rev. Biophys. 31*, 297−355.
39. Buck, M., Radford, S. E., and Dobson, C. M. (1993) *Biochemistry 32*, 669−678.
40. Buck, M., Schwalbe, H., and Dobson, C. M. (1995) *Biochemistry 34*, 13219−13232.
41. Kuroda, Y., Hamada, D., Tanaka, T., and Goto, Y. (1996) *Folding Des. 1*, 255−263.
42. Yang, J. J., Pikeathly, M., and Radford, S. E. (1994) *Biochemistry 33*, 7345−7353.
43. Chou, P. Y., and Fasman, G. D. (1974) *Biochemistry 13*, 211−222.
44. O'Neil, K. T., and DeGrado, W. F. (1990) *Science 250*, 646−651.
45. Russo, A. A., Jeffrey, P. D., Patten, A. K., Massague, J., and Pavletich, N. P. (1996) *Nature 382*, 325−331.

BI0024005