⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 34, pp. 23326–23332, August 22, 2008

Structural Consequences of Nucleophosmin Mutations in Acute Myeloid Leukemia*

Received for publication, March 3, 2008, and in revised form, May 23, 2008 Published, JBC Papers in Press, May 29, 2008, DOI 10.1074/jbc.M801706200

Charles G. Grummitt[‡], Fiona M. Townsley[‡], Christopher M. Johnson[‡], Alan J. Warren[§¶], and Mark Bycroft[‡1]

From the [‡]Medical Research Council Centre for Protein Engineering and [§]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH and the [¶]Department of Haematology, University of Cambridge, Hills Road, Cambridge CB2 2XY, United Kingdom

Mutations affecting NPM1 (nucleophosmin) are the most common genetic lesions found in acute myeloid leukemia (AML). NPM1 is one of the most abundant proteins found in the nucleolus and has links to the MDM2/p53 tumor suppressor pathway. A distinctive feature of NPM1 mutants in AML is their aberrant localization to the cytoplasm of leukemic cells. This mutant phenotype is the result of the substitution of several C-terminal residues, including one or two conserved tryptophan residues, with a leucine-rich nuclear export signal. The exact molecular mechanism underlying the loss of nucleolar retention, and the role of the tryptophans, remains unknown. In this study we have determined the structure of an independently folded globular domain in the C terminus of NPM1 using NMR spectroscopy, and we report that the conserved tryptophans are critical for structure. This domain is necessary for the nucleolar targeting of NPM1 and is disrupted by mutations in AML with cytoplasmic NPM1. Furthermore, we identify conserved surface-exposed lysine residues that are functionally rather than structurally important for nucleolar localization. This study provides new focus for efforts to understand the pathogenesis of AML with cytoplasmic NPM1 and may be used to aid the design of small molecules that target the C-terminal domain of NPM1 to act as novel anti-proliferative and anti-leukemia therapeutics.

The gene encoding nucleophosmin (*NPM1*) is mutated in 50–60% of normal karotype adult acute myeloid leukemia (AML),[2] making *NPM1* mutations the most common genetic lesions identified to date in *de novo* AML. NPM1 (nucleophosmin, also known as B23 or numatrin) is a ubiquitously expressed protein that belongs to the nucleoplasmin family of nuclear chaperones. Found predominantly in the nucleolus, NPM1 is used as a marker for the granular component (1–3) (the site of pre-ribosomal subunit assembly (4)). NPM1 rapidly shuttles between the nucleoplasm and the cytoplasm (5–7) serving as a chaperone for both nucleic acids and proteins (8–11). NPM1 is essential for embryonic development (12) and appears to modulate diverse molecular functions, including ribosome biogenesis (13–16) and genomic stability (12). NPM1 is overexpressed in a number of solid tumors (17), is linked to the p53/MDM2 pathway via the tumor suppressor ARF (18–22), and is a transcriptional target of Myc (23, 24).

A distinctive feature of NPM1 mutants in AML is their aberrant localization to the cytoplasm of leukemic cells (NPMc+ AML) (25). Two consequences of mutation-related sequence changes at the C terminus of NPM1 promote aberrant cytoplasmic localization of mutant NPM1 proteins in NPMc+ AML. First, generation of an additional leucine-rich nuclear export motif reinforces Crm1 (XpoI)-dependent nuclear export of NPM1 leukemic mutants (26). Second, there is loss of tryptophan residues Trp^{288} and Trp^{290} (or Trp^{290} alone) that determine NPM1 nucleolar localization (25, 27, 28). However, the precise molecular mechanisms that regulate the steady state distribution of NPM1 are unknown. In this study, we tested the hypothesis that NPM1 tryptophans Trp^{288} and Trp^{290} are critical structural components of an independently folded domain that is essential for nucleolar targeting of NPM1. These studies provide a structural basis for understanding the mechanisms underlying NPMc+ AML.

EXPERIMENTAL PROCEDURES

Reagents—Mutants NPM1 W288A/W290A and NPM1 F268A/F276A were made by site-directed mutagenesis and mutants A and E by PCR using appropriate primers. NPM1 C-terminal domain variants, with the exception of mutant A, were cloned into a modified pRSETA vector (Invitrogen) expressing the protein fused to a tobacco etch virus protease-cleavable His_6-tagged lipoyl domain from *Bacillus stearothermophilus*. Proteins were expressed in *Escherichia coli* strain C41 and purified using Ni^{2+}-chelating affinity chromatography. Fusion proteins were cleaved by tobacco etch virus protease digestion and removed in a second affinity step before concentration and final purification by size exclusion chromatography on a GF30 column (Amersham Biosciences). For isotopic labeling, K-MOPS minimal media containing $^{15}NH_4Cl$ and/or [^{13}C]glucose was used. Mutant A C-terminal domain was made using a Liberty CEM microwave peptide synthesizer (Matthews, NC) and standard Fmoc (*N*-(9-fluorenyl)methoxy-

* This work was supported in part by the Leukaemia Research Fund. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

⌘ *Author's Choice*—Final version full access.

The atomic coordinates and structure factors (code 2vxd) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence should be addressed. Tel.: 44-1223-402129; Fax: 44-1223-402140; E-mail: mxb@mrc-lmb.cam.ac.uk.

[2] The abbreviations used are: AML, acute myeloid leukemia; NES, nuclear export signal; NOESY, nuclear Overhauser effect spectroscopy; eGFP, enhanced green fluorescent protein; MES, 4-morpholineethanesulfonic acid; PBS, phosphate-buffered saline; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single quantum correlation spectroscopy; MOPS, 4-morpholinepropanesulfonic acid.

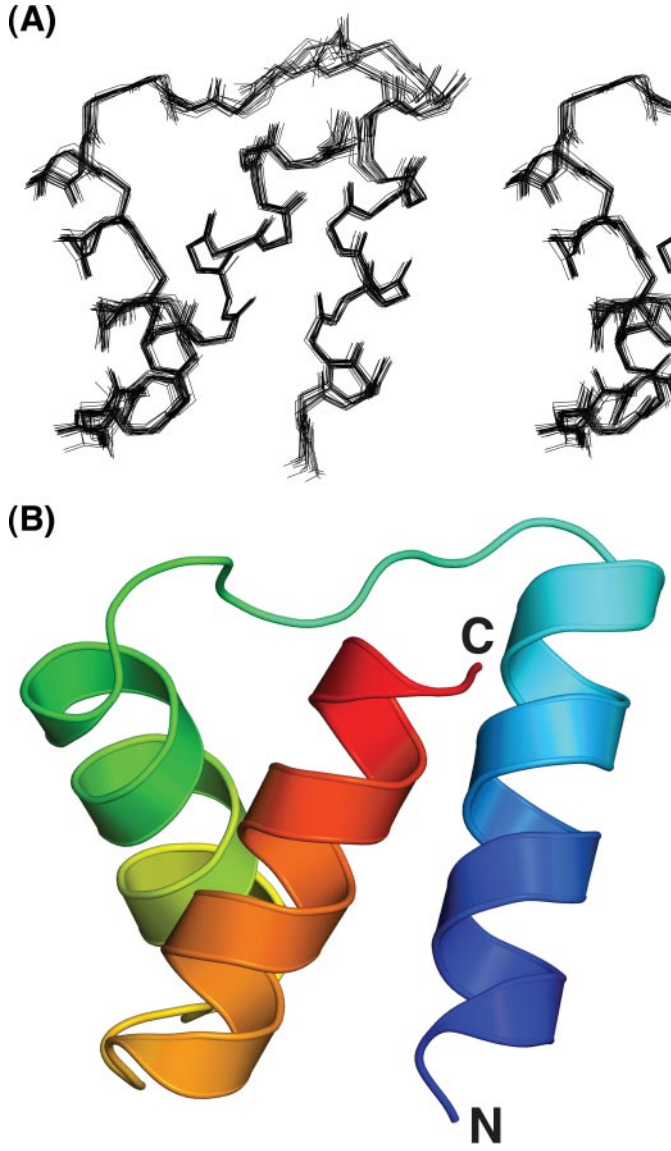

FIGURE 1. **Solution structure of the C-terminal domain of nucleophosmin.** *A,* overlay of the backbone atoms of the 20 lowest energy structures in stereo. *B,* schematic representation of the lowest energy structure colored *blue* to *red* from the N to C terminus (same orientation as in *A*). Prepared using the program PyMOL (53).

carbonyl)/t-butyl chemistry and purified by reverse-phase high pressure liquid chromatography. All C-terminal constructs have an additional two glycine residues at the N terminus as a result of the cloning and purification strategy; residue numbering is in context of the full-length protein and omits these glycines.

SDS-PAGE and electrospray mass spectrometry were used to confirm proteins were of the expected mass and to assess their purity. Protein concentration was determined spectroscopically using a Varian Cary 500 spectrometer (Varian, NC), and extinction coefficients were calculated for 280 nm (29, 30). All reagents were purchased from Sigma and were Anal-R grade or higher, except for isotopically enriched reagents, which were purchased from Cambridge Isotopes (Cambridge, UK).

Spectroscopic Measurements—The NMR spectra were recorded on Bruker Advance-800, Advance-II$^+$ 700, and Advance-600 spectrometers. Two-dimensional NOESY, TOCSY, DQF-COSY, ^{15}N-HSQC, constant time ^{13}C-HSQC and ^{13}C-TOCSY were recorded at 298 K. The mixing times chosen were 80 ms for TOCSY and 120 ms for NOESY. Spectra were referenced relative to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate for signals of proton and carbon, or liquid ammonium for that of nitrogen. Approximately half the Hb resonances were assigned stereospecifically using a combination of HNHB and DQF-COSY spectra. All the Val(Hγ) and Leu(Hδ) resonances were assigned stereospecifically using a 10% ^{13}C-labeled sample of NPM1 C-terminal domain (31). All the NMR spectra were analyzed with Sparky (32).

All samples used for structure determination were 1.2 mM and prepared in 10 mM MES, pH 6.5, 150 mM NaCl, 5 mM dithiothreitol, 0.02% sodium azide, and 5% D_2O (homonuclear samples were in a deuterated version of this buffer). For hydrogen exchange experiments, the ^{15}N-labeled NPM1 C-terminal domain was exchanged into NMR buffer containing 100% D_2O using a NAP-10 column (Amersham Biosciences), and a series of ^{1}H-^{15}N-FHSQC spectra (33) were recorded over the course of 6 h. All one-dimensional ^{1}H NMR experiments were conducted in biophysics buffer (10 mM KP_i, pH 6.5, 150 mM NaCl, 5% D_2O).

Structure Determination—The distance constraints derived from the NOESY spectra were classified into four categories corresponding to inter-proton distance constraints of 1.8–2.8, 1.8–3.5, 1.8–4.75, and 1.8–6.0 Å, respectively. Hydrogen bond constraints of 1.8–2.1 Å were imposed on the distance between the hydrogen and the acceptor oxygen, whereas another constraint of 2.7–3.1 Å was imposed on the distance between the donor nitrogen and the acceptor oxygen. Torsion angle constraints were obtained from stereospecific assignment of residue side chains and incorporated in the structure calculation, along with the backbone ϕ and ψ angle constraints determined with the program TALOS (34). The structures were calculated using a standard torsion angle dynamics simulated annealing protocol with the program CNS (35). 150 structures were accepted out of 150, produced with no distance violations greater than 0.25 Å, and no angle violations greater than 5°. Of these the 20 lowest energy structures were selected for presentation and statistical representation.

Circular Dichroism—CD spectra of NPM1 were recorded between 260 and 190 nm at 30–60 μM in the biophysics buffer

using a Jasco-J815 spectropolarimeter (Jasco Inc., Easton, MD) and a 1-mm path length cuvette. The base line due to the buffer only was recorded and subtracted prior to normalizing the data for concentration. Thermal denaturation was followed at 222 nm with heating from 277 to 371 K at a rate of 1 K·min^{-1}. Data were fit to a standard two-state denaturation as described elsewhere (36).

Immunofluorescence—Full-length NPM1 variants were cloned into the mammalian expression vector pEGFP-C1 (BD Biosciences). NPM1 variants were thus expressed *in vivo* with an N-terminal enhanced green fluorescent protein (eGFP) tag.

NIH-3T3 murine fibroblasts were grown in Dulbecco's modified Eagle's medium GlutaMAX™ (Invitrogen) supplemented with 10% fetal calf serum. A day before transfection a confluent flask of cells was subcultured at a 1:8 dilution in 6-well plates containing glass coverslips. Transfection was performed using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions. After 24 h coverslips were removed, washed with PBS, and fixed in 4% paraformaldehyde for 20 min. Cells were permeabilized with PBS + 0.1% Triton X-100 prior to blocking for 30 min with 20% fetal calf serum, 0.5% Tween 20 in PBS (FATS). Cells were stained with mouse anti-fibrillarin (ab4566, AbCam, UK) at 1:500 in FATS for 1 h. After three washes with PBS + 0.1% Triton X-100, the cells were incubated in the dark for 1 h with Alexa Fluor™ 594-conjugated goat anti-mouse antibody (Molecular Probes, Leiden, Netherlands) at 1:200 in FATS. The coverslips were dried in air and mounted using Fluoromount-G™ (Southern Biotech, Birmingham, AL). Images were collected using an MRC-2000 plus confocal apparatus (Bio-Rad) mounted on a Nikon eclipse E800 microscope using the ×60 oil objective lens.

RESULTS

The Structure of the NPM1 C-terminal Domain—Manual inspection of the sequence of NPM1, with assistance from the secondary structure prediction tool Jpred (37), indicated the potential for three short helices separated by short loops within the last 50 residues. This potential domain is separated from the N-terminal oligomerization domain by an ~120-residue acid-rich region that is predicted to be disordered. The DNA sequence corresponding to residues 243–294 of NPM1 was cloned and overexpressed in *E. coli*. Inspection by SDS-PAGE and electrospray-mass spectrometry following purification identified a pure protein of the expected mass. Both one-dimensional ^{1}H NMR and ^{15}N-HSQC experiments produced well dispersed spectra that are characteristic of a folded protein (Fig. 3*A*).

The NMR spectra of residues Ser^{243} to Leu^{294} of NPM1 were assigned, and the solution structure was determined using standard techniques (38, 39). Residues 244–294 adopt a well defined tertiary structure with a root mean square deviation of 0.4 Å for backbone atoms (Fig. 1*A*). Experimental restraints and structural statistics for the 20 accepted lowest energy structures are summarized in Table 1. The coordinates for the structure are available from the Protein Data Bank (entry code 2vxd).

The C-terminal domain of NPM1 forms a well defined right-handed 3-helix bundle (Fig. 1*B*). Helix 1 is formed of residues Val^{244} to Glu^{256}, with Lys^{257} and Gly^{258} then forming an H-bonded turn into the first loop Gly^{259}–Pro^{264}. Helix 2 runs from Glu^{265} to Phe^{276} until the second loop of Arg^{277}–Asp^{280}, and helix 3 is composed of residues Gln^{281} to Lys^{292}. The core of the protein is formed largely by aromatic residues (Tyr^{271}, Phe^{268}, Phe^{276}, Trp^{288}, and Trp^{290}) with further contribution from residues Met^{251}, Ile^{255}, Leu^{261}, Val^{272}, and Leu^{287}. Sequence alignment of homologous proteins (Fig. 2*D*) reveals the hydrophobic core is largely conserved, with Trp^{288}, Trp^{290}, and Phe^{268} and Phe^{276} strictly conserved. Deletions, seen in the more distantly related proteins from *Lytechinus pictus* and *Asterina pectinifera*, are confined solely to loop 1, and it is highly likely that the fold is conserved throughout these species. Other residues strictly conserved are Pro^{262}, which ends loop 1, and three lysines as follows: Lys^{263} and Lys^{267} cluster on the surface of the protein, whereas Lys^{248} appears to form a salt bridge with Asp^{286} (Fig. 2*C*) and may be structurally important.

The NPM1 C-terminal domain has no significant sequence similarity to any protein or domain, but a search for structurally similar proteins using the program DALI (40) returned several matches, among which the most significant single domain protein was the recently published archaeal ribosomal protein S17e (41).

Effect of the Leukemia Mutations on the Structure of the NPM1 C-terminal Domain—NPMc+ mutations alter the sequence of the C terminus of the third helix in the structure. In the most common mutation, mutant A, the wild-type sequence WQWRKSL is replaced by the sequence CLAVEEVSLRK, whereas in the less common mutant E it is replaced with WQS-

TABLE 1

Summary of conformational constraints and statistics for the 20 lowest energy accepted structures of C-terminal domain of nucleophosmin

Structural constraints	
Intra-residue	823
Sequential	196
Medium range ($2 \leq \|i-j\| \leq 4$)	219
Long range ($\|i-j\| > 4$)	118
χ1 angle constraints	15
ALOS constraints	77
Hydrogen bond constraints	50
Total	1498
Statistics for accepted structures	
Statistics parameter (±S.D.)	
r.m.s.d.[a] for distance constraints	0.0084 ± 0.0003 Å
r.m.s.d. for dihedral constraints	0.345 ± 0.029°
Mean CNS energy term (kcal·mol^{-1} ± S.D.)	
E (overall)	50.3 ± 4.2
E (van der Waals)	15.8 ± 2.9
E (NOE and hydrogen bond constraints)	3.1 ± .2
E (χ-1 dihedral and TALOS constraints)	0.56 ± 0.1
r.m.s.d. from the ideal geometry (±S.D.)	
Bond lengths	0.0015 ± 0.00009 Å
Bond angles	0.32 ± 0.007°
Improper angles	0.24 ± 0.013°
Average atomic r.m.s.d. from the mean structure (±S.D.)	
Residues 3–54 (N, C-α, C atoms)	0.40 ± 0.07 Å
Residues 3–54 (all heavy atoms)	0.94 ± 0.08 Å
Structural quality	
Residues in most favored region of Ramachandran plot	94.1%
Residues in additional allowed region of Ramachandran plot	5.2%
Residues in disallowed region of Ramachandran plot	0.3%

[a] r.m.s.d. means root mean square deviation; NOE means nuclear Overhauser effect.

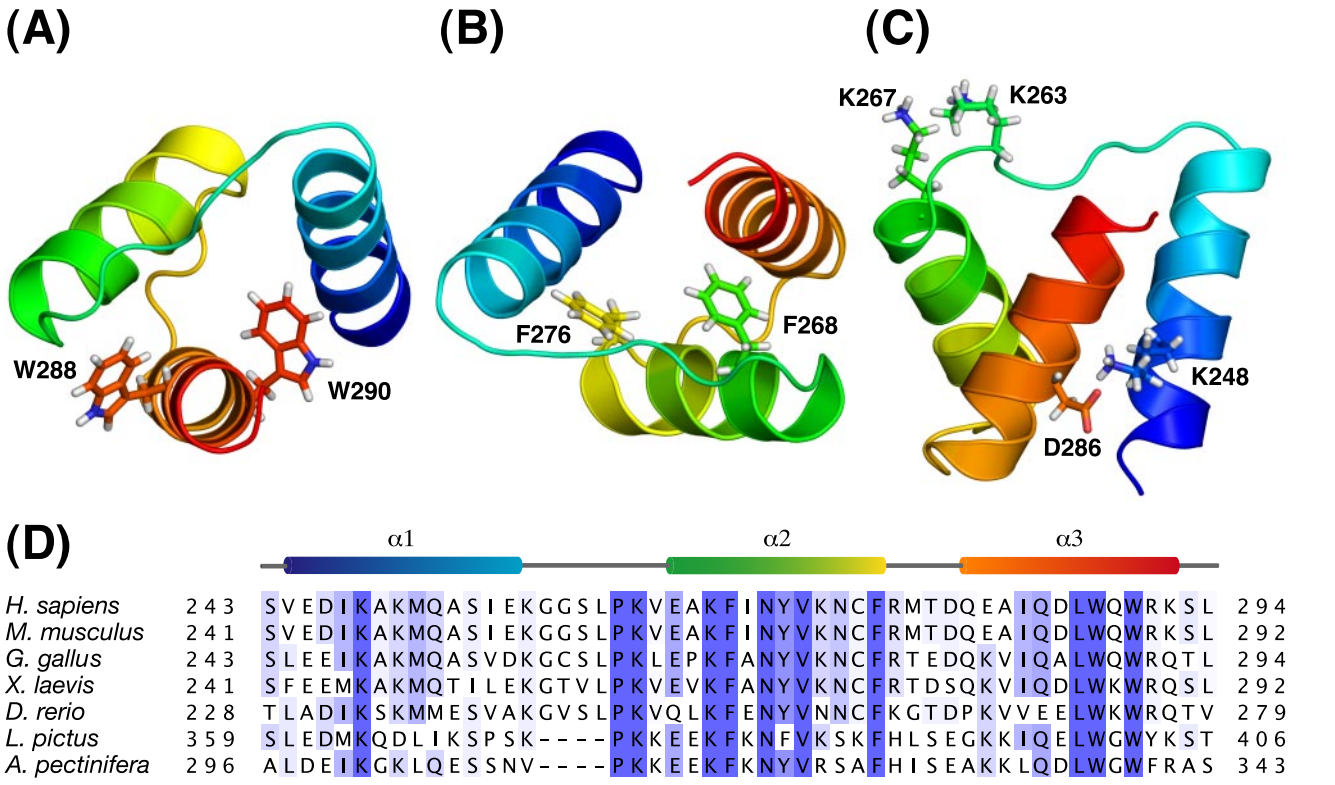

FIGURE 2. **Sequence alignment and the positions of structurally and functionally important residues.** *A–C,* schematic representations of the C-terminal domain of nucleophosmin colored *blue* to *red* from the N to C terminus with structurally and functionally important residues highlighted as *sticks*. Figure was prepared using the program PyMOL (53). *A,* Trp[288] and Trp[290] showing their positions in the core. *B,* Phe[268] and Phe[276], for comparison with *A*. *C,* residues Lys[248] and Asp[286] that form a salt bridge and Lys[263] and Lys[267], which cluster on the surface. *D,* structure-based sequence alignment of the C-terminal domain of nucleophosmin and related proteins. Sequences were aligned using the programs ClustalW (54) and Jalview (55) and shaded *blue* according to the degree of conservation. Proteins and their NCBI accession numbers are as follows: nucleophosmin 1 from *Homo sapiens* gi:10835063, *Mus musculus* gi:56206422, *Danio rerio* gi:31418910, *Gallus gallus* gi:45383996. NO38 from *Xenopus laevis* gi:64919, Mitotic Apparatus Protein p62 from *L. pictus* gi:1841528 and nucleolar protein from *A. pectinifera* gi:18146874.

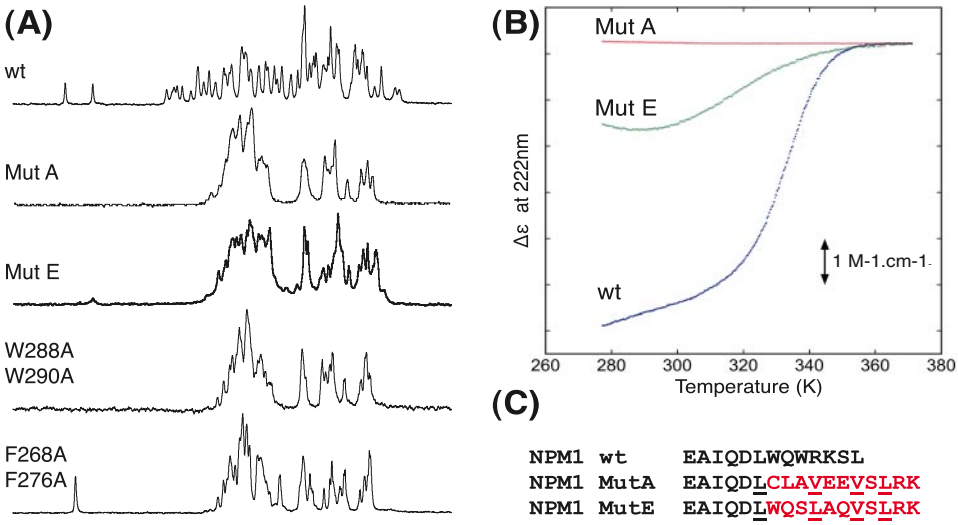

FIGURE 3. **Biophysical analysis of nucleophosmin variants.** *A,* amide region of one-dimensional ^{1}H NMR spectra of the C-terminal domains of the nucleophosmin variants: wild-type (*wt*); mutant A (*Mut A*); mutant E (*Mut E*); W288A/W290A (*DW*), and F268A/F276A. All mutants show a decrease in dispersion compared with wild type, indicating that all mutants are unfolded. *B,* traces from thermal melts of nucleophosmin C-terminal domain variants followed by differential absorbance at 222 nm by circular dichroism spectroscopy. The *blue points* correspond to wild type (*wt*), the *green* to Mutant E (*Mut E*), and the *red* to Mutant A (*Mut A*). Wild-type describes a typical two-state sigmoidal transition. Differential scanning calorimetry was used to corroborate this result and agreed within experimental error to a T_m of 62 °C and $\Delta H_{unfolding}$ of 40 kcal·mol^{-1}. Mutant A shows no sign of any transition, whereas mutant E exhibits some degree of transition although appears to never fully occupy the folded state of wild-type NPM1. *C, table* displaying the C-terminal sequences of wild-type NPM1 compared with mutants A and E. All naturally occurring mutations found in AML, including A and E, are because of a frameshift that results in the substitution of the extreme C terminus (including one or both of the tryptophans Trp[288] and Trp[290]) with a leucine-rich NES. Substituted residues are colored *red* and those forming the NES motif are *underlined*.

LAQVSLRK (Fig. 3*C*). Several of the altered residues are at the interface between helix 3 and helices 2 and 1 and contribute to the hydrophobic core of the domain. In the NPMc+ mutants, these residues are replaced by smaller amino acids, and this would be predicted to greatly destabilize the domain. To test this hypothesis, we used NMR spectroscopy. The one-dimensional ^{1}H NMR spectra of both mutants A and E show a decrease in dispersion compared with NPM1wt (Fig. 3*A*), indicative of a loss of tertiary structure. Fig. 3*B* shows the CD spectra of NPM1wt and mutants A and E as a function of temperature. NPM1wt describes a sigmoidal unfolding transition typical of a folded protein, whereas mutant A shows no evidence of any such transition and yields a trace similar to that expected from an unstructured protein. Mutant E shows an intermediate response; there is evidence of a transition,

but the protein never appears to fully occupy the folded state of the wild type.

These results suggest that the mutations found in AML lead to the destabilization of the C-terminal globular domain of NPM1. Mutant A, which lacks both Trp^{288} and Trp^{290}, has no residual regular secondary structure. Interestingly, mutant E (in which Trp^{288} is retained), although largely destabilized, is clearly not completely unstructured. This mutant may retain some ability to localize NPM1 to the nucleolus, and this could explain why it has a stronger nuclear export signal than mutant A (42), which itself is completely unfolded.

A Folded C-terminal Domain Is Required for Nucleolar Targeting—Expressing NPM1 as an eGFP fusion in mammalian cells allows for the direct observation of its subcellular localization (43). In this manner the mutation of Trp^{288} and Trp^{290} to alanine in NPM1 has been shown to lead to loss of nucleolar localization and the accumulation of the protein in the nucleus (27, 28). To date it has not been clear how these residues contributed to the direction of NPM1 to the nucleolus. We have shown that Trp^{288} and Trp^{290} are part of the hydrophobic core of a folded domain (Fig. 2*A*), and it would be expected that their mutation would lead to a loss of structure. Indeed, the NPM1 C-terminal domain with Trp^{288} and Trp^{290} mutated to alanine lacks any distinct tertiary structure (as judged by one-dimensional ^{1}H NMR; see Fig. 3*A*) and does not undergo a cooperative unfolding transition when heated (Fig. 3*B*). This would suggest that mutation of these residues prevents nucleolar localization simply by preventing the correct folding of the C-terminal domain of NPM1. To test this hypothesis, we decided to make other mutations in the C-terminal domain that would be expected to disrupt its structure while retaining the tryptophans. To this end we chose to mutate phenylalanines 268 and 276 to alanine (F268A and F276A) as these residues are also strictly conserved, large hydrophobic residues that pack at the interfaces of the helices and make many contacts within the core (Fig. 2*B*, and compared with Fig. 2*A*). As expected the NPM1 domain containing the F268A/F276A mutations lacks any distinct tertiary structure (Fig. 3*A*). We next examined the effects of these mutations on subcellular localization by transfecting full-length green fluorescent protein-tagged NPM1 F268A/F276A into NIH-3T3 cells. The mutant eGFP-NPM1 F268A/F276A is nuclear and excluded from the nucleolus (Fig. 4), thus proving that nucleolar localization is a property of the C-terminal domain as a whole.

It has been reported by Falini *et al.* (27) that re-inserting Trp^{288} and Trp^{290} into mutant A leads to complete relocalization to the nucleolus. We investigated the C terminus of this mutant A C288W/A290W construct and found it to be folded, as judged by CD and NMR. Presumably, refolding the C terminus of mutant A allows it carry on its normal function within the nucleolus while making the nuclear export signal unavailable to the export machinery. Hence, to better distinguish the role of the NES from nucleolar targeting, we added the nuclear export signal from mutant A (LAVEEVSLRK) to the end of wild-type C-terminal domain with a GGS linker and found the full-length construct to be nucleolar. This indicates that a loss of structure and the addition of a nuclear export signal are necessary to cause the aberrant cytoplasmic localization observed in NPMc+ AML.

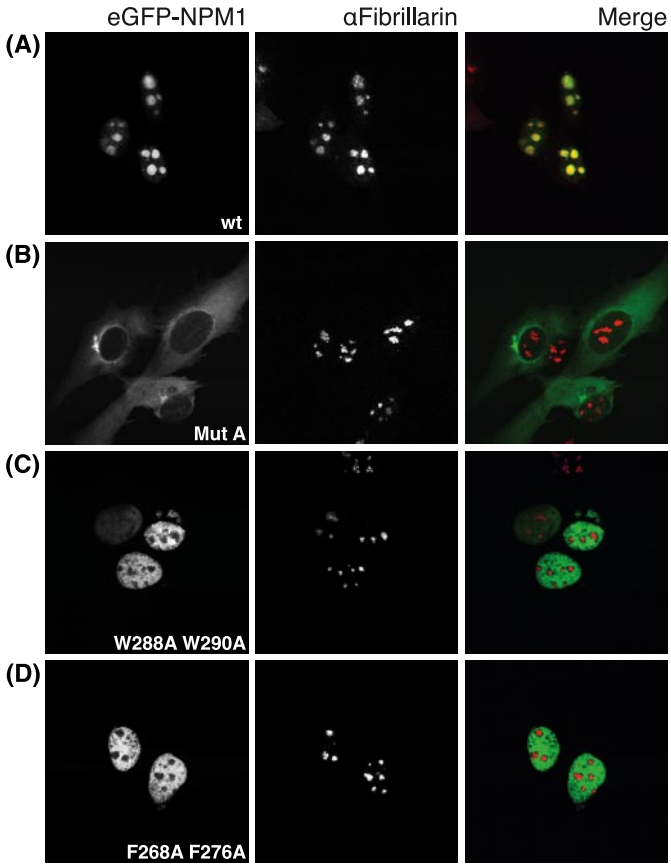

FIGURE 4. **Cellular localization of C-terminal domain variants, structural mutants.** NIH-3T3 cells were transfected with eGFP tagged full-length nucleophosmin variants, as indicated. *A,* wild type (*wt*); *B,* mutant A (*Mut A*); *C,* W288A/W290A; *D,* F268A/F276A (*left panel*). Cells were co-stained with anti-fibrillarin antibody to highlight the nucleolus (*middle panel*). As expected, wild-type NPM1 is entirely nucleolar and mutant A is entirely cytoplasmic. Mutant eGPF-NPM1 W288A/W290A is nuclear and exclusive of the nucleolus, as too is mutant eGFP-NPM1 F268A/F276A.

Functionally Important Residues in the C-terminal Domain of NPM1—Within the C-terminal domain of NPM1 lysine residues 248, 263, and 267 are strictly conserved and surface-exposed (Fig. 5), thus indicating their potential as functionally important residues. To examine this, these residues were mutated to alanine, and the effects on domain stability and nucleolar localization were monitored.

Although Lys^{248} is surface-exposed, it appears to form a long range salt bridge with Asp^{286}. Indeed the mutation K248A significantly destabilizes the domain and as such leads to Lys^{248} being classified as structurally important as opposed to being directly functional (data not shown).

Lysines 263 and 267 cluster and orientate freely on the surface of the domain (Fig. 5, and as judged by a lack of side chain nuclear Overhauser effects), and mutation of either Lys^{263} or Lys^{267}, or both, increases the stability of the domain in line with the previously described effects of removing unpaired surface charges (44). Once established as not structurally detrimental, the nucleolar localization of the full-length NPM1 mutants K263A, K267A, and the double mutant K263A/K267A was studied *in vivo*. Mutant K263A was partially delocalized to the

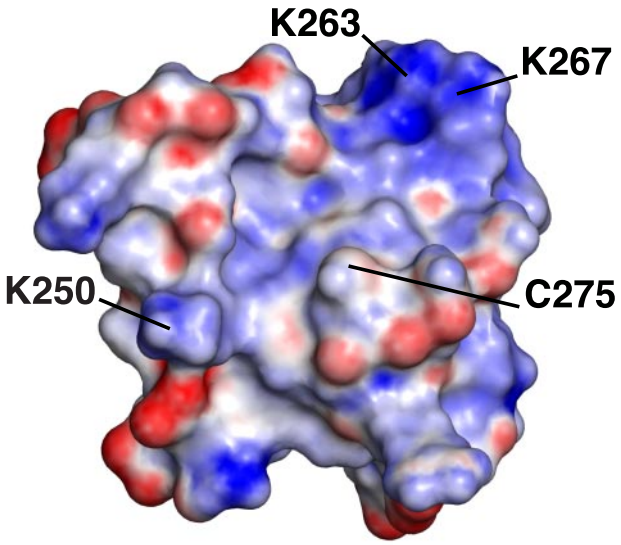

FIGURE 5. **Electrostatic representation of the C-terminal domain showing functionally important residues.** The electrostatic surface potential of the C-terminal domain of NPM1 as calculated by the program APBS (56) and colored using a linear color ramp from −20.0 kiloteslas (*red*) to +20.0 (*blue*). Lys^{263} and Lys^{267} are required for nucleolar localization, but Lys^{250} is not (see Fig. 6). The natural product avrainvillamide bound to Cys^{275} and displaced NPM1 from the nucleolus to the nucleoplasm (45).

nucleoplasm, and K267A and the double mutant were completely delocalized (Fig. 6).

As a control, the freely orienting surface lysine Lys^{250} was also mutated. The C-terminal mutant K250A was, predictably, more stable than the wild-type domain yet had no significant effect on nucleolar localization (Fig. 6). This important result emphasizes the specific functional requirement of residues Lys^{263} and Lys^{267} for nucleolar localization over any nonspecific charge effects.

DISCUSSION

Nucleophosmin is a member of the nucleoplasmin family of molecular chaperones and shares the N-terminal pentamerization domain and central acidic region characteristic of these proteins. In this study we have shown that NPM1 has an additional folded domain at its C terminus, which is unique to this subfamily. Although nucleoplasmin is found throughout the nucleus, NPM1 is one of the most abundant proteins in the nucleolus, and this C-terminal domain is necessary for its nucleolar localization. How it directs the protein to the nucleolus is not clear. It seems likely that it binds to a protein or nucleic acid ligand that is abundant in the nucleolus, and it is possible that the structural similarity to ribosomal protein S17e is functionally significant.

The natural product avrainvillamide has been shown to bind to the C-terminal domain of NPM1 via Cys^{275} (45) and sensitizes cells to apoptosis. It seems likely that avrainvillamide competes with the natural binding partner of this domain. Interestingly, Cys^{275} is on the same face as the two lysine residues that we have shown to be important for nucleolar localization (Fig. 5) suggesting that this region of the domain is involved in ligand binding.

Ribosome biogenesis is metabolically expensive and a rate-limiting step in cell division and growth. Consequently, ribosomal integrity is closely monitored by a variety of pathways (46).

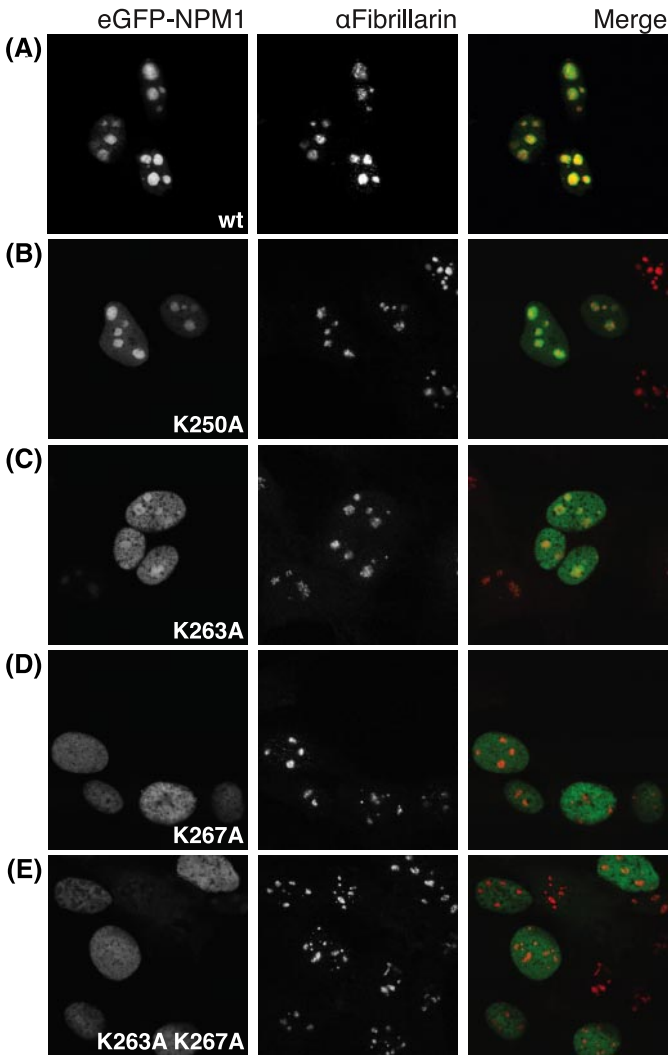

FIGURE 6. **Cellular localization of C-terminal domain mutants, functional mutants.** NIH-3T3 cells were transfected with eGFP tagged full-length nucleophosmin variants, as indicated. *A,* wild type (*wt*); *B,* K250A; *C,* K263A; *D,* K267A; *E,* K263A/K267A (*left panel*). Cells were co-stained with α-fibrillarin antibody to highlight the nucleolus (*middle panel*). Wild-type NPM1 is entirely localized to the nucleolus, and mutant eGPF-NPM1 K250A is broadly similar. Mutant eGPF-NPM1 K263A is largely displaced into the nucleoplasm but displays some residual nucleolar retention. Mutant eGPF-NPM1 267A and the double mutant eGPF-NPM1 K263A/K267A show no residual nucleolar binding and are evenly distributed throughout the nucleoplasm.

Free ribosomal proteins, for example, activate the tumor suppressor p53 by inhibiting its negative regulator MDM2. There is growing evidence that defects in these pathways can lead to disease. Several blood disorders, for example, are caused by mutations in ribosomal proteins or in ribosome assembly factors (47–51), and mutations that prevent ribosomal protein binding to MDM2 have been reported in human cancers (52). These pathways also seem to be disrupted in NPMc+ AML. NPM1 binds to the p19/ARF tumor suppressor and is also regulated by other pathways that control cell growth (18–24). It seems likely that misdirection of $p19^{ARF}$, or another key regulatory molecule that binds NPM1, to the cytoplasm may play a key role in AML. The results of this study show that the aberrant cytoplasmic accumulation of NPM1 in NPMc+ AML results from both the gain of a nuclear export signal and the disruption of the structure of its C-terminal domain. Molecules

that target the C-terminal domain of NPM1 interfere with this mislocalization could therefore have potential applications in leukemia therapy.

REFERENCES

1. Michalik, J., Yeoman, L. C., and Busch, H. (1981) *Life Sci.* **28,** 1371–1379
2. Spector, D. L., Ochs, R. L., and Busch, H. (1984) *Chromosoma* (*Berl.*) **90,** 139–148
3. Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I., and Mann, M. (2005) *Nature* **433,** 77–83
4. Tschochner, H., and Hurt, E. (2003) *Trends Cell Biol.* **13,** 255–263
5. Borer, R. A., Lehner, C. F., Eppenberger, H. M., and Nigg, E. A. (1989) *Cell* **56,** 379–390
6. Yun, J. P., Chew, E. C., Liew, C. T., Chan, J. Y., Jin, M. L., Ding, M. X., Fai, Y. H., Li, H. K., Liang, X. M., and Wu, Q. L. (2003) *J. Cell Biochem.* **90,** 1140–1148
7. Leung, A. K., Gerlich, D., Miller, G., Lyon, C., Lam, Y. W., Lleres, D., Daigle, N., Zomerdijk, J., Ellenberg, J., and Lamond, A. I. (2004) *J. Cell Biol.* **166,** 787–800
8. Hingorani, K., Szebeni, A., and Olson, M. O. (2000) *J. Biol. Chem.* **275,** 24451–24457
9. Okuwaki, M., Tsujimoto, M., and Nagata, K. (2002) *Mol. Biol. Cell* **13,** 2016–2030
10. Szebeni, A., and Olson, M. O. (1999) *Protein Sci.* **8,** 905–912
11. Wang, D., Baumann, A., Szebeni, A., and Olson, M. O. (1994) *J. Biol. Chem.* **269,** 30994–30998
12. Grisendi, S., Bernardi, R., Rossi, M., Cheng, K., Khandker, L., Manova, K., and Pandolfi, P. P. (2005) *Nature* **437,** 147–153
13. Olson, M. O., Wallace, M. O., Herrera, A. H., Marshall-Carlson, L., and Hunt, R. C. (1986) *Biochemistry* **25,** 484–491
14. Herrera, J. E., Savkur, R., and Olson, M. O. (1995) *Nucleic Acids Res.* **23,** 3974–3979
15. Savkur, R. S., and Olson, M. O. (1998) *Nucleic Acids Res.* **26,** 4508–4515
16. Yu, Y., Maggi, L. B., Jr., Brady, S. N., Apicelli, A. J., Dai, M. S., Lu, H., and Weber, J. D. (2006) *Mol. Cell. Biol.* **26,** 3798–3809
17. Grisendi, S., Mecucci, C., Falini, B., and Pandolfi, P. P. (2006) *Nat. Rev. Cancer* **6,** 493–505
18. den Besten, W., Kuo, M. L., Williams, R. T., and Sherr, C. J. (2005) *Cell Cycle* **4,** 1593–1598
19. Brady, S. N., Yu, Y., Maggi, L. B., Jr., and Weber, J. D. (2004) *Mol. Cell Biol.* **24,** 9327–9338
20. Bertwistle, D., Sugimoto, M., and Sherr, C. J. (2004) *Mol. Cell. Biol.* **24,** 985–996
21. Colombo, E., Marine, J. C., Danovi, D., Falini, B., and Pelicci, P. G. (2002) *Nat. Cell Biol.* **4,** 529–533
22. Itahana, K., Bhat, K. P., Jin, A., Itahana, Y., Hawke, D., Kobayashi, R., and Zhang, Y. (2003) *Mol. Cell* **12,** 1151–1164
23. Boon, K., Caron, H. N., van Asperen, R., Valentijn, L., Hermus, M. C., van Sluis, P., Roobeek, I., Weis, I., Voute, P. A., Schwab, M., and Versteeg, R. (2001) *EMBO J.* **20,** 1383–1393
24. Zeller, K. I., Haggerty, T. J., Barrett, J. F., Guo, Q., Wonsey, D. R., and Dang, C. V. (2001) *J. Biol. Chem.* **276,** 48285–48291
25. Falini, B., Mecucci, C., Tiacci, E., Alcalay, M., Rosati, R., Pasqualucci, L., La Starza, R., Diverio, D., Colombo, E., Santucci, A., Bigerna, B., Pacini, R., Pucciarini, A., Liso, A., Vignetti, M., Fazi, P., Meani, N., Pettirossi, V., Saglio, G., Mandelli, F., Lo-Coco, F., Pelicci, P. G., and Martelli, M. F. (2005) *N. Engl. J. Med.* **352,** 254–266
26. Nakagawa, M., Kameoka, Y., and Suzuki, R. (2005) *N. Engl. J. Med.* **352,** 1819–1820
27. Falini, B., Bolli, N., Shan, J., Martelli, M. P., Liso, A., Pucciarini, A., Bigerna, B., Pasqualucci, L., Mannucci, R., Rosati, R., Gorello, P., Diverio, D., Roti, G., Tiacci, E., Cazzaniga, G., Biondi, A., Schnittger, S., Haferlach, T., Hiddemann, W., Martelli, M. F., Gu, W., Mecucci, C., and Nicoletti, I. (2006) *Blood* **107,** 4514–4523
28. Nishimura, Y., Ohkubo, T., Furuichi, Y., and Umekawa, H. (2002) *Biosci. Biotechnol. Biochem.* **66,** 2239–2242
29. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem.* **182,** 319–326
30. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci.* **4,** 2411–2423
31. Neri, D., Szyperski, T., Otting, G., Senn, H., and Wuthrich, K. (1989) *Biochemistry* **28,** 7510–7516
32. Goddard, T. D., and Kneller, D. G. (2007) *SPARKY 3*, University of California, San Francisco
33. Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. (1995) *J. Magn. Reson.* **108,** 94–98
34. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
35. 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. Sect. D Biol. Crystallogr.* **54,** 905–921
36. Schmid, F. X. (1989) *Spectral Methods of Characterizing Protein Conformation and Conformational Changes*, (Creighton, T. E. ed) pp. 251–285, IRL Press at Oxford University Press, Oxford, UK
37. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) *Bioinformatics* (*Oxf.*) **14,** 892–893
38. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acid*, John Wiley & Sons, Inc., New York
39. Bax, A. (1994) *Curr. Opin. Struct. Biol.* **4,** 738–744
40. Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20,** 478–480
41. Wu, B., Yee, A., Huang, Y. J., Ramelot, T. A., Cort, J. R., Semesi, A., Jung, J. W., Lee, W., Montelione, G. T., Kennedy, M. A., and Arrowsmith, C. H. (2008) *Protein Sci.*, **17,** 583–588
42. Bolli, N., Nicoletti, I., De Marco, M. F., Bigerna, B., Pucciarini, A., Mannucci, R., Martelli, M. P., Liso, A., Mecucci, C., Fabbiano, F., Martelli, M. F., Henderson, B. R., and Falini, B. (2007) *Cancer Res.* **67,** 6230–6237
43. Dundr, M., Misteli, T., and Olson, M. O. (2000) *J. Cell Biol.* **150,** 433–446
44. Serrano, L., Horovitz, A., Avron, B., Bycroft, M., and Fersht, A. R. (1990) *Biochemistry* **29,** 9343–9352
45. Wulff, J. E., Siegrist, R., and Myers, A. G. (2007) *J. Am. Chem. Soc.* **129,** 14444–14451
46. Ruggero, D., and Pandolfi, P. P. (2003) *Nat. Rev. Cancer* **3,** 179–192
47. Cmejla, R., Cmejlova, J., Handrkova, H., Petrak, J., and Pospisilova, D. (2007) *Hum. Mutat.* **28,** 1178–1182
48. Draptchinskaia, N., Gustavsson, P., Andersson, B., Pettersson, M., Willig, T. N., Dianzani, I., Ball, S., Tchernia, G., Klar, J., Matsson, H., Tentler, D., Mohandas, N., Carlsson, B., and Dahl, N. (1999) *Nat. Genet.* **21,** 169–175
49. Ebert, B. L., Pretz, J., Bosco, J., Chang, C. Y., Tamayo, P., Galili, N., Raza, A., Root, D. E., Attar, E., Ellis, S. R., and Golub, T. R. (2008) *Nature* **451,** 335–339
50. Gazda, H. T., Grabowska, A., Merida-Long, L. B., Latawiec, E., Schneider, H. E., Lipton, J. M., Vlachos, A., Atsidaftos, E., Ball, S. E., Orfali, K. A., Niewiadomska, E., Da Costa, L., Tchernia, G., Niemeyer, C., Meerpohl, J. J., Stahl, J., Schratt, G., Glader, B., Backer, K., Wong, C., Nathan, D. G., Beggs, A. H., and Sieff, C. A. (2006) *Am. J. Hum. Genet.* **79,** 1110–1118
51. Menne, T. F., Goyenechea, B., Sanchez-Puig, N., Wong, C. C., Tonkin, L. M., Ancliff, P. J., Brost, R. L., Costanzo, M., Boone, C., and Warren, A. J. (2007) *Nat. Genet.* **39,** 486–495
52. Lindstrom, M. S., Jin, A., Deisenroth, C., White Wolf, G., and Zhang, Y. (2007) *Mol. Cell. Biol.* **27,** 1056–1068
53. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System,* DeLano Scientific, San Carlos, CA
54. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) *Nucleic Acids Res.* **31,** 3497–3500
55. Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004) *Bioinformatics* (*Oxf.*) **20,** 426–427
56. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10037–10041