Structural polymorphism in the N-terminal oligomerization domain of NPM1

Diana M. Mitrea[a], Christy R. Grace[a], Marija Buljan[b], Mi-Kyung Yun[a], Nicholas J. Pytel[a], John Satumba[a], Amanda Nourse[a], Cheon-Gil Park[a], M. Madan Babu[b], Stephen W. White[a,c], and Richard W. Kriwacki[a,c,1]

[a]Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN 38105; [b]Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom; and [c]Department of Microbiology, Immunology and Biochemistry, University of Tennessee Health Science Center, Memphis, TN 38163

Edited by Carl Frieden, Washington University School of Medicine, St. Louis, MO, and approved February 19, 2014 (received for review November 11, 2013)

Nucleophosmin (NPM1) is a multifunctional phospho-protein with critical roles in ribosome biogenesis, tumor suppression, and nucleolar stress response. Here we show that the N-terminal oligomerization domain of NPM1 (Npm-N) exhibits structural polymorphism by populating conformational states ranging from a highly ordered, folded pentamer to a highly disordered monomer. The monomer–pentamer equilibrium is modulated by posttranslational modification and protein binding. Phosphorylation drives the equilibrium in favor of monomeric forms, and this effect can be reversed by Npm-N binding to its interaction partners. We have identified a short, arginine-rich linear motif in NPM1 binding partners that mediates Npm-N oligomerization. We propose that the diverse functional repertoire associated with NPM1 is controlled through a regulated unfolding mechanism signaled through posttranslational modifications and intermolecular interactions.

NMR | X-ray crystallography

Nucleophosmin (NPM1) is a highly abundant nucleolar phosphoprotein with functions associated with ribosome biogenesis (1, 2), maintenance of genome stability (1), nucleolar stress response (3), modulation of the p53 tumor suppressor pathway (4), and regulation of apoptosis (5). Importantly, genetic alterations that affect the NPM1 protein sequence or expression level are associated with oncogenesis. For example, NPM1 overexpression was observed in a variety of solid tumors, and mutations within the protein and genetic translocations involving *NPM1* are associated with hematological malignancies (reviewed in ref. 6).

NPM1 primarily resides in the nucleolus which is a membraneless compartment and the site of rRNA synthesis, processing, and assembly with ribosomal proteins (7). In the nucleolus, NPM1 is involved in processing preribosomal RNA (4), chaperoning the nucleolar entry of ribosomal (1, 8) and viral (9) proteins, and stabilizing the alternate reading frame (ARF) tumor suppressor protein (4, 5, 10, 11), while also playing a role in the shuttling of preribosomal particles assembled in the nucleolus to the cytoplasm (12–14).

NPM1 is a member of the nucleoplasmin protein family, which includes the histone chaperones NPM2 and NPM3. These proteins share a conserved N-terminal oligomerization domain that mediates homopentamerization (15). Disruption of NPM1 oligomerization by a small molecule (16) or an RNA aptamer (17) causes exclusive nucleoplasmic localization, loss of colocalization with ARF, and induction of p53-dependent apoptosis (16, 17). These observations suggest that changes in the oligomeric state of NPM1 may influence its biological functions. However, although it is hypothesized (1) that NPM1 function is modulated through control of its oligomeric state, experimental data are currently lacking. Intriguingly, NPM1 exhibits 40 putative phosphorylation sites, the majority of which are evolutionarily conserved (18, 19). Modification of these sites that is influenced by subcellular localization and cell cycle phase (20, 21) modulates the biological function of NPM1 (1, 6). Approximately one third of these phosphorylation sites are located within the N-terminal oligomerization domain, indicating their possible involvement in regulation of the oligomerization state of NPM1.

What is the molecular mechanism that underlies the various functions of NPM1? Here we show by using in vitro biophysical and structural methods that the N-terminal oligomerization domain of NPM1 (Npm-N) exhibits structural polymorphism by populating a range of conformations with various degrees of structural disorder that span two extreme structural states: a folded pentameric state and a disordered monomeric state. Several conserved phosphorylation sites in pentameric Npm-N are positioned within the hydrophobic interior of the pentameric structure (18) and therefore are inaccessible to kinases. Interestingly, other conserved sites are solvent accessible, and we show that these posttranslational modification (PTM) sites serve as molecular switches for modulating the oligomerization and the folding state of Npm-N. We propose that, when exposed through initial destabilizing phosphorylation events, the otherwise inaccessible sites act as molecular locks that, when phosphorylated (22–24), destabilize the pentameric form and lock Npm-N in the monomeric state. We demonstrate that the monomer–pentamer equilibrium is modulated by protein binding partners and have identified a short, arginine-rich (R-rich) linear motif that mediates this interaction. Our results suggest that the diverse cellular functions and subcellular localization of NPM1 are influenced through a regulated unfolding mechanism, signaled through PTMs and intermolecular interactions.

Significance

Nucleophosmin (NPM1) is a multifunctional protein with critical roles in ribosome biogenesis, centrosome duplication, and tumor suppression. Despite the established importance of NPM1 as a tumor marker and potential drug target, little is currently known about the molecular mechanisms that govern its various functions. Our manuscript describes that the N-terminal domain of NPM1 (Npm-N) exhibits phosphorylation-dependent structural polymorphism along a broad conformational landscape between two extreme states: a stable, folded pentamer and a globally disordered monomer. We propose that phosphorylation-induced "regulated unfolding" of Npm-N provides a means to modulate NPM1 function and subcellular localization. Our findings will drive future structure-based studies on the roles of regulated unfolding in NPM1 biology and will provide a foundation for NPM1-targeted anticancer drug development.

Author contributions: D.M.M., M.B., A.N., M.M.B., and R.W.K. designed research; D.M.M., C.R.G., M.B., M.-K.Y., N.J.P., J.S., A.N., and C.-G.P. performed research; D.M.M., C.R.G., M.B., M.-K.Y., A.N., and S.W.W. analyzed data; and D.M.M., M.B., and R.W.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 4N8M).

[1]To whom correspondence should be addressed. E-mail: richard.kriwacki@stjude.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321007111/-/DCSupplemental.

Results

NPM1 contains three functional regions (Fig. 1). The N-terminal structured domain mediates oligomerization and protein chaperone activity and contains two putative nuclear export signals (15). The central region, important for interactions with histones (25), is disordered (*SI Appendix*, Fig. S1*A*) and highly acidic, and contains a bipartite nuclear localization signal (15). Last, the C-terminal domain, which evolved later and is not conserved in other members of the nucleoplasmin family (25), folds into a three-helix bundle (26, 27), contains a nucleolar localization signal, and binds to nucleic acids (28). The NPM1 primary structure is highly conserved among mammals and amphibians (*SI Appendix*, Fig. S1*B*).

Mouse Npm-N Exhibits Two Radically Different Conformations. Toward understanding the role of the N-terminal oligomerization domain in NPM1 function, we solved the structure of mouse Npm-N by using X-ray crystallography (residues 1–130; Protein Data Bank ID code 4N8M; *SI Appendix*, Table S1). The construct studied includes the previously characterized core domain, which incorporates acidic tract A1 (29), and additionally contains the complete N terminus and a disordered C-terminal segment (residues 123–130) containing a CK2 phosphorylation site (Ser125) (20) and acidic tract A2 associated with histone chaperone activity (30). Consistent with previously published structures of the nucleoplasmin family core domains (*SI Appendix*, Table S2), Npm-N crystallized as a pentamer (Fig. 2*A*). Each Npm-N protomer (Fig. 2*B*) is comprised of eight antiparallel β-strands stabilized by extensive intra- and interprotomer hydrophobic and H-bond interactions; >1,000 Å^2 of surface area is buried at each protomer–protomer interface. Interestingly, several highly conserved, putative phosphorylation sites (Ser48, Ser88, Ser112, and Thr75) (22, 24) form an extensive intra- and interprotomer H-bond network that stabilizes the pentamer core (*SI Appendix*, Fig. S2).

The distribution of charged residues (31) within the Npm-N pentamer is highly asymmetric, with negatively charged residues clustered on one face of the oligomer ("top"; adjacent to the N and C termini) and mixed electrostatic features observed on the other (Fig. 2*A*). Herrera et al. (32) previously demonstrated that the oligomeric state of NPM1 was dependent on ionic strength; in the absence of mono- and divalent cations, NPM1 sedimented as a monomer, whereas, in their presence, it formed an oligomer. We crystallized pentameric Npm-N in the presence of 0.2 M NaCl but were unable to crystallize monomeric Npm-N in the absence of salt. Therefore, we used a variety of solution techniques to characterize the monomeric form of Npm-N under low ionic strength conditions. All experiments, unless otherwise noted, were performed in phosphate buffer at pH 7.

CD spectra confirmed that, in solution, Npm-N exhibited β-sheet secondary structure in the presence of 0.15 M NaCl (termed the "high-salt" form; Fig. 2*C*, blue trace). However, in the absence of NaCl (termed the "low-salt" form; Fig. 2*C*, red trace), Npm-N was disordered. Analysis using sedimentation velocity analytical ultracentrifugation (AUC) confirmed that the high-salt form of Npm-N was pentameric (Fig. 2*D*, blue trace, and *SI Appendix*, Table S3) and that the disordered, low-salt

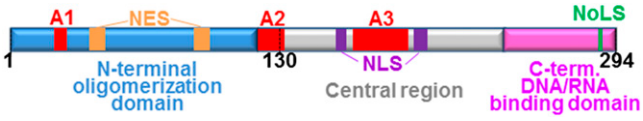

Fig. 1. Domain organization of NPM1. NPM1 is composed of an N-terminal oligomerization domain (blue), containing two nuclear export signals (NES), a central disordered region (gray) that incorporates the bipartite nuclear localization signal (NLS), and a C-terminal DNA/RNA binding domain (magenta) harboring the nucleolar localization signal (NoLS). The three evolutionarily conserved acidic tracts, A1, A2, and A3, are represented in red.

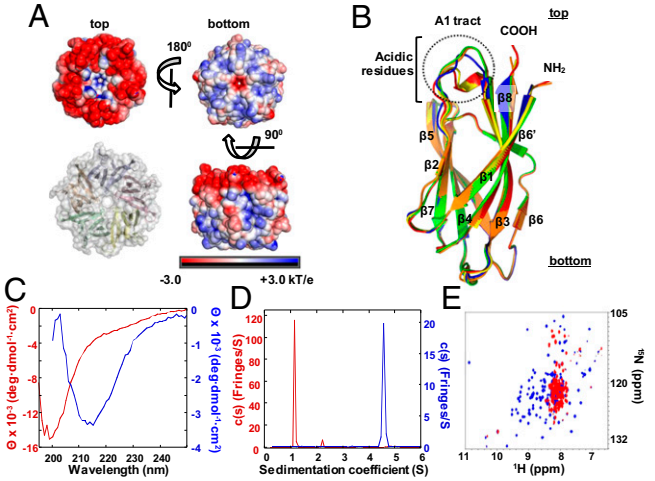

Fig. 2. Npm-N converts from a folded pentamer to a disordered monomer as a function of Na^+ concentration. (*A*) Crystal structure of NPM1 N-terminal domain in pentameric form, solved at 1.8 Å. The electrostatic potential map of Npm-N calculated in PyMOL Molecular Graphics System with Adaptive Poisson-Bolzmann Solver (APBS) shows that the N- and C-termini face (*Upper*) of the pentamer has a large negative net charge, whereas the opposite face (*Lower*) is charge-neutral. (*B*) Superimposition of isolated subunits of the pentamer; these contain nine β-strands connected by flexible linkers (residues 13–120). The N- and C-termini are highly flexible and are not observed in the structure. The acidic loop connecting β2 and β3 (A1 tract) exhibits a high degree of flexibility between the otherwise overlapping structures of the five subunits of the pentamer. (*C*) CD wavelength scans of disordered Npm-N in the absence of NaCl (red) and folded Npm-N in the presence of 0.2 M NaCl (blue). (*D*) Velocity sedimentation AUC analysis shows the folded state to be pentameric and the disordered state to be monomeric. (*E*) Two-dimensional transverse relaxation-optimized spectroscopy (TROSY) spectral overlay of the monomeric (red) and pentameric (blue) folds of $^2H/^{15}N$-labeled Npm-N.

form was an elongated monomer (Fig. 2*D*, red trace, and *SI Appendix*, Table S3). These findings were confirmed by using NMR spectroscopy; 2D [^1H,^{15}N] TROSY (33) spectra showed that the high-salt form of Npm-N was folded (Fig. 2*E*, blue spectrum) in a conformation consistent with that observed in the crystal structure (*SI Appendix*, Fig. S3*A*), and that the low-salt form was disordered (Fig. 2*E*, red spectrum, and *SI Appendix*, Fig. S3*B*).

These observations collectively suggest that the extreme charge asymmetry within the Npm-N pentamer confers the ionic strength dependence of oligomerization. Cations are required to shield the otherwise repulsive electrostatic forces that destabilize the pentamer structure. Physiological ionic strength conditions stabilize the folded pentameric state of Npm-N. Although a regulatory mechanism involving modulation of cellular ionic strength is not physiologically relevant, we investigated whether other factors could modulate the NPM1 disordered monomer–folded pentamer equilibrium.

Phosphorylation Modulates the Oligomerization State of NPM1. PTMs alter the physicochemical properties of proteins and often serve as switches in biological signaling networks (34, 35). We hypothesized that addition of negative charges through phosphorylation of Ser, Thr, and Tyr residues can oppose the stabilizing effect of metal cations on pentameric NPM1 and cause the monomeric form to be populated in cells.

Npm-N contains 16 putative phosphorylation sites (18, 19) (*SI Appendix*, Table S4), evenly distributed between solvent-accessible and solvent-inaccessible locations on the pentameric structure. Intriguingly, large-scale MS studies of cell extracts showed that solvent-exposed and buried sites were phosphorylated

(22, 24), suggesting the monomeric form of the N-terminal domain of NPM1 is populated in living cells.

Our past computational analysis (18) showed that phosphorylation of solvent-exposed and buried sites likely destabilized the pentameric state of Npm-N (*SI Appendix*, Table S4). This suggested a regulatory mechanism wherein phosphorylation of solvent exposed sites leads to destabilization of the Npm-N pentamer and exposure of buried sites for further phosphorylation, effectively shifting the conformational equilibrium from the pentameric to monomeric state. We tested this hypothesis by using a combination of (*i*) phosphosite mimicry (by mutating the phosphorylatable residues to Glu or Asp) and (*ii*) phosphorylation with PKA, which is predicted to phosphorylate Npm-N at five structurally occluded sites (Ser43, Ser48, Ser88, Ser112, and Thr75).

Thr95 and Ser125 were chosen as examples of solvent exposed phosphosites. Thr95 lies at the tip of the loop connecting β6′ and β7, whereas Ser125 is located within the disordered C-terminal acidic tract, A2 (Figs. 2*B* and 3*D*). Phosphomimetic mutations at Thr95 (Npm-N^{T95D}) or Ser125 (Npm-N^{S125E}) subtly affected Npm-N structure and stability. Both mutants were folded under high-salt conditions based on CD (*SI Appendix*, Fig. S5*A*). Sedimentation equilibrium AUC analysis (*SI Appendix*, Table S3) showed that the S125E mutation caused a ∼10-fold increase in the K_d value for formation of pentamers, whereas NaCl titrations monitored by CD revealed that Npm-N^{T95D} required greater electrostatic screening than Npm-N (corresponding to a higher concentration of NaCl; $[NaCl]_{1/2}$) to form the pentamer (*SI Appendix*, Fig. S5*B*). Although destabilizing, these phosphomimetic mutations were not sufficient to allow phosphorylation of the buried PKA putative phosphosites (Fig. 3 *E* and *F*).

Interestingly, the double Npm-N phosphomimetic mutant Npm-$N^{T95D/S125E}$ exhibited a slight decrease in the β-sheet content (*SI Appendix*, Fig. S5*A*) and a shift in $[NaCl]_{1/2}$ from 17 mM to 70 mM (*SI Appendix*, Fig. S5*B*), as measured by CD, in comparison with these features for Npm-N. In addition, this mutant exhibited a >100-fold increase in the K_d value for pentamerization compared with that for Npm-N (*SI Appendix*, Table S3). As a result of weakened protomer–protomer interactions, 10% of Npm-$N^{T95D/S125E}$ sedimented as monomers (*SI Appendix*, Table S3). This mutant was effectively phosphorylated by PKA, albeit to a lesser degree than more disordered mutants (e.g., Npm-N^{S48E} and Npm-$N^{S48E/T95D}$; Fig. 3*E*), and this modification increased the population of monomeric Npm-$N^{T95D/S125E}$ (Fig. 3*F*). These observations suggest that phosphorylation of multiple solvent-accessible sites on Npm-N will cause an increase in the population of monomeric Npm-N and promote exposure of other sites buried within the pentamer structure for subsequent phosphorylation.

Ser48 and Ser88 are examples of solvent inaccessible phosphosites, and are both located at the interface between protomers within the β3 and β6′ regions, respectively (Fig. 3*D*). Particularly, the five isosteric Ser88 residues engage in a network of H-bonds, stabilizing the pentamer core (*SI Appendix*, Fig. S4). These residues are poised to act as structural switches, stabilizing the pentamer in their unmodified state, but becoming sterically and electrostatically unfavorable when phosphorylated.

Consistent with the large unfavorable change in free Gibbs energy (ΔΔG) value calculated for phosphorylation at Ser48 and Ser88 (*SI Appendix*, Table S4), both mutations disrupted the pentameric structure of Npm-N. Interestingly, Npm-N^{S88E} phosphomimic exhibited CD and NMR features similar to those of Npm-N (Fig. 3*B* and *SI Appendix*, Fig. S5*A*), although the 2D TROSY spectrum displayed a larger number of resonances than observed for the WT protein. This latter observation is consistent with a monomer–dimer equilibrium identified by sedimentation equilibrium and sedimentation velocity AUC (*SI Appendix*, Table S3). PKA phosphorylated Npm-N^{S88E} at a maximum of four sites upon overnight incubation; low levels of modification were observed in the short-term radiolabeled assay (Fig. 3*E*), implying a kinetically controlled mechanism for phosphosite exposure.

CD and 2D NMR analysis showed that the S48E mutant (Npm-N^{S48E}) was largely disordered under high-salt, pentamer-promoting conditions (Fig. 3*A* and *SI Appendix*, Fig. S5*A*). This mutant sedimented as an elongated monomer by AUC as signified by a large frictional ratio value (*SI Appendix*, Table S3), and was easily phosphorylated by PKA at as many as four sites (Fig. 3*E*). Addition of T95D phosphomimic mutation increased the phosphorylation efficiency by PKA (Fig. 3*E*). Interestingly, the aggregation that occurred after overnight incubation of these mutants at 37 °C was counteracted in the presence of PKA (Fig. 3*F*).

Taken together, these results support a model wherein sequential phosphorylation of solvent exposed Ser and Thr residues modulates the thermodynamic stability of the NPM1 oligomerization domain to promote exposure of other sites buried within the pentamer structure for subsequent phosphorylation and drive the structural switch from pentamer to disordered monomer.

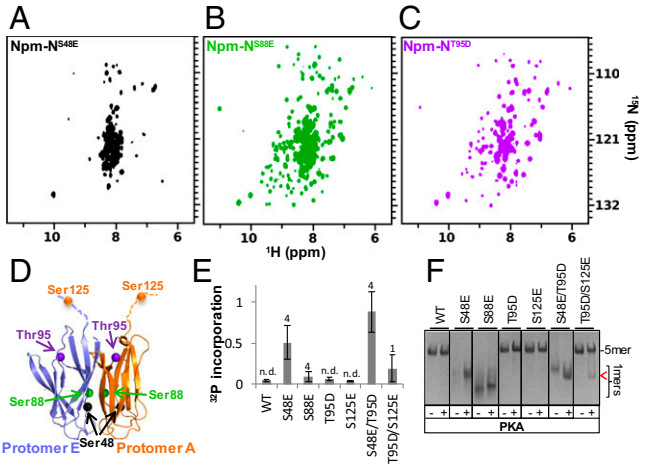

Fig. 3. Phosphorylation modulates Npm-N oligomerization and folding states. [^{1}H,^{15}N] TROSY spectra of ^{15}N-labeled S48E (*A*), S88E (*B*), and T95D (*C*) highlight the disorder–order landscape of Npm-N generated through phosphomimetic mutations. (*D*) Location of the phosphorylation sites used in this study to introduce phosphomimetic mutations are represented on two adjacent subunits of Npm-N pentamer. Ser48 (black) and Ser88 (green) are buried at the hydrophobic interface between two protomers, Thr95 (purple) is solvent-exposed, at the tip of the tight loop connecting β6–β6′, and Ser125 (orange), not observed in the crystal structure, is part of acidic tract A2, in the C-terminal flexible tail. (*E*) The PKA phosphorylation sites are differentially protected from modification, as evidenced by radiolabeled ATP incorporation assays after 30 min incubation at 37 °C. The maximum number of modifications achieved after overnight incubation with PKA for each mutant, as determined by MS, is indicated above the bars (n.d., not detected). The intensity of ^{32}P-radiolabeled gel bands was normalized relative to a histone H1 control, as well as the total predicted PKA phosphorylation sites for each mutant (mean of three independent experiments ± SD). (*F*) Native gel electrophoresis indicates that destabilization of pentameric Npm-N by the double T95D/S125E mutations, mimicking solvent-exposed phosphorylation, allows PKA phosphorylation, which consequently drives dissociation of the pentamer into monomers (red arrow).

ARF Binding Promotes NPM1 Oligomer Assembly. The ARF tumor suppressor interacts with and inhibits the E3 ubiquitin ligase function of Mdm2 (36) and the preribosomal RNA processing function of NPM1 (4), among other identified targets and functions (reviewed in ref. 37). The highly conserved N-terminal domain of ARF binds to the oligomerization domain of NPM1 (4, 38) and also to the intrinsically disordered acidic domain of Mdm2 (39, 40). Similar to the ARF/Mdm2 interaction, which leads to the assembly of large, β-sheet–containing supramolecular structures (39), NPM1 and ARF form large aggregates (4). A fragment of

mouse $p19^{ARF}$ comprised of 37 N-terminal residues (ARF-N37) colocalizes with Mdm2 in mouse fibroblast cells and causes cell cycle arrest, as does full-length $p19^{ARF}$ (41, 42). We used native gel electrophoresis to study the interaction of ARF-N37 with Npm-N and various phosphomimetic Npm-N mutants that exhibit different oligomeric states (Fig. 4*A* and *SI Appendix*, Fig. S6*A*). Under physiological conditions (high-salt buffer), in the absence of ARF-N37, Npm-N^{S88E} migrated faster than Npm-N, consistent with a monomeric form. Upon addition of ARF-N37, additional, more slowly migrating bands were observed, indicating formation of Npm-N^{S88E} oligomers. These Npm-N^{S88E}/ARF-N37 complexes exhibited slower mobility than those formed with Npm-N, suggesting that phosphorylation (of Npm-N, mimicked here by Ser-to-Glu mutagenesis) affects the stoichiometry of the NPM1/ARF complexes. We used size-exclusion chromatography coupled with multiangle light scattering (SEC-MALS) to determine the mass of Npm-N/ARF-N37 complexes (Table 1). Npm-N eluted as a pentamer (73.6 kDa) whereas the predominant species for the Npm-N:ARF-N37 complex exhibited a mass corresponding to 5:2 stoichiometry (83.3 kDa). Npm-N^{S88E} exhibited a mass slightly less than that of a dimer (theoretical mass 29.2 kDa), consistent with the monomer–dimer equilibrium described earlier. Interestingly, in the presence of ARF-N37, Npm-N^{S88E} exhibited multiple heteromeric complexes. The smallest complex exhibited a mass (35.5 kDa) between 2:1 and 2:2 Npm-N^{S88E}:ARF-N37 stoichiometry. A more abundant species, with apparent 5:3 Npm-N^{S88E}:ARF-N37 stoichiometry (88 kDa), demonstrated that ARF-N37 binding promoted assembly of Npm-N oligomers. Notably, visible precipitation was observed upon addition of ARF-N37 at stoichiometries greater than 1:1 (or 5:5) for all Npm-N constructs tested, consistent with the previously reported formation of NPM1–ARF supramolecular complexes (4). Under circumstances in which Npm-N was disordered and monomeric (e.g., Npm-N in low-salt buffer or Npm-N^{S48E} mutant in high-salt buffer), the addition of ARF-N37 caused precipitation even at low concentrations, suggestive of nonspecific, polyvalent interactions.

NPM1 Interacts with Its Nucleolar Partners Through an R-Rich Short Linear Motif. An 8–10-residue-long consensus motif within the ARF N terminus (termed the Arf motif), composed of at least two Arg residues separated by a stretch of hydrophobic residues, mediates binding to and coassembly with the acidic domain of Mdm2 (40). The Arg residues within this short linear motif (SLiM) (35) are required for binding to Mdm2 (39). Arf6, an ARF-derived peptide corresponding to the first 6 aa of mouse

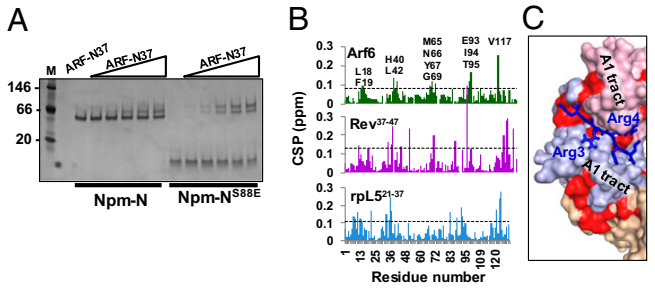

Fig. 4. Peptides containing and R-rich motif bind to Npm-N pentamer and promote assembly of Npm-N monomers. (*A*) Titrations of ARF-N37 (0, 1.0, 2.5, 5.0, 7.5, and 10 μM) in WT and S88E (10 μM) assayed by native gel electrophoresis. (*B*) Histograms of chemical shift perturbations observed in 2D TROSY correlation spectra of 0.4 mM [^{15}N]Npm-N pentamer upon interaction with 3× molar excess of Arf6 (green), Rev^{37-47} (purple), and $rpL5^{21-37}$ (blue) peptides. (*C*) Arg side chains of Arf6 peptide form electrostatic contacts with acidic groups from adjacent protomers within Npm-N pentamer; one molecule of Arf6 (blue) is shown docked between chains A (light blue) and E (light magenta) shown in surface representation, with the residues perturbed by Arf6 interaction highlighted in red.

Table 1. SEC-MALS analysis of 50 μM Npm-N or Npm-N^{S88E} in the presence or absence of 75 μM ARF-N37 shows that Npm-N and Npm-N^{S88E} constructs bind ARF-N37 with different stoichiometry

Sample	Total mass, %	MW, kDa	Npm-N:ARF-37*
WT	100	73.6	5:0 (73.5)
WT plus ARF-N37	94.3	83.3	5:2 (82.9)
S88E	100	27.0	2:0 (29.4)
S88E plus ARF-N37	31.3	88.0	5:3 (87.6)
	65.1	35.3	2:1 (34.1)

*Stoichiometry (theoretical mass in kilodaltons) of proposed oligomeric complex.

$p19^{ARF}$ (Arf6; MGRRFL) and the first Arg cluster, circumvents the aggregation issues experienced with ARF-N37, while retaining the ability to bind to Npm-N (*SI Appendix*, Figs. S6*B* and S7). We performed site-directed mutagenesis to determine (*i*) the contribution of individual residues in Arf6 and (*ii*) the effect of net charge in stabilizing pentameric Npm-N under low-salt conditions. A minimum of two Arg residues were required for stabilization of the folded form of Npm-N (*SI Appendix*, Table S5). Interestingly, an Arf6 peptide in which the two Arg residues were mutated to Lys was unable to stabilize Npm-N in the absence of salt (*SI Appendix*, Table S5), suggesting that binding specificity for Arg residues is an important contributor to the interaction. However, pentamer stabilization via charge screening can be achieved with a polylysine peptide (6K in *SI Appendix*, Table S5).

We next investigated whether NPM1 could be interacting with other protein partners by recognizing the same motif. For this, we obtained a list of the reported NPM1 interaction partners from the BioGrid database (http://thebiogrid.org) and searched those for the R-rich motifs (Dataset S1). We compared the proteins with the highest density of R-rich motifs (i.e., top 50 proteins) to those without a predicted motif (149 proteins) and found that the first group was enriched in proteins annotated as ribonucleoproteins (32% vs. 10%; $P < 0.0005$, χ^2 test), supporting our hypothesis that this motif is involved in mediating interactions of many other ribosomal and nucleolar proteins with NPM1.

The R-Rich SLiM of ARF Binds Within an Acidic Groove of the Mouse NPM1 Pentamer. We further explored the relevance of R-rich motifs in NPM1 interactions by using NMR to study interactions between Npm-N and peptides containing R-rich motifs from three NPM1 binding partners: $p19^{ARF}$ (^{1}MG**RR**FL^{6}, Arf6) (43), HIV-1 Rev (^{37}A**RRN**R**RRRWR**EYC^{47}, Rev^{37-47}) (9), and ribosomal protein L5 (^{21}**RRRR**EGKTDYYA**RKRR**LV^{37}, $rpL5^{21-37}$) (44). We assigned the polypeptide backbone of Npm-N pentamer (available at www.bmrb.wisc.edu/) and identified the interacting residues with the R-rich peptides by analyzing the chemical shift perturbations. All three peptides bound and caused chemical shift perturbations within an acidic groove of pentameric [^{15}N]Npm-N (Fig. 4*B*). Specifically, these interactions mapped to Npm-N residues within the flexible acidic N- and C-termini, acidic tract A1, and the loops connecting strands β4 and β5 and β6 and β6′ (Fig. 2*B* and *SI Appendix*, S8*A*). These data were used as restraints to dock a single Arf6 peptide to the crystal structure of Npm-N by using the program HADDOCK (36). The lowest-energy structure within the cluster with the lowest docking score was selected for structural analysis (*SI Appendix*, *SI Methods* and Table S6). In this low-resolution computational model, the extended backbone of the Arf6 peptide docks in an acidic groove formed between adjacent protomers (Fig. 4*C*). Acidic residues Asp36 and Glu39 within the A1 tract and Glu93 within β6/β6′ participate in H-bonds with the side chains of Arg3 and Arg4 of Arf6 (*SI Appendix*, Fig. S8*A*), providing an explanation for Arf6 peptide-dependent oligomerization of Npm-N under electrostatically unfavorable, low-salt conditions. Residues within the acidic tract A1 exhibit elevated crystallographic

B-factor values (*SI Appendix*, Fig. S8*B*) and variable conformations (Fig. 2*B*), suggesting that the Arf6 binding site is a malleable groove that can accommodate diverse R-rich motifs, in agreement with the large number of identified NPM1 binding partners.

Discussion

NPM1 has proliferative and tumor suppressor functions (1). Although NPM1 shuttles freely between the nucleolus and nucleoplasm (20), accumulation of NPM1 in the nucleolus correlates with protein oligomerization and cell proliferation. Conversely, nucleoplasmic accumulation of NPM1 has been associated with its monomerization (17, 45), early responses to nucleolar stress (3), and induction of apoptosis (17). Upon UV irradiation, NPM1 translocates to the nucleoplasm (46, 47) and dissociates from ARF (46). Furthermore, monomeric mutants exhibit reduced $t_{1/2}$ values [~10 h (45)] compared with the WT protein [>24 h (38, 45, 48)]. These cellular observations suggest a strong correlation between NPM1 oligomerization state, function, and subcellular localization.

Here we report that the multifunctional protein NPM1 exhibits structural polymorphism in its N-terminal oligomerization domain. Repulsive electrostatic interactions between the A1 and A2 tract segments of adjacent protomers in Npm-N must be neutralized by small cations or arginine-containing peptide motifs for the pentameric structure to fold and oligomerize. Regulated unfolding of the pentameric structure is achieved by increasing the repulsive electrostatic field via phosphorylation, phosphomimetic mutations, or changes in solution ionic strength. Distinct phosphorylation signatures allow Npm-N to populate a conformational landscape of order and disorder (Figs. 3 *A*–*C* and 5), which may be responsible for controlling the multiple cellular functions of NPM1 by controlling the exposure of (*i*) binding sites that exist only within the folded pentamer, (*ii*) binding sites or motifs that lie at the protomer–protomer interface, and/or (*iii*) motifs that mediate interactions with partners only in the disordered state. Although these results were obtained by using phosphomimic mutations (net charge, −1; COO^-), similar or exacerbated effects are expected upon phosphorylation of the same residues (net charge, −2; PO_4^{2-}). This prediction is supported by energy calculations comparing the effects of phosphorylation and phosphomimicry (*SI Appendix*, Table S4).

We propose a mechanism wherein phosphorylation of single or multiple solvent exposed sites, such as Thr95 (putative site for PDK, cdc2, and p38MAPK modification; www.cbs.dtu.dk/services/NetPhosK/) (21) or Ser125 (known CK2 site) (20), shift the thermodynamic equilibrium toward the monomeric form, thereby facilitating kinase accessibility to residues that were structurally occluded in the pentameric fold, such as Ser48 [putative site for PKA (www.cbs.dtu.dk/services/NetPhosK/), and ATM and Aurora (19)] (24) and Ser88 (known Nek2 kinase site) (22, 49). Modification of these buried phosphosites structurally locks the protein in one of the monomeric conformations until phosphatases reverse the process or until a binding partner, such as ARF, mediates monomer reassembly into oligomers (Fig. 5). This mechanism is supported by our in vitro results with PKA, in which highly destabilized mutants are much more readily phosphorylated, including at sites buried within the pentamer interior, than WT Npm-N. Importantly, the dual mutant, Npm-$N^{T95D/S125E}$, which mimics phosphorylation at two sites that are solvent accessible in the pentameric state but remains predominantly pentameric, is partially shifted to the monomeric state although phosphorylation by PKA, further supporting our regulated unfolding model.

The nucleolus functions as a sensor for cellular stresses, including UV damage, hypoxia, and chemotoxicity associated with certain chemotherapy agents (3, 46, 50). The structure and composition of the nucleolus is highly dynamic as a result of a constant flux of its protein and nucleic acid constituents into and out of the nucleoplasm (2). The partitioning of proteins between the nucleolus and nucleoplasm is controlled by PTMs that modulate their affinity for nucleolar constituents (7, 20). NPM1 interacts specifically with proteins containing an R-rich SLiM [e.g., ARF (4), ribosomal proteins (8, 12, 44), and HIV-1 protein Rev (9)] and sequesters them in the nucleolus. The R-rich SLiMs bind the pentameric form of Npm-N in a pocket formed at the interface between protomers (Fig. 4*C*). Because the obligatory disordered monomer S48E mutant was unable to bind the Arf6 peptide (*SI Appendix*, Fig. S6), we conclude that particular phosphorylation signatures may trigger unloading of these nucleolar cargoes by shifting the pentamer to monomer equilibrium. As observed with the S88E mutant (Table 1), however, different phosphorylation signatures may signal changes in stoichiometry or even enhance the affinity (49) for particular binding partners. We conclude that a finely tuned regulated unfolding mechanism controls NPM1's subcellular localization, turnover rates, and function. The NPM1 oligomerization domain has generally been understood as having a highly ordered, pentameric structure. Now we show that this domain exhibits otherwise cryptic disordered states that are likely to contribute to NPM1 function. We suggest that cryptic disorder lies within other oligomeric proteins as a means to diversify their structural states and functions.

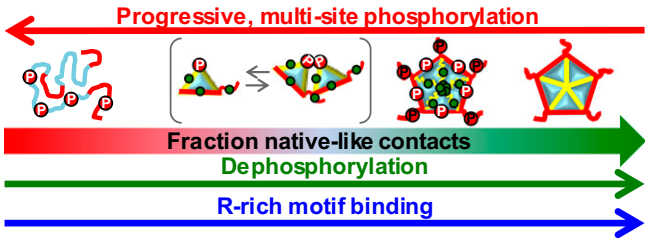

Fig. 5. Navigating the Npm-N conformational landscape. Structural polymorphism of Npm-N is facilitated by repulsive electrostatic forces as a result of close spatial proximity of acidic tracts A1 and A2 (red lines) within the pentameric structure. Documented and putative solvent-exposed and buried phosphorylation sites are illustrated (green dots). Unphosphorylated Npm-N predominantly populates the pentameric state (*Right*). Phosphorylation (red circles) of pentameric Npm-N at solvent exposed sites slightly destabilizes the oligomeric structure (indicated by wavy yellow lines), reducing the energy barrier to make the conformational transition to other, less native-like structures, thus exposing additional, otherwise buried sites for subsequent phosphorylation. Phosphorylation at buried sites dramatically destabilizes the oligomeric structure, stabilizing monomeric folded or monomeric disordered structures (*Left*). Both destabilization scenarios will expose additional interior sites for multisite phosphorylation to structurally lock Npm-N in the monomeric, disordered state. This process can be reversed by dephosphorylation by phosphatases. In addition, binding to target proteins containing R-rich motifs stabilizes the pentameric form, counteracting the destabilization associated with phosphorylation.

Methods

Cloning, Protein Expression, and Purification. All Npm-N constructs were subcloned in frame with an N-terminal polyhistidine tag in pET28a(+) (Novagen), containing a thrombin cleavage site. Npm-N constructs were expressed in *Escherichia coli*, strain BL21(DE3), and purified via Ni-NTA affinity column, followed by proteolytic removal of the polyhistidine tag and size-exclusion chromatography. All concentrations reported in this manuscript are calculated relative to monomeric Npm-N. ARF-N37 expression and purification was previously described (42).

CD. CD measurements were performed using an Aviv 202 spectropolarimeter by using a 1-mm cuvette at 15 μM protein concentration at 25 °C.

AUC. Experiments were carried out in a ProteomeLab XL-I analytical ultracentrifuge with an eight-hole rotor (An-50Ti; Beckman) and cells containing sapphire or quartz windows and charcoal-filled Epon double-sector center pieces (Beckman Coulter). Sedimentation velocity experiments were performed at 20 °C. Sedimentation equilibrium was attained at a rotor temperature of 4 °C at increasing speeds of 10,000 rpm (for 63 h) and 20,000 rpm (for 38 h).

SEC-MALS. The experiment was carried out by using a Shodex PROTEIN KW-803 (exclusion limit, 170,000 Da) size-exclusion column [Showa Denko; three detectors connected in series: an Agilent 1200 UV detector (Agilent Technologies) and a Wyatt DAWN-HELEOS multiangle light-scattering and a Wyatt Optilab rEX differential refractive index detector (Wyatt Technologies)] at 25 °C.

Crystallization. Crystals were grown by the hanging-drop vapor diffusion method at 18 °C by using the following well solution: 0.6 M 1,6-hexanediol, 10 mM cobalt chloride, and 0.1 M sodium acetate (pH 4.6). Well solution (1 μL) was added to 1 μL protein solution (20 mg/mL in 25 mM Tris, pH 8.0, 100 mM NaCl, 1 mM DTT) containing 4.5 mM of the ARF peptide (GRRFLVTVR).

NMR Spectroscopy. All NMR spectra were collected at 25 °C on a Bruker Avance 600 MHz or Bruker Avance 800 MHz spectrometer equipped with $^{1}H/^{13}C$ detect, "TCI" triple-resonance cryogenic probes. NMR data were processed by using Topspin and analyzed using CARA (52) software.

Native Gel Electrophoresis. Stock solutions of 50 μM of all proteins were prepared in reaction buffer. Binding reactions were prepared in the same buffer, with 10 μM Npm-N variant and the indicated concentration of ARF-N37, and were incubated at room temperature for 1 h. Reaction samples were then mixed 1:1 with native PAGE loading buffer and loaded on 8–16% TGX gels (BioRad) and run at 150 V for 65 min at room temperature using prechilled running buffer, and stained with Coomassie Blue.

Phosphorylation Assays. Reactions containing 10 μM Npm-N and 3 μM PKA were incubated at 37 °C for 30 min. The details of the assay are described elsewhere (53). For native PAGE and intact mass determination, the same reactions were prepared by using unlabeled ATP and were incubated at 37 °C overnight.

ACKNOWLEDGMENTS. The authors thank E. Watson and J. Bates for technical assistance, R. Cassell and Dr. P. Rodriguez for peptide synthesis, Dr. S. Taylor for the PKA plasmid, and members of the R.W.K. and Mittag groups for stimulating discussions. Use of the Southeast Regional Collaborative Access Team Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. This work was funded by National Institutes of Health Grants 5RO1-CAO82491 (to R.W.K.), 5R01-GM083159 (to R.W.K.), and P30CA21765 (to St. Jude Children's Research Hospital); MRC Grant MC_U105185859; HFSP Grant RGY0073/2010 (to M.M.B. and M.B.); and ALSAC.

1. Lindström MS (2011) NPM1/B23: A multifunctional chaperone in ribosome biogenesis and chromatin remodeling. *Biochem Res Int* 2011:195209.
2. Chen D, Huang S (2001) Nucleolar components involved in ribosome biogenesis cycle between the nucleolus and nucleoplasm in interphase cells. *J Cell Biol* 153(1):169–176.
3. Avitabile D, et al. (2011) Nucleolar stress is an early response to myocardial damage involving nucleolar proteins nucleostemin and nucleophosmin. *Proc Natl Acad Sci USA* 108(15):6145–6150.
4. Bertwistle D, Sugimoto M, Sherr CJ (2004) Physical and functional interactions of the Arf tumor suppressor protein with nucleophosmin/B23. *Mol Cell Biol* 24(3):985–996.
5. Li Z, Hann SR (2009) The Myc-nucleophosmin-ARF network: A complex web unveiled. *Cell Cycle* 8(17):2703–2707.
6. Colombo E, Alcalay M, Pelicci PG (2011) Nucleophosmin and its complex network: A possible therapeutic target in hematological diseases. *Oncogene* 30(23):2595–2609.
7. Olson MO, Dundr M (2005) The moving parts of the nucleolus. *Histochem Cell Biol* 123(3):203–216.
8. Lindström MS (2012) Elucidation of motifs in ribosomal protein S9 that mediate its nucleolar localization and binding to NPM1/nucleophosmin. *PLoS ONE* 7(12):e52476.
9. Szebeni A, et al. (1997) Nucleolar protein B23 stimulates nuclear import of the HIV-1 Rev protein and NLS-conjugated albumin. *Biochemistry* 36(13):3941–3949.
10. Bolli N, et al. (2009) A dose-dependent tug of war involving the NPM1 leukaemic mutant, nucleophosmin, and ARF. *Leukemia* 23(3):501–509.
11. Saporita AJ, et al. (2011) RNA helicase DDX5 is a p53-independent target of ARF that participates in ribosome biogenesis. *Cancer Res* 71(21):6708–6717.
12. Yu Y, et al. (2006) Nucleophosmin is essential for ribosomal protein L5 nuclear export. *Mol Cell Biol* 26(10):3798–3809.
13. Maggi LB, Jr., et al. (2008) Nucleophosmin serves as a rate-limiting nuclear export chaperone for the Mammalian ribosome. *Mol Cell Biol* 28(23):7050–7065.
14. Rees-Unwin KS, et al. (2010) Ribosome-associated nucleophosmin 1: Increased expression and shuttling activity distinguishes prognostic subtypes in chronic lymphocytic leukaemia. *Br J Haematol* 148(4):534–543.
15. Hingorani K, Szebeni A, Olson MO (2000) Mapping the functional domains of nucleolar protein B23. *J Biol Chem* 275(32):24451–24457.
16. Qi W, et al. (2008) NSC348884, a nucleophosmin inhibitor disrupts oligomer formation and induces apoptosis in human cancer cells. *Oncogene* 27(30):4210–4220.
17. Jian Y, et al. (2009) RNA aptamers interfering with nucleophosmin oligomerization induce apoptosis of cancer cells. *Oncogene* 28(47):4201–4211.
18. Mitrea DM, Kriwacki RW (2012) Cryptic disorder: an order-disorder transformation regulates the function of nucleophosmin. *Pac Symp Biocomput* 152–163.
19. Ramos-Echazábal G, Chinea G, García-Fernández R, Pons T (2012) In silico studies of potential phosphoresidues in the human nucleophosmin/B23: Its kinases and related biological processes. *J Cell Biochem* 113(7):2364–2374.
20. Negi SS, Olson MO (2006) Effects of interphase and mitotic phosphorylation on the mobility and location of nucleolar protein B23. *J Cell Sci* 119(Pt 17):3676–3685.
21. Wang W, Budhu A, Forgues M, Wang XW (2005) Temporal and spatial control of nucleophosmin by the Ran-Crm1 complex in centrosome duplication. *Nat Cell Biol* 7(8):823–830.
22. Cantin GT, et al. (2008) Combining protein-based IMAC, peptide-based IMAC, and MudPIT for efficient phosphoproteomic analysis. *J Proteome Res* 7(3):1346–1351.
23. Oppermann FS, et al. (2009) Large-scale proteomics analysis of the human kinome. *Mol Cell Proteomics* 8(7):1751–1764.
24. Olsen JV, et al. (2010) Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal* 3(104):ra3.
25. Eirín-López JM, Frehlick LJ, Ausió J (2006) Long-term evolution and functional diversification in the members of the nucleophosmin/nucleoplasmin family of nuclear chaperones. *Genetics* 173(4):1835–1850.
26. Gallo A, et al. (2012) Structure of nucleophosmin DNA-binding domain and analysis of its complex with a G-quadruplex sequence from the c-MYC promoter. *J Biol Chem* 287(32):26539–26548.
27. Federici L, et al. (2010) Nucleophosmin C-terminal leukemia-associated domain interacts with G-rich quadruplex forming DNA. *J Biol Chem* 285(48):37138–37149.
28. Wang D, Baumann A, Szebeni A, Olson MO (1994) The nucleic acid binding activity of nucleolar protein B23.1 resides in its carboxyl-terminal end. *J Biol Chem* 269(49): 30994–30998.
29. Lee HH, et al. (2007) Crystal structure of human nucleophosmin-core reveals plasticity of the pentamer-pentamer interface. *Proteins* 69(3):672–678.
30. Akey CW, Luger K (2003) Histone chaperones and nucleosome assembly. *Curr Opin Struct Biol* 13(1):6–14.
31. Unni S, et al. (2011) Web servers and services for electrostatics calculations with APBS and PDB2PQR. *J Comput Chem* 32(7):1488–1491.
32. Herrera JE, Correia JJ, Jones AE, Olson MO (1996) Sedimentation analyses of the salt- and divalent metal ion-induced oligomerization of nucleolar protein B23. *Biochemistry* 35(8):2668–2673.
33. Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94(23):12366–12371.
34. Nishi H, Hashimoto K, Panchenko AR (2011) Phosphorylation in protein-protein binding: Effect on stability and function. *Structure* 19(12):1807–1815.
35. Van Roey K, Gibson TJ, Davey NE (2012) Motif switches: Decision-making in cell regulation. *Curr Opin Struct Biol* 22(3):378–385.
36. de Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. *Nature protocols* 5(5):883–897.
37. Sherr CJ (2006) Divorcing ARF and p53: An unsettled case. *Nat Rev Cancer* 6(9):663–673.
38. Itahana K, et al. (2003) Tumor suppressor ARF degrades B23, a nucleolar protein involved in ribosome biogenesis and cell proliferation. *Mol Cell* 12(5):1151–1164.
39. Sivakolundu SG, et al. (2008) Intrinsically unstructured domains of Arf and Hdm2 form bimolecular oligomeric structures in vitro and in vivo. *J Mol Biol* 384(1):240–254.
40. Bothner B, et al. (2001) Defining the molecular basis of Arf and Hdm2 interactions. *J Mol Biol* 314(2):263–277.
41. Weber JD, et al. (2000) Cooperative signals governing ARF-mdm2 interaction and nucleolar localization of the complex. *Mol Cell Biol* 20(7):2517–2528.
42. DiGiammarino EL, Filippov I, Weber JD, Bothner B, Kriwacki RW (2001) Solution structure of the p53 regulatory domain of the p19Arf tumor suppressor protein. *Biochemistry* 40(8):2379–2386.
43. Kuo ML, den Besten W, Bertwistle D, Roussel MF, Sherr CJ (2004) N-terminal polyubiquitination and degradation of the Arf tumor suppressor. *Genes Dev* 18(15): 1862–1874.
44. Rosorius O, et al. (2000) Human ribosomal protein L5 contains defined nuclear localization and export signals. *J Biol Chem* 275(16):12061–12068.
45. Enomoto T, Lindström MS, Jin A, Ke H, Zhang Y (2006) Essential role of the B23/NPM core domain in regulating ARF binding and B23 stability. *J Biol Chem* 281(27): 18463–18472.
46. Lee C, Smith BA, Bandyopadhyay K, Gjerset RA (2005) DNA damage disrupts the p14ARF-B23(nucleophosmin) interaction and triggers a transient subnuclear redistribution of p14ARF. *Cancer Res* 65(21):9834–9842.
47. Kurki S, et al. (2004) Nucleolar protein NPM interacts with HDM2 and protects tumor suppressor protein p53 from HDM2-mediated degradation. *Cancer Cell* 5(5):465–475.
48. Boisvert FM, et al. (2012) A quantitative spatial proteomics analysis of proteome turnover in human cells. *Mol Cell Proteomics* 11(3):011429.
49. Velimezi G, et al. (2013) Functional interplay between the DNA-damage-response kinase ATM and ARF tumour suppressor protein in human cancer. *Nat Cell Biol* 15(8):967–977.
50. Kurki S, Peltonen K, Laiho M (2004) Nucleophosmin, HDM2 and p53: Players in UV damage incited nucleolar stress response. *Cell Cycle* 3(8):976–979 Epub2004Aug.
51. Zhang HG, Wang J, Yang X, Hsu HC, Mountz JD (2004) Regulation of apoptosis proteins in cancer cells by ubiquitin. *Oncogene* 23(11):2009–2015.
52. Keller R (2004) *The Computer Aided Resonance Assignment Tutorial* (CANTINA Verlag, Goldau, Switzerland).
53. Wang Y, et al. (2011) Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21. *Nat Chem Biol* 7(4):214–221.