

# Deciphering the folding transition state structure and denatured state properties of Nucleophosmin C-terminal domain

Flavio Scaloni<sup>a</sup>, Luca Federici<sup>b</sup>, Maurizio Brunori<sup>a,1</sup>, and Stefano Gianni<sup>a</sup>

<sup>a</sup>Istituto Pasteur-Fondazione Cenci Bolognetti and Istituto di Biologia e Patologia Molecolari del Consiglio Nazionale delle Ricerche, Dipartimento di Scienze Biochimiche A. Rossi Fanelli, Università di Roma La Sapienza, Rome, Italy; and <sup>b</sup>Centro Studi sull'Invecchiamento and Dipartimento di Scienze Biomediche, Università di Chieti G. D'Annunzio, Chieti, Italy

Edited by William A. Eaton, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, and approved February 17, 2010 (received for review September 15, 2009)

**Nucleophosmin (NPM1), one of the most abundant nucleolar proteins, is a frequent target of oncogenic mutations in acute myeloid leukaemia (AML). Mutation-induced changes at the C-terminal domain of NPM1 (Cter-NPM1) compromise its stability and cause the aberrant translocation of NPM1 to the cytosol. Hence, this protein represents a suitable candidate to investigate the relations between folding and disease. Since Cter-NPM1 folds via a compact denatured state, stabilization of the folded state of the mutated variants demands detailed structural information on both the native and denatured states. Here, we present the characterization of the complete folding pathway of Cter-NPM1 and provide molecular details for both the transition and the denatured states. The structure of the transition state was assessed by  $\Phi$ -value analysis, whereas residual structure in the denatured state was mapped by evaluating the effect of mutations as modulated by conditions promoting denatured state compaction. Data reveal that folding of Cter-NPM1 proceeds via an extended nucleus and that the denatured state retains significant malleable structure at the interface between the second and third helices. Our observations constitute the essential prerequisite for structure-based drug-design studies, aimed at identifying molecules that may rescue pathological NPM1 mutants by stabilizing the native-like state.**

kinetics | mutagenesis | protein folding

**N**ucleophosmin (NPM1) is a ubiquitously expressed protein that belongs to the nucleoplasmin family of nuclear chaperones and is one of the most represented nucleolar proteins (1). NPM1 is a key component of several cellular processes, ranging from ribosome biogenesis, control of centrosome duplication, maintenance of genomic stability, and cell-cycle regulation also by its direct interaction with the tumor suppressors p53 and p14arf (2). A distinctive feature of NPM1 resides in its property to rapidly shuttle from the nucleus to the cytoplasm and backwards (3), a property that may require some structural malleability.

The *NPM1* gene has been found overexpressed in several solid human cancers and is also a frequent target of translocations occurring in hematopoietic tumors. Mutation of the *NPM1* gene is the most frequent genetic lesion in acute myeloid leukemia (AML) displaying a normal karyotype (4). Such mutations map at the C-terminal domain of the protein, and all result in its denaturation and appearance of an additional nuclear export signal. These two signatures concur and are both necessary to cause the stable and aberrant cytoplasmic localization of NPM1 mutants (1). Furthermore, NPM1 mutants enhance the translocation in the cytoplasm of several interacting partners, including the tumor suppressor p14arf and the c-Myc E3-ubiquitin ligase Fbw7- $\gamma$  (5, 6). It has been proposed that this feature is involved in the mechanism for the oncogenic properties of NPM mutants (7).

Recent work (8) has shown that some pathological mutations affect the stability of proteins, yielding partial or total denatura-

tion of the mutants even under physiological conditions. Within this perspective, the study of the folding properties of proteins whose stability is altered in specific pathologies is gaining momentum, with the possible goal of restoring the functional native structure of destabilized mutant proteins by appropriate pharmacological chaperones (9, 10). In fact, because folding is an equilibrium between native and denatured states, molecules binding preferentially to the native state can in theory restore the activity of unfolded mutants, if they do not disrupt the activity by binding to or near active sites or hinder critical recognition surfaces.

Recently, we showed that the C-terminal domain of NPM1 (Cter-NPM1) folds via a two-state mechanism, but its denatured state retains significant residual structure (11). In addition, we observed that a compaction of the denatured state results in the acceleration of the folding process, implying that the residual structure should contain mainly native-like interactions; therefore, a structural characterization of the denatured state is a prerequisite to design specific chaperone-active ligands. Because of its elusive nature, unveiling the structure of the denatured state under conditions favoring folding is a challenging task. In fact, it requires reduction of the population of the native state without the addition of denaturants. The experimental approaches used so far include reduction of intramolecular disulfide bridges (12), removal of cofactors (13), truncation of a few residues at the termini of the protein (14, 15), or conservative site-directed mutagenesis (16).

In this work we present the structural dynamics of the folding of Cter-NPM1 using kinetics and protein engineering. We obtained structural information on both the transition state for folding and the denatured state. In particular, while the structure of the transition state was assessed using standard  $\Phi$ -value analysis (17), the residual structure in the denatured state was inferred by comparing the mutation-induced changes in stability as obtained in the absence and in the presence of a stabilizing salt. In fact, as detailed below, quantitative comparison of the datasets obtained under the two experimental conditions allows to estimate for each of the tested positions, a  $\Delta\Delta\Delta G_{D-N}$  value that reflects the contribution of a specific residue to the structure of the denatured state. Data reveal that folding of Cter-NPM1 proceeds via an extended nucleus. Moreover, the denatured state of Cter-NPM1 appears to retain significant residual structure at the interface between helices H2 and H3, whereas H1 is largely unfolded. Interestingly, we demonstrate a linear correlation between  $m_{D-N}$

Author contributions: F.S., L.F., M.B., and S.G. designed research; F.S. and S.G. performed research; F.S., L.F., M.B., and S.G. analyzed data; and F.S., L.F., M.B., and S.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed: E-mail: maurizio.brunori@uniroma1.it.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0910516107/DCSupplemental](http://www.pnas.org/cgi/content/full/0910516107/DCSupplemental).

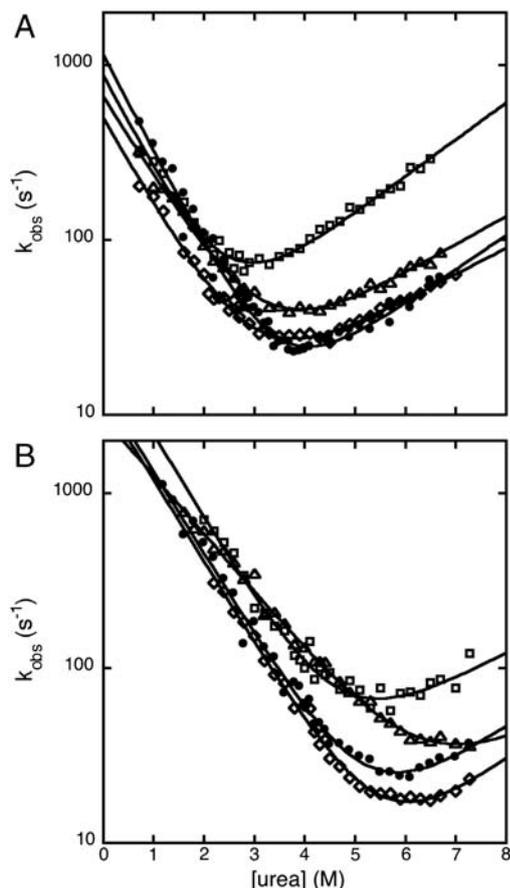
values and  $\Delta\Delta\Delta G_{D-N}$  values, which suggests that the structure of the denatured state is highly malleable and perturbed by mutagenesis.

Our results provide a unique, detailed structural description of the folding transition state and the denatured state of Cter-NPM1. This information paves the way for future studies focused on the design of small molecules aimed at rescuing the activity of naturally occurring pathological NPM1 mutants.

## Results

To characterize quantitatively the folding pathway of Cter-NPM1 we carried out extensive  $\Phi$  value analysis (17). In addition, to obtain structural information on the denatured state we extended the analysis of all mutants also in the presence of 0.5 M sodium chloride, which promotes denatured state compaction. Thirty-two conservative mutants insisting on 25 positions were synthesized (Table 1 and Tables S1 and S2). Five of these (K248A, A248G, M251A, F268A, and F276A) expressed too poorly to be included in the analysis; the remaining 27 were subjected to both equilibrium and kinetic folding experiments. Overall  $\Phi$  values could be calculated for 19 positions (Table 1 and Tables S1 and S2).

Unfolding and refolding kinetics were satisfactorily fitted to single-exponential time courses, under all conditions. Semilogarithmic plots of the observed unfolding and refolding rate constants versus denaturant concentration (chevron plots) for wild-type Cter-NPM1 and selected mutants are shown in Fig. 1A. Fitted parameters are reported in Table 1 and Fig. S1.



**Fig. 1.** Chevron plots of wt and representative mutants of Cter-NPM1. A and B refer to data recorded in the absence and in the presence of 0.5 M NaCl, respectively. Data refer to wt Cter-NPM1 (Closed Circles), V244A (Squares), A266G (Diamonds), and I255V (Triangles). Lines are the best fit to a two-state folding model (51).

Because of its small size (about 6500 Da) (18), Cter-NPM1 displays a low  $m_{D-N}$  value ( $m_{D-N} = 0.9 \pm 0.1$  Kcal mol<sup>-1</sup> M<sup>-1</sup>; as obtained from equilibrium urea-induced denaturation) (11) that jeopardizes an accurate determination of the equilibrium unfolding free energy by equilibrium urea-induced denaturation experiments. Hence, all mutants were investigated by thermal denaturation unfolding (Fig. S2 and Tables S1 and S2). Following Fersht (19), reliable  $\Delta\Delta G_{D-N}$  values induced by mutations were calculated as follows:

$$\Delta\Delta G_{D-N}^{\text{wt-mut}} = \Delta T_m^{\text{wt-mut}} \cdot \Delta S \quad [1]$$

The kinetic and equilibrium  $\Delta\Delta G_{D-N}$  values were both used to calculate  $\Phi$  values (Table 1); these two sets of  $\Phi$  values were found to be essentially the same within experimental error, confirming that, in the absence of stabilizing salt, the protein folds following a simple two-state model. In all cases  $\Phi$  values were fractional (Table 1) and their distribution along the structure revealed the presence of a diffuse nucleus, comprising elements of each of the three helices (Fig. 2). Overall, the structural distribution and the magnitude of  $\Phi$  values both suggest that Cter-NPM1 folds via a nucleation-condensation like mechanism (20–24), whereby folding proceeds via concurrent formation of secondary and tertiary structures, and the transition state should resemble a distorted version of the native state (Fig. S3).

In a recent study, we characterized the folding of Cter-NPM1 under different experimental conditions (11). Analysis of the kinetic  $m$  values according to Sanchez and Kiefhaber (25) revealed that the denatured state retains some residual native-like compact structure in the presence of stabilizing salt (0.5 M NaCl), while it corresponds to an extended random coil at destabilizing conditions (such as low pH); at physiological ionic strength and pH conditions a somewhat intermediate behavior was observed. Characterizing the structural features of a denatured state is a particularly difficult task. In this case we resorted to compare the effect of mutagenesis at low ionic strength with that obtained in the presence of a stabilizing salt (in this case NaCl). The rationale of such an approach is as follows.

While the  $\Phi$  values analysis allows characterization of the structure of folding transition states, it is extremely difficult to unveil interaction patterns in denatured states by protein engineering. In fact, by reference to a square thermodynamic cycle (19), changes in stability upon mutation may be calculated as follows:

$$\Delta\Delta G_{D-N}^{\text{wt-mut}} = \Delta\Delta G_{D-D'} - \Delta\Delta G_{N-N'} \quad [2]$$

where D' and N' represent the denatured and native states of the mutant protein, respectively. Because residues are generally taken not to interact in completely random unfolded conformations, the observed  $\Delta\Delta G_{D-N}^{\text{wt-mut}}$  typically equals the  $\Delta\Delta G_{N-N'}$ : i.e., in the absence of residual structure in the denatured state, the observed change in stability reflects the perturbation in the free energy of the native state induced by the mutation. In this work, we focused on the analysis of the residual structural features of the denatured state of Cter-NPM1, as induced by high salt concentrations (11). Hence, considering a cubic thermodynamic scheme (Fig. 3A), and comparing the changes in the stability induced by a mutation as obtained in the absence and in the presence of 0.5 M NaCl, the coupling free energy  $\Delta\Delta\Delta G_{D-N}^{0 \rightarrow 0.5 \text{ M}[\text{NaCl}]}$  can be calculated by the following equation:

$$\Delta\Delta\Delta G_{D-N}^{0 \rightarrow 0.5 \text{ M}[\text{NaCl}]} = (\Delta\Delta\Delta G_{\text{wt-mut}}^{0 \rightarrow 0.5 \text{ M}[\text{NaCl}]})_N - (\Delta\Delta\Delta G_{\text{wt-mut}}^{0 \rightarrow 0.5 \text{ M}[\text{NaCl}]})_D \quad [3]$$

This equation, formally similar to the analysis of double-mutant cycles (19, 26), implies that the measured coupling free

**Table 1. Kinetic and thermodynamic parameters for the folding of Cter-NPM1 site-directed mutants, at pH 7.0 and 283 K.**

|                     | $m_{D-N}^*$<br>(kcal mol <sup>-1</sup> M <sup>-1</sup> ) | $\Delta\Delta G_{D-N}^*$<br>(kcal mol <sup>-1</sup> ) | $\Phi^*$       | $\Delta\Delta G_{D-N}^\dagger$<br>(kcal mol <sup>-1</sup> ) | $\Phi^\dagger$                   | $m_{D-N}^\dagger$<br>(kcal mol <sup>-1</sup> M <sup>-1</sup> ) | $\Phi^\dagger$ | $\Delta\Delta\Delta G_{D-N}$<br>(kcal mol <sup>-1</sup> ) |
|---------------------|--|---|----------------|---|----------------------------------|--|----------------|---|
|                     | Low ionic strength                                       |   |                |   | High ionic strength (0.5 M NaCl) |  |                |   |
| V244A               | 1.02 ± 0.06  | 0.98 ± 0.16   | 0.16 ± 0.09    | 0.70 ± 0.08   | 0.23 ± 0.12                      | 0.80 ± 0.06  | -0.44 ± 0.19   | 0.32 ± 0.20   |
| E245A               | 0.91 ± 0.03  | 0.03 ± 0.14   | - <sup>§</sup> | -0.23 ± 0.03  | - <sup>§</sup>                   | 0.72 ± 0.04  | - <sup>§</sup> | -0.06 ± 0.23  |
| A245G               | 0.92 ± 0.04  | 0.56 ± 0.11   | 0.31 ± 0.10    | 0.32 ± 0.03   | 0.54 ± 0.16                      | 0.75 ± 0.04  | 0.16 ± 0.13    | -0.15 ± 0.18  |
| D246A               | 0.94 ± 0.03  | 0.19 ± 0.17   | - <sup>§</sup> | -0.28 ± 0.03  | - <sup>§</sup>                   | 0.77 ± 0.17  | - <sup>§</sup> | 0.42 ± 0.19   |
| A246G               | 0.97 ± 0.04  | 0.31 ± 0.16   | - <sup>¶</sup> | 0.40 ± 0.15   | - <sup>¶</sup>                   | 0.78 ± 0.11  | - <sup>¶</sup> | 0.23 ± 0.18   |
| I247V               | 1.00 ± 0.04  | 0.50 ± 0.14   | 0.14 ± 0.12    | 1.03 ± 0.11   | 0.07 ± 0.06                      | 0.70 ± 0.14  | 1.49 ± 1.78    | 0.44 ± 0.16   |
| K248A <sup>  </sup> | -  | -   | -              | -   | -                                | -  | -              | -   |
| A248G <sup>  </sup> | -  | -   | -              | -   | -                                | -  | -              | -   |
| A249G               | 0.89 ± 0.07  | 1.42 ± 0.16   | 0.29 ± 0.07    | 1.59 ± 0.18   | 0.26 ± 0.06                      | 0.66 ± 0.09  | 0.05 ± 0.03    | -0.11 ± 0.17  |
| K250A               | 0.85 ± 0.04  | -0.60 ± 0.27  | - <sup>§</sup> | -0.51 ± 0.07  | - <sup>§</sup>                   | 0.80 ± 0.02  | - <sup>§</sup> | 0.23 ± 0.28   |
| A250G               | 0.85 ± 0.08  | 1.11 ± 0.13   | 0.45 ± 0.08    | 0.85 ± 0.06   | 0.67 ± 0.10                      | 0.75 ± 0.05  | 0.05 ± 0.05    | -0.01 ± 0.27  |
| M251A <sup>  </sup> | -  | -   | -              | -   | -                                | -  | -              | -   |
| A253G               | 0.99 ± 0.07  | 1.30 ± 0.16   | 0.26 ± 0.08    | 1.75 ± 0.20   | 0.19 ± 0.06                      | 0.62 ± 0.07  | 0.25 ± 0.05    | 0.09 ± 0.18   |
| I255V               | 0.81 ± 0.04  | 0.77 ± 0.19   | 0.38 ± 0.13    | 1.18 ± 0.14   | 0.25 ± 0.07                      | 0.64 ± 0.13  | 0.62 ± 0.34    | 0.45 ± 0.22   |
| L261A               | 0.95 ± 0.24  | 1.95 ± 0.31   | -0.04 ± 0.07   | 2.06 ± 0.25   | -0.04 ± 0.07                     | 0.94 ± 0.05  | -0.22 ± 0.05   | 0.89 ± 0.31   |
| A266G               | 1.02 ± 0.05  | 0.49 ± 0.17   | 0.58 ± 0.29    | 0.40 ± 0.05   | 0.71 ± 0.27                      | 0.83 ± 0.07  | 1.65 ± 1.09    | 0.68 ± 0.20   |
| F268A <sup>  </sup> | -  | -   | -              | -   | -                                | -  | -              | -   |
| I269V               | 1.00 ± 0.03  | 0.23 ± 0.16   | - <sup>¶</sup> | 0.13 ± 0.02   | - <sup>¶</sup>                   | 0.86 ± 0.11  | - <sup>¶</sup> | 0.82 ± 0.20   |
| N270A               | 0.84 ± 0.04  | -0.43 ± 0.24  | - <sup>§</sup> | -0.72 ± 0.07  | - <sup>§</sup>                   | 0.73 ± 0.02  | - <sup>§</sup> | 0.00 ± 0.25   |
| A270G               | 0.93 ± 0.04  | 1.07 ± 0.16   | 0.84 ± 0.14    | 0.87 ± 0.16   | 1.03 ± 0.21                      | 0.74 ± 0.04  | 0.52 ± 0.14    | -0.27 ± 0.18  |
| V272A               | 1.22 ± 0.08  | 1.23 ± 0.15   | 0.40 ± 0.09    | 1.35 ± 0.16   | 0.36 ± 0.09                      | 0.86 ± 0.05  | 0.36 ± 0.08    | 0.55 ± 0.18   |
| K273A               | 0.92 ± 0.04  | 0.78 ± 0.16   | - <sup>§</sup> | 0.41 ± 0.06   | - <sup>§</sup>                   | 0.77 ± 0.05  | - <sup>§</sup> | 0.02 ± 0.18   |
| A273G               | 0.93 ± 0.08  | 0.74 ± 0.16   | 0.92 ± 0.08    | 1.39 ± 0.16   | 0.49 ± 0.09                      | 0.78 ± 0.04  | 0.70 ± 0.30    | -0.08 ± 0.18  |
| F276A <sup>  </sup> | -  | -   | -              | -   | -                                | -  | -              | -   |
| A283G               | 0.88 ± 0.04  | 0.84 ± 0.17   | 0.56 ± 0.14    | 1.11 ± 0.14   | 0.42 ± 0.09                      | 0.86 ± 0.07  | -0.17 ± 0.28   | 1.20 ± 0.20   |
| I284V               | 1.12 ± 0.06  | 0.98 ± 0.15   | 0.13 ± 0.10    | 1.56 ± 0.21   | 0.08 ± 0.06                      | 0.74 ± 0.06  | 0.86 ± 0.26    | 0.61 ± 0.17   |
| L287A               | 1.58 ± 0.52  | 2.31 ± 0.18   | 0.06 ± 0.05    | 2.09 ± 0.30   | 0.06 ± 0.06                      | 0.66 ± 0.15  | 0.24 ± 0.20    | -0.30 ± 0.54  |
| W288F               | 1.15 ± 0.25  | 2.20 ± 0.26   | 0.15 ± 0.10    | 2.95 ± 0.28   | 0.11 ± 0.08                      | 0.83 ± 0.05  | -0.39 ± 0.11   | 1.03 ± 0.28   |
| W290F               | 0.91 ± 0.10  | 2.24 ± 0.14   | 0.28 ± 0.04    | 1.69 ± 0.10   | 0.38 ± 0.05                      | 0.85 ± 0.09  | 0.27 ± 0.28    | 0.84 ± 0.40   |
| K292A               | 0.94 ± 0.04  | 0.67 ± 0.16   | - <sup>§</sup> | 1.66 ± 0.07   | - <sup>§</sup>                   | 0.72 ± 0.07  | - <sup>§</sup> | 0.20 ± 0.19   |
| A292G               | 0.94 ± 0.09  | 0.97 ± 0.17   | 0.36 ± 0.13    | 1.07 ± 0.05   | 0.33 ± 0.10                      | 0.66 ± 0.08  | 0.23 ± 0.14    | -0.17 ± 0.23  |
| L294A               | 0.67 ± 0.14  | 2.15 ± 0.23   | 0.26 ± 0.06    | 2.20 ± 0.03   | 0.25 ± 0.06                      | 0.95 ± 0.07  | -0.13 ± 0.07   | 1.67 ± 0.24   |

The errors reported are the standard errors. Repetition of individual experiments and calculation of standard deviations on wild-type Cter-NPM1 and some selected mutants revealed that obtained standard deviations were in the same order of magnitude of the standard errors.

\*Calculated from chevron plot analysis. Data recorded in the absence of NaCl.

†Calculated from thermal melt equilibrium denaturations in the absence of NaCl.

‡Calculated from chevron plot analysis. Data recorded in the presence of 0.5 M NaCl.

§Ala-Gly scanning mutants. A solvent exposed residue in the helix was mutated into Ala and Gly, and a  $\Phi$  value was calculated by comparing the folding kinetics of the Gly variant against its Ala counterpart.

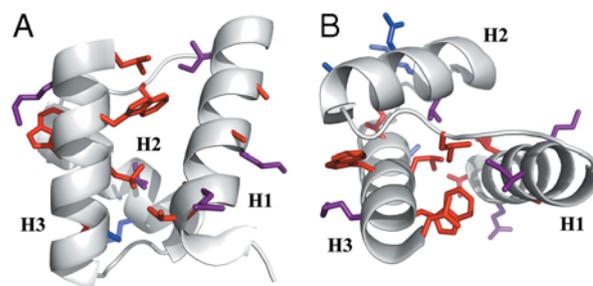
¶ $\Delta\Delta G_{D-N}$  was too low to calculate a reliable  $\Phi$  value.

||Mutants expressed poorly and could not be characterized.

energy for each mutant in the absence and presence of 0.5 M NaCl is equal to the difference in coupling energy between the native and the denatured states. If we assume that the free energy of the native state in high salt solutions is unlikely to be affected by conservative mutations mostly involving buried hydrophobic side-chains (27),  $(\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]})_N$  may be taken to be zero. Since we have shown that there is a compaction of the denatured state upon addition of salt, we consider that  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}$  is a measure of the contribution of a specific residue to the structure of the denatured state alone. It should be noted, however, that while  $\Phi$  values represent direct structural indexes,  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}$  simply reflects the contribution of each mutated side-chain to the stabilization of the structure nucleus of the denatured state when comparing data in the absence and in the presence of a stabilizing salt. Furthermore, a  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]} = 0$  does not necessarily imply that the probed residue is unstructured in the denatured state, but that there is no change in coupling free energy when considering the denatured state upon addition of salt. Hence, while analysis of  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}$  can in theory provide structural information

about compact denatured states, the calculated values should be interpreted with caution.

A complete equilibrium and kinetic characterization of all the mutants in the absence and in the presence of 0.5 M NaCl allowed calculating folding parameters and the resulting  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}$  values, all listed in Table 1. As depicted in a



**Fig. 2.**  $\Phi$  value analysis on Cter-NPM1.  $\Phi$  values are highlighted on the structure and clustered in three groups:  $0 < \Phi < 0.3$  kcal mol<sup>-1</sup> (Red);  $0.3 < \Phi < 0.5$  kcal mol<sup>-1</sup> (Magenta);  $\Phi > 0.5$  kcal mol<sup>-1</sup> (Blue). A broad diffused folding nucleus, consistent with a nucleation-condensation type folding mechanism, may be identified. The two panels (A and B) refer to two different orientations of the molecule.

color-coded representation in Fig. 3 *B* and *C*, a structural distribution of  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N}$  clearly shows that mutations displaying the highest values ( $>0.5$  Kcal mol $^{-1}$ ) all map onto the helices H2 and H3 and their interface, whereas helix H1 includes only residues with smaller  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N}$  values ( $<0.5$  Kcal mol $^{-1}$ ). These data clearly suggest that the region comprised between helices H2 and H3 is likely to partially retain its native conformation in the denatured state.

There is a growing consensus that denatured states of proteins do not behave as fully unfolded chains (15, 16, 28–31). However, little is known about the mechanism whereby their residual structure may be affected by denaturants or mutagenesis. Is the denatured state under physiological conditions in equilibrium with the fully unfolded conformation or does it unfold via a gradual melting of structure? Does the denatured state unfold via a first- or a second-order type transition? Taking advantage of the folding and unfolding parameters obtained for the site-directed mutants of Cter-NPM1, we addressed this issue by analyzing the dependence of the observed  $m_{D-N}$  values on  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N}$ . In fact, while according to the analysis outlined above  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N}$  values reflect the effect of any given mutation on the free energy of the denatured state, the  $m_{D-N}$  value would indicate its overall degree of compaction. Interestingly, data in Fig. 4 show a linear dependence of  $m_{D-N}$  on  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N}$  ( $R = 0.73$ ), with greater values of  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N}$  (i.e., stronger destabilization of D) corresponding to greater values of  $m_{D-N}$  (i.e., more unfolded D). Furthermore, it may be noted that such dependence is mainly due to an effect on the  $m_F$  value, the calculated  $m_U$  values being essentially independent of  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N}$  ( $R = 0.17$ ). These data strongly suggest that the denatured state responds to destabilizing mutations by altering its structure and indicate that the variable two-state scenario (32) would represent the best model to describe the folding and unfolding behavior of Cter-NPM1 and its mutants.

## Discussion

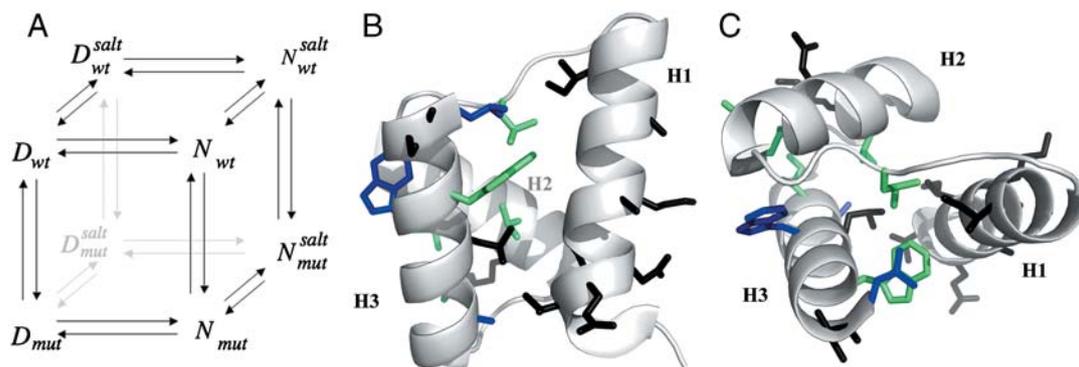
Nucleophosmin is a key nucleolar chaperone that was discovered to be mutated in several haematological malignancies. In general, mutations compromise the physiological cellular traffic of the protein: Indeed, more than 30% of adult patients affected by AML show aberrant nucleophosmin localization in the cell cytoplasm (33). This led to the recent inclusion of AML with mutated NPM1 into the World Health Organization classification of myeloid neoplasms as a new provisional entity (34). All NPM1 mutations, in spite of remarkable heterogeneity (about 50 mutations

identified so far), result in a shift of the reading frame and cause changes at the C terminus of the protein that are responsible for its partial or total denaturation (1). This observation suggests that the design of small molecules capable of stabilizing NPM1 native fold compromised in pathologic mutants might have a therapeutic value, possibly enhancing relocation of the mutants in their native compartment. Recently, we have shown that Cter-NPM1 retains some residual structure in the denatured state in the presence of salt (11). Because folding relies on the equilibrium between the native and denatured states, mapping the residual structure in the denatured state is a key step in any attempt to design structure-based drugs (8, 9, 35).

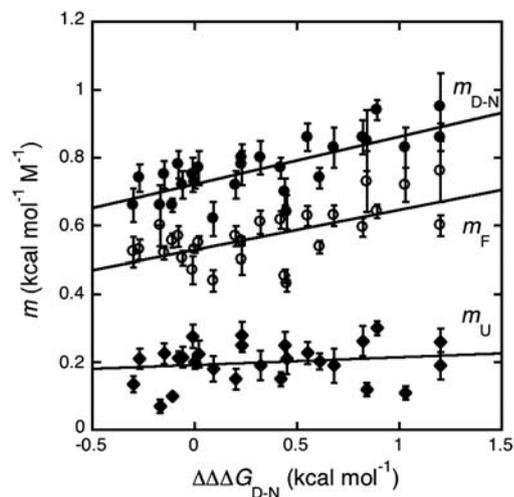
In this work we have characterized the folding transition state structure of Cter-NPM1 and deciphered the overall structural features of its denatured state nucleus through kinetics and extensive protein engineering. In particular, we performed a conventional  $\Phi$  value analysis on about 30 conservative mutants. This protein displays a folding rate constant that is only marginally slower than what predicted from its calculated contact order (36), i.e., contact order of 10.5% corresponding to a folding rate constant of circa  $10^4$  s $^{-1}$  vs. the experimentally determined value of  $4 \times 10^3$  s $^{-1}$ . Data revealed the presence of a diffused nucleus for folding, as mirrored by the presence of fractional  $\Phi$  values in essentially all probed positions. Furthermore, these values were broadly distributed along the Cter-NPM1 structure, suggesting that this small protein domain folds via a nucleation-condensation mechanism (20, 21, 24). This observation is somewhat surprising, because this type of mechanism is generally invoked for proteins that display little or no residual structure in their denatured states (23, 37). Indeed, this behavior, in analogy to what recently observed for the homeodomain-like proteins, has been associated to a hybrid between the nucleation-condensation and diffusion-collision mechanisms (23, 38, 39).

In an effort to unveil the structural features of the denatured state, we subjected all mutants to equilibrium and kinetic folding experiments under conditions promoting denatured state structure, i.e., in the presence of a stabilizing salt (0.5 M NaCl). The rationale for this approach, represented in Fig. 3*A* and discussed in *Results*, relies on the emerging view that, because the denatured state ensemble can contain energetically significant interactions, mutations can exert on the denatured state effects that are comparable to those seen for the native or the transition states, (40) and references therein.

The results of the  $\Phi$  value analysis in the presence of salt are shown in Table 1. Interestingly, while all the  $\Phi$  values calculated in the absence of salt were fractional, five of the  $\Phi$  values obtained in its presence appear to be nonconventional, i.e., negative. This finding may be explained following Cho and Raleigh (40), who argued that when a mutation changes the free energy



**Fig. 3.** Denatured state structure of Cter-NPM1 as obtained from protein engineering. (A) Double perturbation cycle. The cube depicts the different states of the native (N) and denatured (D) conformations populated as a result of a double perturbation (i.e., mutagenesis and addition of stabilizing salt). (B and C) Coupling free energies mapped on the structure shown in two orientations ( $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N} < 0.5$  kcal mol $^{-1}$  (Black);  $0.5$  kcal mol $^{-1} < \Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N} < 1$  kcal mol $^{-1}$  (Green);  $\Delta\Delta\Delta G_{wt-mut}^{0 \rightarrow 0.5 M[NaCl]}_{D-N} > 1$  kcal mol $^{-1}$  (Blue).



**Fig. 4.** Correlation between  $m_F$  (Open Circles),  $m_U$  (Diamonds), and  $m_{D-N}$  (Closed Circles) values (recorded in the presence of 0.5 M NaCl) versus  $\Delta\Delta\Delta G_{wt-mut}^{0-0.5 M[NaCl]}$  values. Lines are the best fit to a linear function ( $m_F$ ,  $R = 0.84$ ;  $m_U$ ,  $R = 0.17$ ;  $m_{D-N}$ ,  $R = 0.73$ ). As described in the text, the observed correlation of both  $m_F$  and  $m_{D-N}$  suggests that the denatured state retains a malleable structure, which is perturbed by mutagenesis.

level of all the states (denatured, native, and transition), the resulting  $\Phi$  value may be negative because of a change in the free energy of the denatured state. Detailed analysis of the effect of mutations on the free energy of the denatured state has been previously reported (41–43).

As described in *Results*, a comparison between the datasets obtained in the absence and in the presence of salt allows an estimate of the contribution of a given mutation to the stabilization of the denatured state residual structure. Notably for Cter-NPM1, the residues with the highest values of  $\Delta\Delta\Delta G_{wt-mut}^{0-0.5 M[NaCl]}$  ( $>0.5$  Kcal mol $^{-1}$ ) all map on helices H2 and H3 and at the surface in between, contrary to residues in H1 ( $\Delta\Delta\Delta G_{wt-mut}^{0-0.5 M[NaCl]}$   $<0.5$  Kcal mol $^{-1}$ ). These results indicate that H2 and H3 retain some residual structure in the denatured state of Cter-NPM1. Furthermore, the linear dependence of  $m_{D-N}$  on  $\Delta\Delta\Delta G_{wt-mut}^{0-0.5 M[NaCl]}$  indicates that such a residual structure is malleable, at variance with discrete folding intermediates (44, 45) and similarly to what was recently observed in the Engrailed homeodomain (16), where the denatured state under physiological conditions unfolds in a gradual fashion upon destabilization. According to our data on Cter-NPM1, helix H1 would represent the structural portion of the protein to be targeted by chemicals in an effort to favor the native structure, and thereby our findings add previously undescribed information for the selection of drugs that counteract leukemic mutations. So far, in fact, the region comprising helix H2 has been mainly envisaged for ligand binding, as it includes important functional residues (e.g. K263 and

K267) (46) and the Cys 275, which is responsible for the binding of the natural antitumoral compound Avrainvillamide (47).

In the context of the fight against AML, this work paves the way to an investigation of libraries of pharmaceuticals that, examined by appropriate revealing assays (48), may restore the native fold of leukemic variants of NPM1 by stabilizing helix H1 and ultimately restore the physiological subcellular localization of nucleophosmin.

## Materials and Methods

**Mutagenesis, Expression, and Purification.** All mutants were realized by PCR using a QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) and using the Cter-NPM1 cDNA cloned into a pET28a vector (Novagen, San Diego, CA) as template. The construct codes for the C-terminal domain of NPM1 ( $\Delta N241$ ) fused to a His6-tag. The presence of the His6-tag had a negligible effect on the folding and unfolding kinetics of Cter-NPM1 (Fig. S1). The expression system was *Escherichia coli* strain BL21(DE3) and proteins were purified using Ni $^{2+}$  chelating affinity chromatography, followed by ion-exchange and size exclusion chromatography. SDS-PAGE and mass spectrometry were used to confirm the protein purity and mass. All reagents were purchased from Sigma-Aldrich Corp. (St. Louis, MO) and were of analytical grade.

**Circular Dichroism (CD).** CD spectra of Cter-NPM1 wt and mutants were recorded between 260 and 190 nm at a concentration of 10  $\mu$ M using a Jasco spectropolarimeter (Jasco, Inc., Easton, MD) and a 1-mm pathlength quartz cuvette (Hellma, Plainview, NY). Thermal denaturation was followed at 222 nm with heating from 277 to 363 K at a rate of 1 K min $^{-1}$  (Fig. S2). Data were fitted to a standard 2-state denaturation (19). An equation that takes into account the pre- and post-transition baselines was used to fit the observed unfolding transition (49). The buffers used were 50 mM sodium phosphate pH 7.0, 1 mM DTT either with or without 0.5 M sodium chloride.

**Stopped-Flow Measurements.** Single mixing kinetic folding experiments were carried out on a  $\Pi$  stopped-flow instrument (Applied Photophysics, Leatherhead, UK); the excitation wavelength was 280 nm and the fluorescence emission was measured using a 320-nm cut-off glass filter. In all experiments, performed at 283 K, refolding and unfolding were initiated by an 11-fold dilution of the denatured or the native protein with the appropriate buffer. Final protein concentrations were typically 1  $\mu$ M. The kinetics observed were always independent of protein concentration, as expected for a monomolecular reaction without effects due to transient aggregation (50).

**Data Analysis.** Analysis of observed time courses was performed by nonlinear least-squares fitting of single exponential time courses using the fitting procedures provided in the Applied Photophysics software. The chevron plots were fitted by numerical analysis based on a two-state model following the equation:

$$k_{obs} = k_F + k_U \quad [4]$$

where  $k_F$  and  $k_U$  represent the folding and unfolding rate constants, respectively. The logarithm of each microscopic rate constant was assumed to vary linearly with denaturant concentration (51).  $\Phi$  values were calculated from folding rate constants using standard equations (17).

**ACKNOWLEDGMENTS.** We express our thanks to Professor Brunangelo Falini (Perugia, Italy) for introducing us to nucleophosmin and for his precious collaboration. This work was supported by a grant from the Associazione Italiana Ricerca sul Cancro to M.B.

- Falini B, et al. (2009) Altered nucleophosmin transport in acute myeloid leukaemia with mutated NPM1: Molecular basis and clinical implications. *Leukemia* 23:1731–1743.
- Grisendi S, Mecucci C, Falini B, Pandolfi PP (2006) Nucleophosmin and cancer. *Nat Rev Cancer* 6:493–505.
- Borer RA, Lehner CF, Eppenberger HM, Nigg EA (1989) Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell* 56:379–390.
- Falini B, Nicoletti I, Martelli MF, Mecucci C (2007) Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): Biologic and clinical features. *Blood* 109:874–885.
- Bolli N, et al. (2009) A dose-dependent tug of war involving the NPM1 leukemic mutant, nucleophosmin, and ARF. *Leukemia* 23:501–509.
- Bonetti P, et al. (2008) Nucleophosmin and its AML-associated mutant regulate c-Myc turnover through Fbw7 gamma. *J Cell Biol* 182:19–26.
- Di Fiore PP (2008) Playing both sides: Nucleophosmin between tumor suppression and oncogenesis. *J Cell Biol* 182:7–9.
- Joerger AC, Fersht AR (2007) Structure-function-rescue: The diverse nature of common p53 cancer mutants. *Oncogene* 26:2226–2242.
- Bullock AN, Fersht AR (2001) Rescuing the function of mutant p53. *Nat Rev Cancer* 1:68–76.
- Pey AL, et al. (2008) Identification of pharmacological chaperones as potential therapeutic agents to treat phenylketonuria. *J Clin Invest* 118:2858–2867.
- Scaloni F, Gianni S, Federici L, Falini B, Brunori M (2009) Folding mechanism of the C-terminal domain of nucleophosmin: Residual structure in the denatured state and its pathophysiological significance. *FASEB J* 23:2360–2365.
- Schwalbe H, et al. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry* 36:8977–8991.

13. Bai Y, Chung J, Dyson HJ, Wright PE (2001) Structural and dynamic characterization of an unfolded state of poplar apo-plastocyanin formed under nondenaturing conditions. *Protein Sci* 10:1056–1066.
14. Neira JL, et al. (1996) Towards the complete structural characterization of a protein folding pathway: The structures of the denatured, transition and native states for the association/folding of two complementary fragments of cleaved chymotrypsin inhibitor 2. Direct evidence for a nucleation-condensation mechanism. *Fold Des* 1:189–208.
15. Shortle D, Meeker AK (1989) Residual structure in large fragments of staphylococcal nuclease: Effects of amino acid substitutions. *Biochemistry* 28:936–944.
16. Religa TL, Markson JS, Mayor U, Freund SM, Fersht AR (2005) Solution structure of a protein denatured state and folding intermediate. *Nature* 437:1053–1056.
17. Fersht AR, Matouschek A, Serrano L (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J Mol Biol* 224:771–782.
18. Myers JK, Pace CN, Scholtz JM (1995) Denaturant  $m$  values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci* 4:2138–2148.
19. Fersht AR (1999) *Structure and mechanism in protein science* (Freeman, New York).
20. Abkevich VI, Gutin AM, Shakhnovich EI (1994) Specific nucleus as the transition state for protein folding: Evidence from the lattice model. *Biochemistry* 33:10026–10036.
21. Fersht AR (1995) Optimization of rates of protein folding: The nucleation-condensation mechanism and its implications. *Proc Natl Acad Sci USA* 92:10869–10873.
22. Gianni S, et al. (2007) A PDZ domain recapitulates a unifying mechanism for protein folding. *Proc Natl Acad Sci USA* 104:128–133.
23. Gianni S, et al. (2003) Unifying features in protein-folding mechanisms. *Proc Natl Acad Sci USA* 100:13286–13291.
24. Itzhaki LS, Otzen DE, Fersht AR (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analyzed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J Mol Biol* 254:260–88.
25. Sanchez IE, Kiefhaber T (2003) Hammond behavior versus ground state effects in protein folding: Evidence for narrow free energy barriers and residual structure in unfolded states. *J Mol Biol* 327:867–84.
26. Horovitz A (1996) Double-mutant cycles: a powerful tool for analyzing protein structure and function. *Fold Des* 1:R121–R126.
27. Timasheff SN (1993) The control of protein stability and association by weak interactions with water: How do solvents affect these processes?. *Annu Rev Biophys Biomol Struct* 22:67–97.
28. McCarney ER, Kohn JE, Plaxco KW (2005) Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins. *Crit Rev Biochem Mol Biol* 40:181–189.
29. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA (2007) Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc Natl Acad Sci USA* 104:1528–1533.
30. Mok KH, et al. (2007) A pre-existing hydrophobic collapse in the unfolded state of an ultrafast folding protein. *Nature* 447:106–109.
31. van Gunsteren WF, Bürgi R, Peter C, Daura X (2001) The key to solving the protein-folding problem lies in an accurate description of the denatured state. *Angew Chem Int Ed Engl* 40:351–355.
32. Ferreon AC, Bolen DW (2004) Thermodynamics of denaturant-induced unfolding of a protein that exhibits variable two-state denaturation. *Biochemistry* 43:13357–13369.
33. Falini B, et al. (2005) Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med* 352:254–266.
34. Falini B, Martelli MP (2009) Anaplastic large cell lymphoma: changes in the World Health Organization classification and perspectives for targeted therapy. *Haematologica* 94:897–900.
35. Friedler A, et al. (2002) A peptide that binds and stabilizes p53 core domain: chaperone strategy for rescue of oncogenic mutants. *Proc Natl Acad Sci USA* 99:937–942.
36. Plaxco KW, Simons KT, Baker D (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* 277:985–994.
37. Daggett V, Fersht AR (2003) Is there a unifying mechanism for protein folding?. *Trends Biochem Sci* 28:18–25.
38. Karplus M, Weaver DL (1994) Protein folding dynamics: the diffusion-collision model and experimental data. *Protein Sci* 3:650–668.
39. White GV, et al. (2005) Simulation and experiment conspire to reveal cryptic intermediates and a slide from the nucleation-condensation to framework mechanism of folding. *J Mol Biol* 350:757–775.
40. Cho JH, Raleigh DP (2006) Denatured state effects and the origin of nonclassical  $\phi$  values in protein folding. *J Am Chem Soc* 128:16492–16493.
41. Cho JH, Raleigh DP (2009) Experimental characterization of the denatured state ensemble of proteins. *Methods Mol Biol* 490:339–351.
42. Fersht AR, Sato S (2004)  $\phi$ -value analysis and the nature of protein-folding transition states. *Proc Natl Acad Sci USA* 101:7976–7981.
43. Scott KA, Alonso DO, Sato S, Fersht AR, Daggett V (2007) Conformational entropy of alanine versus glycine in protein denatured states. *Proc Natl Acad Sci USA* 104:2661–2666.
44. Dalby PA, Oliveberg M, Fersht AR (1998) Folding intermediates of wild-type and mutants of barnase. I. Use of  $\phi$ -value analysis and  $m$ -values to probe the cooperative nature of the folding pre-equilibrium. *J Mol Biol* 276:625–646.
45. Uversky VN, Ptitsyn OB (1996) All-or-none solvent-induced transitions between native, molten globule and unfolded states in globular proteins. *Fold Des* 1:117–122.
46. Grummitt CG, Townsley FM, Johnson CM, Warren AJ, Bycroft M (2008) Structural consequences of nucleophosmin mutations in acute myeloid leukemia. *J Biol Chem* 283:23326–23332.
47. Wulff JE, Siegrist R, Myers AG (2007) The natural product avrainvillamide binds to the oncoprotein nucleophosmin. *J Am Chem Soc* 129:14444–51.
48. Kau TR, Way JC, Silver PA (2004) Nuclear transport and cancer: from mechanism to intervention. *Nat Rev Cancer* 4:106–117.
49. Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27:8063–8.
50. Silow M, Oliveberg M (1997) Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc Natl Acad Sci USA* 94:6084–6.
51. Jackson SE, Fersht AR (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* 30:10428–10435.