Biochimica et Biophysica Acta 1865 (2017) 176–185

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Insights into amyloid-like aggregation of H2 region of the C-terminal domain of nucleophosmin

Anna Russo [a], Carlo Diaferia [a], Sara La Manna [a], Cinzia Giannini [b], Teresa Sibillano [b], Antonella Accardo [a], Giancarlo Morelli [a], Ettore Novellino [a], Daniela Marasco [a,*]

[a] Department of Pharmacy, CIRPEB: Centro Interuniversitario di Ricerca sui Peptidi Bioattivi- University of Naples "Federico II", 80134 Naples, Italy
[b] Institute of Crystallography (IC), National Research Council, Bari 70125, Italy

ARTICLE INFO

Article history:
Received 2 September 2016
Received in revised form 29 October 2016
Accepted 14 November 2016
Available online 17 November 2016

Keywords:
Misfolding
Aggregation
Acute myeloid leukemia
Wide-angle X-ray scattering

ABSTRACT

Nucleophosmin (NPM1) is a multifunctional protein involved in a variety of biological processes including the pathogenesis of several human malignancies and is the most frequently mutated gene in Acute Myeloid Leukemia (AML). To deepen the role of protein regions in its biological activities, lately we reported on the structural behavior of dissected C-terminal domain (CTD) helical fragments. Unexpectedly the H2 (residues 264–277) and H3 AML-mutated regions showed a remarkable tendency to form amyloid-like assemblies with fibrillar morphology and β-sheet structure that resulted as toxic when exposed to human neuroblastoma cells. More recently NPM1 was found to be highly expressed and toxic in neurons of mouse models of Huntington's disease (HD). Here we investigate the role of each residue in the β-strand aggregation process of H2 region of NPM1 by performing a systematic alanine scan of its sequence and structural and kinetic analyses of aggregation of derived peptides by means of Circular Dichorism (CD) and Thioflavin T (Th-T) assay. These solution state investigations pointed out the crucial role exerted by the basic amyloidogenic stretch of H2 (264–271) and to shed light on the initial and main interactions involved in fibril formation we performed studies on fibrils deriving from the related Ala peptides through the analysis of fibrils with birefringence of polarized optical microscopy and wide-angle X-ray scattering (WAXS). This analysis suggested that the presence of branched Ile^{269} conferred preferential packing patterns that, instead, appeared geometrically hampered by the aromatic side-chain of Phe^{268}. Present investigations could be useful to deepen the knowledge of AML molecular mechanisms and the role of cytoplasmatic aggregates of NPM1c+.

1. Introduction

In processes of protein misfolding the amyloid fibrillization is a unique ordered state governed by specific patterns of molecular interactions and not an aspecific aggregation [1,2]. The process that leads to the formation of amyloid aggregates is a heterogeneous multistep reaction with many parallel events [3] and, nevertheless the common structural features of final fibrils [4,5], initiation steps of the aggregation by different proteins can be significantly diverse. Even if the primary sequences do not determine uniquely the structures of final fibrils [6–8], the details of structural variations can be markedly sequence-dependent [9]. The determination of conformational elements that underlay molecular recognition and self-assembly is crucial to understand amyloid formation [10]. Normally folded proteins can access to amyloidogenic states as a result of thermal fluctuations of the native state [11] or disruption of the quaternary structure [12] and the amyloidogenic state is often considered as an ensemble of native-like conformations with locally unfolded elements [13,14]. The characterization of these amyloidogenic species is crucial to elucidate potential aggregation under native conditions and for *in vivo* aggregation events [15]. Many peptides and proteins convert *in vitro* into highly organized amyloid structures and the investigations of the effects of sequence changes on the aggregation can help the development of therapeutical strategies for amyloid-associated diseases [16].

Nucleophosmin (NPM1, B23, No38 and numatrin) is a multifunctional protein, present in the granular region of nucleoli [17,18], belonging to the nucleoplasmin family of nuclear chaperones [19] that shuttles between the nucleus and cytoplasm [20]. In human cells this protein is expressed in three distinct isoforms: B23.1, the longest and the most abundant (294 residues), B23.2 and B23.3 splicing variants lacking the

Abbreviations: TIS, Triisopropylsilane; TFA, Trifluoroacetic acid; DMF, Dimethylformamide; HBTU, 1-H-Benzotriazolium, 1-[bis(dimethylamino)methylene]-Hexafluorophosphate(1-),3 oxide; HOBt, N-hydroxybenzotriazole; DIEA, Di-isopropylethylamine; Fmoc, Fluorenylmethoxycarbonyl; HPLC, High Performance Liquid Chromatography; LC-MS, Liquid Chromatography Mass Spectrometry; CD, Circular dichroism; CTD, C-terminal domain; NPM1, Nucleophosmin 1; AML, Acute myeloid leukemia; ThT, Thioflavin T; CR, Congo Red; WAXS, Wide-angle X-ray scattering.

* Corresponding author at: Department of Pharmacy, University "Federico II", Via Mezzocannone, 16, 80134 Naples, Italy.
E-mail address: daniela.marasco@unina.it (D. Marasco).

http://dx.doi.org/10.1016/j.bbapap.2016.11.006
1570-9639/

C-terminal 35 amino-acids and a 29 amino-acid stretch (residues 195–223) in the basic region, respectively [21]. Structural investigations showed that the protein is endowed with a modular structure: i) the N-terminal region is an oligomerization domain crucial for chaperone function [22,23], the central portion presents two IDRs (Intrinsically Disordered Regions) [24] and is crucial for DNA/RNA recognition mechanism [25–27] and the C-terminal domain (CTD) with a three helix bundle tertiary structure [28]. A NMR structural study allowed to delineate the first helix, H1, (243–259), the second, H2, (264–277), and the third H3 (280–294), within the wt CTD that demonstrated to fold through a compact transition state with an extended nucleus and unfolds keeping a malleable residual secondary structure at the interface between H2 and H3 helices [29–31].

NPM1 is overexpressed in solid tumors correlating with mitotic index and metastatization [32] and is the most frequently mutated gene in Acute Myeloid Leukemia (AML) patients: specific mutations in the exon 12 of the NPM1 gene occur, accounting for approximately 30% of cases [19,33–39]. Mutations in the H3 region of the CTD cause its unfolding [28,40], the impairment of several functions such as DNA/RNA recognition [39] and the aberrant accumulation of the protein in the cytoplasm of the leukemic cells (thus the term NPM cytoplasmic positive NPMc + AML) [36]. AML NPM1 mutations revealed involved in leukemia cell viability and invasion and, in this process, Matrix metalloproteinases MMPs (regulated by the K-Ras/ERKMAPK signaling pathway) play a key role [41]. The aberrant cytoplasmatic accumulation of AML mutated protein is due both to the loss of the NoLS (nucleolar localization signal) and to the oligomerization level of the protein depending on N-terminal and/or C-terminal domains. Recently a study focused on the role of NPM1 in neurons of mouse models of Huntington's disease (HD) was reported. It demonstrated that NPM1 has complex roles in the regulation of neuronal survival in dependence on its levels of expression, oligomerization and subcellular localization: in particular NPM1 resulted highly expressed and toxic with a cytoplasmatic localization [42].

Lately, following a structure-based protein dissection approach, we demonstrated that the H2 [43] and H3 AML mutated regions [44] of the CTD of NPM1 form amyloid-like assemblies endowed with fibrillar morphology and β-sheet structure that resulted toxic in cell viability assays. Actually, our mechanistic hypothesis is that the AML-associated mutations destabilize the α-helical structure of the H3 region in the native NPM1 and disrupts the CTD tertiary structure predisposing it to the formation of toxic aggregates since it induces the exposure of the H2 region, which is the most amyloidogenic region of the whole NPM1-CTD.

Here we investigate the role of each residue in the aggregation process of H2 region of NPM1 through the characterization of a series of peptides deriving from a systematic alanine scan [45] of its sequence. Bioinformatics analyses pointed out the chameleon nature of this region and that Alanine-substitutions deeply affect the "basic amyloidogenic unit" (264–271) as confirmed by structural and kinetic analyses of aggregation processes by means of Circular Dichorism, Infrared spectroscopies and Th-T assay. Further, to shed light on the initial and main interactions in fibril formation, we analyzed amyloid fibrils of H2 wt and several sequences bearing Ala substitutions with birefringence of polarized optical microscopy and wide-angle X-ray scattering (WAXS).

2. Results and discussion

2.1. Chameleon behavior of H2-Ala scan peptides

To investigate the specific role of the residues of the region H2 of NPM1-CTD (264–277), we performed a systematic alanine scan of the sequence and thirteen peptides (Table 1) were synthesized and purified with good yields (>90%).

A bioinformatic analysis of each primary sequence of the H2 Ala-scan peptides was carried out using the PASTA server [46]. Their tendencies to assume β-strand secondary structure and to aggregate in an amyloid-like way respect to disorder were analyzed and probability profiles were reported in Figs. S1 and S2, respectively. Notably a similar analysis, carried out on the NPM1-CTD wt outlined the chameleon character of H2 sequence and defined a shorter region as the "basic amyloidogenic unit" (named H2 short$^{265–272}$) [43]. Here PASTA analysis indicates that Alanine point substitutions do not alter β secondary structure preferences respect to H2 wt. Exceptions were for the Ala/Asn^{270} (H2A7) and Ala/Lys^{273} (H2A10) (to a lesser extend) replacements (Fig. S1G,J) for which the probability to assume a beta conformation seems comparable to aggregation. Similarly the greater probability to aggregate respect to disorder is not affected by substitutions except for Ala/Ile^{269} (H2A6) (Fig. S2F) that seemed, in turn, to decrease the tendency of the related sequence to aggregate.

Table 1
Sequences and names of analyzed peptides derived from Ala scan of H2 wt. Molecular weight (*MW*) and isoelectronic point (*pI*) for each peptide sequence are also reported.

Name	*Sequence*	*MW (amu)*	*pI*
H2wt	VEAKFINYVKNCFR	1772	10.15
H2A1, Ala/Val^{264}	**A**EAKFINYVKNCFR	1744	10.15
H2A2, Ala/Glu^{265}	V**A**AKFINYVKNCFR	1714	10.76
H2A4, Ala/Lys^{267}	VEA**A**FINYVKNCFR	1715	9.06
H2A5, Ala/Phe^{268}	VEAK**A**INYVKNCFR	1696	10.15
H2A6, Ala/Ile^{269}	VEAKF**A**NYVKNCFR	1730	10.15
H2A7, Ala/Asn^{270}	VEAKFI**A**YVKNCFR	1729	10.15
H2A8, Ala/Tyr^{271}	VEAKFIN**A**VKNCFR	1680	10.64
H2A9, Ala/Val^{272}	VEAKFINY**A**KNCFR	1744	10.15
H2A10, Ala/Lys^{273}	VEAKFINYV**A**NCFR	1715	9.06
H2A11, Ala/Asn^{274}	VEAKFINYVK**A**CFR	1729	10.15
H2A12, Ala/Cys^{275}	VEAKFINYVKN**A**FR	1740	10.78
H2A13, Ala/Phe^{276}	VEAKFINYVKNC**A**R	1696	10.15
H2A14, Ala/Arg^{277}	VEAKFINYVKNCF**A**	1687	9.01

The effects of Ala substitutions on conformational features were experimentally evaluated by CD spectra over time. H2 wt sequence revealed able to self-assemble in an amyloid-like way at physiological pH but basic pHs demonstrated to speed this process and in order to compare the β-sheet aggregating propensities of Ala sequences in the absence of net positive charges, we performed studies at pH 10, that is the closest value to theoretical pIs (Table 1) [43]. H2 Ala derived peptides freshly dissolved in 10 mM borate buffer, at pH 10.0, exhibited different far-UV CD spectra (Fig. 1). Each spectrum was followed within 24–48 h. Most spectra presented decreases over time in the overall signal, resulting from the formation of large aggregates that escape experimental detection. The large mean residue ellipticity in the 215–220 nm range indicated a β-structure content [47]. This feature was mainly detected in H2A9, H2A11, H2A12 and H2A14 sequences: notably all of them, except H2A9, present Ala substitutions outside the indicated basic amyloidogenic stretch (Fig. 1H,J,K,M). Visually H2A9 did not formed fibrils in the explored time-course (differently from several other sequences), presumably the substitution Ala/Val^{272} allowed a major solubility of prefibrillar oligomers suggesting that Val^{272} is not determinant in the oligomerization mechanism and could only enhance insolubility of wt sequence as confirmed by the solubility profile reported in Fig. S3-I, see below. Conversely, spectra of H2A5, H2A6 and H2A8 did not assume β-strand profiles within 24–48 h: while H2A5 and H2A8 showed only decreased Cotton effects with a substantial random profile H2A6 presented a transition from random to α + β mixed conformation, as predicted by PASTA analysis (Fig. 1D,E,G). Aromatic interactions can greatly affect to the self-assembly process: indeed stacking interactions can provide energetic contributions, as well as order and directionality in the aggregation [48]. H2 sequence contains three aromatic residues: two inside the shortest amyloidogenic region, Phe^{268} and Tyr^{271}, and one outside it, Phe^{276}. Their Ala substitutions caused that none of related peptides presented β-like CD spectra (Fig. 1D,G,L) at micromolar concentrations: notably this effect is more evident for Phe^{268} with respect to Phe^{276}, pointing out the dependence of the local context on conformational behavior. This dependence is also evident in the