Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

The ubiquitin *versus* NEDD8 binding preference of NEDD4 binding protein 1 (N4BP1) is based on a mutual conformational perturbation.

Ridvan Nepravishta[1#], Federica Ferrentino[2#], Walter Mandaliti[3#], Anna Mattioni[2], Luisa Castagnoli[2], Gianni Cesareni[2,4*], Maurizio Paci[3*] and Elena Santonico[2*]

[1]School of Pharmacy East Anglia, University of Norwich, UK; [2]Department of Biology, University of Tor Vergata, Rome, Italy; [3]Dpt. of Chemical Sciences and Technologies, Tor Vergata University, Rome, Italy; [4]Fondazione Santa Lucia Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), 00143 Rome, Italy.

* Corresponding authors
# These authors contributed equally to this work

To whom correspondence should be addressed: Gianni Cesareni, University of Rome Tor Vergata, Via della ricerca scientifica, 00133 Rome, Italy. E-mail: Cesareni@uniroma2.it; Maurizio Paci, University of Rome Tor Vergata, Via della ricerca scientifica, 00133 Rome, Italy. E-mail: Paci@uniroma2.it. Elena Santonico, University of Rome Tor Vergata, Via della ricerca scientifica, 00133 Rome, Italy. E-mail: Elena.Santonico@uniroma2.it.

Keywords: CUBAN; Ubiquitin-binding domains (UBDs); KHNYN; NEDD8; NMR spectroscopy.

Abstract

Ubiquitin binding domains (UBDs) are modular elements that bind non-covalently to ubiquitin and act as downstream effectors and amplifiers of the ubiquitination signal. With few exceptions, UBDs recognize the hydrophobic path centered on Ile44 (Leu-8, Ile-44, Val-70). Nevertheless, a variety of different orientations, which can be attributed to specific contacts between each UBD and surface residues surrounding the hydrophobic patch, specify how each class of UBD recognizes ubiquitin. Here, we describe the structure of a novel ubiquitin-binding domain that we identified in NEDD4 binding protein 1 (N4BP1). By performing protein sequence analysis, mutagenesis and NMR spectroscopy of the ^{15}N isotopically labelled protein, we demonstrate that a Phe-Pro motif in N4BP1 recognizes the canonical hydrophobic patch of ubiquitin. This recognition mode resembles the molecular mechanism evolved in the CUE (Coupling of ubiquitin conjugation to ER degradation) domain family, where an invariant proline, usually following a phenylalanine, is required for binding to ubiquitin. Interestingly, the UBD of N4BP1 is evolutionary related to CUBAN (Cullin binding domain associating with NEDD8) (40% identity and 47% similarity), a protein module that also recognizes the ubiquitin-like NEDD8, which is the closest relative of ubiquitin (58% identity and 80% similarity). By performing circular dichroism and ^{15}N NMR chemical shift perturbation of

© 2019 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

N4BP1 in complex with ubiquitin, we demonstrate that the UBD of N4BP1 lacks the NEDD8 binding properties observed in CUBAN and it recognizes the Ile44-centered patch of ubiquitin through a dedicated binding site, which share some of the features observed in the CUE domain family. Moreover, we show that, in addition to mediating the interaction with ubiquitin and ubiquitinated substrates, both the CUBAN and the UBD of N4BP1 are poly-ubiquitinated in cells. This modification is dependent on the presence of a functional domain. We believe that the structural and functional characterization of this novel UBD will allow a deeper understanding of the molecular mechanisms governing N4BP1 function, while at the same time providing a valuable tool for clarifying how the discrimination between ubiquitin and the highly related NEDD8 is achieved.

Introduction

Ubiquitin (Ub) and ubiquitin-like (Ubl) proteins are small polypeptides that are covalently conjugated to various target proteins leading to degradation or changes in their localization and/or activity [1]. In order to correctly interpret these modifications, cell requires Ub/Ubl adaptors called “ubiquitin receptors”, which contain one or more ubiquitin binding domains (UBDs) that transduce the post-translation modification in the appropriate cellular response [2-4]. The ubiquitin molecule consists of a globular core called the β-grasp fold, comprising five-stranded mixed β-sheet and a α-helix, from which a flexible carboxyl-terminal tail protrudes. Most of the ubiquitin binding domains identified to date recognize a hydrophobic surface on ubiquitin characterized by the Ile44-centered patch, including residues Leu8, Ile44, His68 and Val70. Residues surrounding this patch allow for a specific modality of recognition that is peculiar of each domain family, thus explaining how different domains have evolved towards the recognition of a common hydrophobic contact site.

Neural precursor cell expressed developmentally downregulated protein 8 (NEDD8) is the closest relative to ubiquitin (58% identity and 80% similarity). Both molecules present a similar asymmetric surface distribution of charges, with a predominantly acidic face opposed to a hydrophobic one. Notably, the Ile44-centered patch is identical in the ubiquitin and NEDD8 molecules [5]. This high degree of conservation in the hydrophobic surfaces provides the explanation of why NEDD8 is usually involved in the recognition by several UBDs, with affinities that are similar to those reported for the interactions of such domains with ubiquitin. On the other hand, it also gives rise to the risk of misleading interpretations of the experimental results. Therefore, investigating how protein domains can discriminate between NEDD8 and ubiquitin is a challenge that allows tracing the most accurate picture of the biological properties and the cellular functions in which these highly similar proteins are involved.

NEDD4 binding protein 1 (N4BP1) was originally identified as a target for the proteasomal degradation by the HECT E3-ligase Nedd4 [6,7]. It has been suggested that N4BP1 participates in nuclear RNA processing processes and that the Nedd4-mediated monoubiquitination of N4BP1 is required for this function [8]. Moreover, the protein was reported to localize in the nucleolus where it undergoes sumoylation, polyubiquitination and proteasomal turnover at promyelocytic leukaemia nuclear bodies [9].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

By screening protein arrays with the human ubiquitin precursor UbC, Fenner and collaborators identified novel polyubiquitin binding proteins potentially regulating the NF-kappaB activity. Among them, the region of N4BP1, including amino acids 209-896, was identified in the screening and the interaction was experimentally confirmed [10]. By panning a human brain phage-displayed cDNA library using ubiquitin as bait, we identified a shorter fragment of N4BP1 (aa 813-896) that shows ubiquitin binding features and we further reduced the functional region to the last 50 amino acids of human N4BP1 (847-896aa) by means of biochemical approaches. Based on mutational analysis, we have shown that the interaction between N4BP1 and ubiquitin involves the Ile44-hydrophobic patch [11], in analogy with what has been observed in the majority of the UBDs characterized so far.

We present here the structure of the novel ubiquitin-binding domain identified in N4BP1, both alone and in complex with ubiquitin. Structural analysis performed by NMR spectroscopy revealed a novel three alpha-helix bundle domain. Using biochemical assays, we demonstrate that the interacting surface involves the hydrophobic patch in ubiquitin and a key motif in N4BP1 resembling the Phe-Pro (FP) motif typified in the Coupling of ubiquitin conjugation to ER degradation (CUE) domain family. Nevertheless, the structural analysis shows that key differences in the ubiquitin-binding domain of N4BP1 do not support the inclusion of this novel domain neither in the CUE nor in the UBA domain families. Conversely, the UBD of N4BP1 shares common features with the Cullin binding domain associating with NEDD8 (CUBAN), recently identified in the evolutionary related KHNYN protein [11] (*Santonico et al., 2019, submitted*). The observed similarity, both at the primary sequence and structural level, supports the identification of a novel protein domain family having the CUBAN domain and the UBD of N4BP1 as unique members. Interestingly, we show here that the UBD of N4BP1 has a clear preference for ubiquitin with respect to NEDD8. Conversely, the CUBAN domain has acquired the capability to recognize NEDD8 using a specific binding surface that is not present in N4BP1.

Results

A novel ubiquitin-binding domain at the C-terminal end of human N4BP1

Fenner and collaborators recently identified N4BP1 as a potential polyubiquitin chains binding protein. A previous analysis, performed with the phage display approach, enabled us to identify the region spanning residues 813-896 in N4BP1 as the one showing ubiquitin binding features (Figure 1A) [11]. The bioinformatics analysis of this region, performed with Simple Modular Architecture Research Tool (SMART) database, reported the presence of a CUE domain whose primary sequence, however, is too divergent from known CUE domain sequences to be reliably considered a member of this group (Figure 1B) [11]. CUE domains are structurally closely related to the UBA domains. Indeed, they share the three-helix bundle structure and both bind the Ile44 patch of ubiquitin via conserved hydrophobic residues mapping at the C-terminus of the helix-1. Two main features characterize CUE domains: a Phenylalanine-Proline (FP) sequence and a φxx(I/L/V)L motif, where φ is a large hydrophobic residue [20]. The first hydrophobic motif,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

mapping at the end of helix-1, is essential for ubiquitin binding and indeed it directly contacts ubiquitin residues Ile-44, His-68, Leu-69, and Val-70. Like the FP motif, also the C-terminal non-polar φxx(I/L/V)L motif binds the hydrophobic path surrounding Ile-44 of ubiquitin and it is involved in domain stability [21,22].

Interestingly, the minimal ubiquitin-binding region in N4BP1, including residues 849-896, shares 40% identity with the evolutionary related CUBAN domain, recently identified in the evolutionary related KHNYN protein [11]. The CUBAN has been shown to bind NEDD8 with an affinity that is higher than the value measured for monomeric ubiquitin [11]. In addition to interacting with neddylated cullins thanks to the direct recognition of the NEDD8 molecule, the CUBAN domain binds ubiquitin chains and poly-ubiquitinated proteins. This dual capability is due to the presence of two distinct binding sites that are in close proximity, as evidenced by the observation that the interaction of the CUBAN domain with ubiquitin chains competes out NEDD8 binding [11].

In order to evaluate the functional relevance of the FP motif in N4BP1, we inserted the mutation Pro866Ala (PA mutant) in the amino acid fragment spanning residues 813-896 of N4BP1 and we compared the binding specificities of the two constructs by pull-down assay. GFP-N4BP1 and the PA mutant were independently expressed in 293Hek cells and the cell extracts were incubated with beads coated with GST, GST-ubiquitin or GST-NEDD8 (Figure 2A). As shown, GFP-N4BP1 but not the PA mutant, is efficiently recovered by ubiquitin. Moreover, in keeping with our previous report [11], N4BP1 clearly binds ubiquitin while only a faint signal can be detected in the GST-NEDD8 pull-down. Consistent with the previous result, the GST fusion of N4BP1 bound ubiquitinated proteins from a 293Hek cell extract and the interaction was abrogated by the Pro866Ala mutation (Figure 2B). These results demonstrate that the FP motif of N4BP1 acts similarly to the corresponding key motif in the CUE domains. On the other hand, sequence similarities indicate that this domain is structurally similar to the CUBAN, although significantly different in the binding preferences. Nevertheless, the overall divergence highlighted by the SMART search prompts us to speculate that the ubiquitin binding domain of N4BP1 has acquired functional characteristics that are common to the CUE domain family, despite not being evolutionary related.

NMR spectroscopy study of N4BP1 ubiquitin binding domain.

To gain a more accurate description of the N4BP1 ubiquitin binding domain, the NMR spectroscopy study of the C-terminal end (residues 847-896) using the 2D and 3D NMR methodology was performed, leading to the assignment of 47 out of 50 amino acids in the ^{15}N labelled protein. The ^{15}N HSQC spectrum is reported in Figure 3A. The spectral dispersion was sufficient to achieve the assignment of the resonances obtained by 3D ^{15}N –TOCSY and 3D ^{15}N –NOESY spectra. The cross peaks were labelled and the assignments were further confirmed by ^{1}H-^{1}H TOCSY and NOESY spectra, obtained with different mixing times as reported in Materials and Methods (Figure 3B). As shown, the NOEs number in the fingerprint region of the ^{1}H-^{1}H NOESY spectrum of N4BP1 is not so large as expected for a fully structured protein of this size (Figure 3C), thus indicating the presence of an extended random coil region in the secondary structure. The analysis of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

^{15}N chemical shifts of N4BP1 performed with the algorithms of Wishart & Sykes [23] led to the identification of the secondary structure elements of N4BP1, consisting in three regions (α1, α2 and α3), whose chemical shift deviations are typical of a helix conformation. Only a single value characteristic of a beta secondary structure without any persistence of this chemical shift deviation has been identified, thus excluding the presence of secondary structures other than α-helices *(not shown)*. The complete assignment of resonances was performed and reported in Table 1. Based on the NOEs, the structure of the N4BP1 C-terminal end was calculated using the conventional XPLOR [24]. Simulated annealing protocol in vacuo [25], inter-protons distances and torsion angles (φ and ψ) extracted from NMR data (NOEs and Chemical Shift Index) were used as restraints during the molecular dynamics simulations. Beyond the three helices, a large part of the protein appears in a random coil conformation indicating a pronounced structural flexibility, also confirmed by the limited number of NOEs that were detected, thus recommending the structure to be interpreted with some caution. Figure 3D reports the average of a cluster of 10 structures chosen by the minimum energy level out of the 200 generated by XPLOR together with the primary sequence of the domain. The domain conformation scheme recalls the structural features of the largest group of ubiquitin-binding domains, represented by the three-helix bundle fold [2]. Notably, the long loop connecting helices α1 and α2 presents a high level of flexibility, a feature resembling the extended loop1 of the evolutionary related CUBAN domain [11].

Conformational studies of N4BP1 in complex with either ubiquitin or NEDD8 performed by circular dichroism

To gain more information on changes in the secondary structure of the N4BP1 UBD upon binding, the CD spectra was obtained in the region from 195 to 260nm, as reported in Figure 4A. By applying different optimization and deconvolution techniques, we obtained an estimate of secondary structure content that resulted to be about 46% α-helix and 3% β-sheet (or β-turn) when using the K2D3 method [26], a value confirmed also by using DICHROWEB [27] with a deconvolution optimized by the fitting routine CONTIN (~45% α-helix content) [28]. In both cases, the estimate was in agreement with the results obtained by NMR. Indeed, the helix content found in the NMR structure and deduced by the chemical shift index, involves the 47% of the sequence, particularly the tracts 6-12, 23-32 and 37-45. Finally, the NMR result is also supported by the Agadir prediction (http://agadir.crg.es) [29] which gives an estimate of 45% total helix content, corresponding to the tracts 6-18, 23-31 and 43-47.

We then used the CD spectroscopy to report changes in the mean residue ellipticity related to the secondary structure elements of the different proteins in the N4BP1/ubiquitin complex and we compared them with the CD spectra obtained for the N4BP1/NEDD8 complex. Figure 4A shows the CD spectra of N4BP1 and ubiquitin, both alone and in presence of an equimolar amount of the binding partner (experimental details are reported in Materials and methods). We also measured the spectral differences between each component in the mixture and the single, isolated component. As shown, the point by point differences between the CD spectrum of N4BP1 and ubiquitin in the mixture and their CD spectra when measured alone, reveal that

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

small conformational changes, although different in shape and intensity, occurred in the CD spectra of both partners. The small values found in both dichroic profiles are in line with partial adjustments of the conformation upon binding between the two proteins, thus indicating that the mutual perturbation occurred due to an intermolecular interaction. The observed tiny differences were amplified and reported in Figure 4B. A similar experiment, performed with NEDD8 in place of ubiquitin, showed that, within the experimental error, traces in the differences of the CD spectra are close to zero for both proteins (Figure 4C), thus confirming that no interaction occurred between N4BP1 and NEDD8, in agreement with previous observations.

^{15}N HSQC NMR investigation of the N4BP1/ubiquitin complex by chemical shift perturbation.

We then investigated the N4BP1/Ubiquitin complex by monitoring the chemical shift changes(s) of ^{15}N labelled N4BP1 upon addition of unlabelled ubiquitin, at different molar ratio. The typical CSP plot is reported in Figure 5A (1:0.5) and 5B (1:1). As shown, residues that are primarily involved in the interaction between ^{15}N labelled N4BP1 and ubiquitin are A14, L15, I18, S22, I28, Q30 and D40 at both molar ratios. Residues N9, K27 and H35 are only perturbed at 1:1 molar ratio. The complementary experiment performed using ^{15}N labelled ubiquitin and the unlabelled N4BP1 showed that a large number of residues are perturbed following the addition of the unlabelled protein at both 1:0.5 (Figure 6A) and 1:1 molar ratios (Figure 6B) of ^{15}N-labelled ubiquitin/unlabelled N4BP1. Fixing a threshold of 1σ (indicated by a line) a view of the residues involved in the interaction can be highlighted. As shown, the hydrophobic residues mapping in the β-sheet region of ubiquitin are the main determinants of the interaction, followed by a large number of polar and neutral residues that are also involved, thus indicating a complex recognition mechanism between the two proteins upon the early hydrophobic recognition step. In order to determine the regions involved in the interaction, CSPs detected by both experiments were analysed by the HADDOCK procedure [30]. The model obtained, reported in figure 7 (*upper panels*), shows that helix-1 and loop-1 in N4BP1 are both involved in the ubiquitin binding. On the other hand, ubiquitin recognizes the UBD of N4BP1 by means of residues mapping in the β-sheet secondary structure. Particularly, a detailed analysis of the hydrophobic, polar and neutral residues, whose chemical shifts were perturbed by the interaction, revealed that Leu8, His68 and Val70 in ubiquitin are all involved in the interaction with the UBD of N4BP1. These residues, together with Ile44, define the well-known interacting surface of the Ile44-patch, which is central to the majority of UBD–ubiquitin interactions characterized to date. Even though Ile44 lacks of a measurable CSP, the mutation Ile44Ala abrogates the interaction, thus definitely confirming that the hydrophobic patch of ubiquitin is the primary contact site. On the other hand, residues V48 and L49 at the end of helix-3 do not show any perturbation. It is important to remark that, beyond the hydrophobic contacts, many polar and neutral residues are also involved, thus confirming the steering role of electrostatic and/or dipolar interactions during the protein-protein recognition and until reaching the final conformation. Further investigation will allow us to better refine the binding requirements of this novel interaction.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

Finally, figure 8 shows the relevant CSP perturbations of K27 (at 1:1 molar ratio), I28 and Q30 (both at 1:0.5 and 1:1) in N4BP1, which are located quite far from the interacting surface (see also Figures 5A and B) and and therefore should not be affected by the interaction. A detailed observation of the complex structure may lead to an intriguing interpretation of this result. Indeed, given that CSPs are due to perturbations of the ^{15}N backbone chain caused by conformational changes of either the side chain or the secondary structure tract itself, the CSP observed for these residues can be considered as an indicator of magnetic environment perturbations of the ^{15}N of the amide bond, due to a secondary structure perturbation in that tract. Based on this premise, a possible explanation for the relevant perturbations of K27, I28 and Q30 upon binding with ubiquitin could be that the tract of helix-2 containing these residues is subjected to helical geometry distorting forces, generated by the pulling down effects of the two flanking loops 1 and 2 toward the ubiquitin interacting surface.

Differences and similarities between the UBD of N4BP1 and the CUBAN domain of KYNYN.

CUE domains are closely related to the UBA. Indeed, they share the three-helix bundle structure and both interact with the Ile44 patch of ubiquitin via conserved hydrophobic residues at the C-terminus of the helices α1 and α3. As previously described, the Ile44 patch contact sites in CUE domains are represented by the FP pair and the hydrophobic sequence φxx(I/L/V)L [21,22]. Figure 9A shows the alignment of the UBD of N4BP1 with the evolutionary related CUBAN domain of hKHNYN, the CUE domains of human AMFR, AUP1 and Tollip and the yeast CUE domain of Vps9p. As shown, the sequence φxx(I/L/V)L is represented by the motif LSxx(L/V)L, with a Ser following the hydrophobic leucine in both proteins. Moreover, the FP motif in KHNYN is substituted by an FW pair. We therefore inspected the spatial distribution of Phe-Pro side chains respect to the hydrophobic cluster respectively in the UBD of N4BP1, the CUBAN domain of KHNYN and the typical CUE of hAMFR (Figure 9B, C, D). Differences and similarities are clearly visible: while the FP motif is positioned at the C-terminal end of helix-1 in the CUE domain of AMFR, both the FP and FW pairs in N4BP1 and KHNYN occupy a central position inside the extended loop1. Nevertheless, the Trp647 points towards the core of the domain and constitutes, together with flanking residues, a sort of pivot that reduces the flexibility of loop1 in the CUBAN domain of KHNYN [11]. On the contrary, the FP motif of N4BP1 is highly exposed to the solvent, at least in the absence of the ligand. Given the high flexibility of this region, we expect that its spatial position would be significantly affected in the protein complex. In keeping with this observation, the FP pair is directly involved in the interaction with ubiquitin (see Figures 2 and [11].

We therefore attempted to clarify whether the FW and the FP motifs, respectively in CUBAN and N4BP1 carboxyl-terminal domains, are functionally similar. To this end, we compared by pull-down assay the effect of the mutations Trp648Ala and Pro866Ala in the interaction with endogenous neddylated and poly-ubiquitinated proteins. The GST fusions of KHNYN and N4BP1 C-terminal ends, both wild-type and mutated, were incubated with a cell extract from T-Rex Flag-NEDD8 cell line, which expresses low levels of flag-tagged NEDD8, under the control of a tetracycline inducible promoter [11]. As shown in Figure 10A,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

while the mutation Pro866Ala in N4BP1 (P866A) abrogates the interaction with poly-ubiquitinated proteins, the corresponding KHNYN mutant (W647A) retains the capability to interact with both neddylated CUL1 and CUL2 as well as with ubiquitinated substrates, despite showing a significant reduction in the binding efficiency towards monomeric NEDD8. This effect is probably the result of conformational variations in the mutated domain, which mainly affect the interaction having the weaker binding affinity. Therefore, while the FW pair in KHNYN shows structure stabilizing features rather then being involved in protein-protein interactions, the FP motif in N4BP1 directly participates in the interaction of the UBD with ubiquitin.

Finally, it has been shown by other that the ubiquitin binding domains UIM and CUE, apart from mediating the interaction with ubiquitinated partners, are required for intra-molecular ubiquitination within the same protein [20,31]. We therefore set out to find if both the CUBAN domain and the UBD of N4BP1 would share similar properties. To this end, we transiently expressed both domains as GFP fusions in 293Ph cells and we analysed the anti-GFP immunoprecipitates with anti-GFP antibody (Figure 10B). As shown, both domains are subjected to several post-translational modifications, compatible with mono- and multiple ubiquitinations. The result indicates that both the CUBAN domain and the UBD identified in N4BP1 mediate not only the interaction with ubiquitinated proteins, but also covalent modifications that are aimed at regulating the functional properties of the proteins themselves. We therefore investigated whether the FP and FW motifs were also involved in the post-translation modifications of the respective protein domains. To this end, we transfected 293Ph cells with plasmids coding for the mutants $KHNYN_{Ct}$-W647A and $N4BP1_{Ct}$-P866A fused to the GFP, and we performed the immunoprecipitation with anti-GFP antibody. The immunoprecipitated samples were analysed by immunoblotting with anti-ubiquitin antibody (Figure 10C,D). The results confirmed that both domains are ubiquitinated when transiently transfected in cells. Whilst the mutation in the FW motif of KHNYN does not show any effect on protein ubiquitination, the inactivation of the FP motif in N4BP1 abrogates the post-translational modifications of the C-terminal end (Figure 10D). Taken together, these results clearly demonstrate that, despite the two evolutionary related domains share the capability to interact with poly-ubiquitin and mediate the multi-ubiquitination of the proteins themselves, they must substantially differ in the way these recognitions takes place.

Discussion

Sixteen different ubiquitin-binding domains have been identified so far. They can be grouped in different classes depending primarily on the specific structural features. The vast majority belongs to the group of three-alpha helical bundle domains, all sharing the common features of a similar structural fold and the recognition of the Ile44-centered hydrophobic patch in the ubiquitin molecule. Recently, a novel ubiquitin-binding domain has been identified at the C-terminal end of KHNYN (also called KIAA0323) protein. This domain, called CUBAN for Cullin Binding domain Associating with NEDD8, shows a clear preference for monomeric NEDD8 and mediates the interaction of KHNYN with neddylated cullins [11]. Although the CUBAN also binds ubiquitin, primarily in the multimeric form, it has been shown that the recognition of ubiquitin and NEDD8 involves different binding sites, thus supporting the notion that the CUBAN domain

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

has evolutionary gained the capability to discriminate between the two highly related molecules. N4BP1 (KIAA0615), a close relative of KHNYN, also contains an ubiquitin-binding domain at the C-terminal end of its amino acid chain (853-896aa) that shares 40% identity and 47% similarity with CUBAN [11]. The inspection of the carboxyl-terminal ends of KHNYN and N4BP1 structures, both obtained by NMR, revealed, respectively, a moderate and a pronounced flexibility, which makes difficult a rigorous and comprehensive comparison between the two conformational models. Nevertheless, the 3D conformation of the N4BP1 UBD domain is more similar to CUBAN with respect to CUE and UBA domains. Despite the high level of similarity, these domains show significant differences in their binding preferences. The circular dichroism spectroscopy confirmed that the interaction of N4BP1 with ubiquitin occurs with slight mutual changes in the dichroic profiles of both molecules, in line with the structural stability of the ubiquitin molecule, characterized by β-strands flanked one another with a network of hydrogen bonds. Moreover, the same experiment performed with the close relative NEDD8 did not report any variation in the dichroic profiles, thus confirming that N4BP1 does not bind to NEDD8.

The interaction study performed by NMR spectroscopy revealed that a hydrophobic region in N4BP1 contacts the hydrophobic patch located in the β-region of ubiquitin. This observation is in line with the molecular details characterizing other ubiquitin binding domains showing a similar structure, such as UBA and CUE which, although with several different micro-environmental details depending on their local structural characteristics, contact the Ile44-hydrophobic patch of ubiquitin [2]. The residue Ile44 is mostly buried at the intermolecular contact area of ubiquitin, thus giving an explanation of the absence of a measurable CSP of the amide ^{15}N of the backbone for this residue in our NMR binding experiments. Nevertheless, as expected, the mutational analysis confirmed a clear involvement of this residue in the interaction with N4BP1 (Figure 2 and [11]. Taken together, our data definitely demonstrate that the primary contact site of the N4BP1 UBD is represented by the hydrophobic patch of ubiquitin, consisting of Leu8, Ile44, Val70 side chains flanking His68. Further steering contacts take place in this interaction, involving polar and neutral residues, that electrostatically or dipolarly strengthen the recognition and also appear to play a role in the protein-protein selectivity, in agreement with the observation that different UBDs contact the hydrophobic patch of ubiquitin with modalities that show peculiar differences for each binding domain [2].

The analysis of the residues that differ between ubiquitin and NEDD8, but whose identity is conserved between the various NEDD8 orthologs, has led to the identification of two distinct patterns, constituted by “conserved/divergent” residues and aligned along opposite faces, which are expected to mediate NEDD8-specific interactions [5] (Figure 11A). One of these clusters includes residues F4, K12, E64 in ubiquitin and K4, E12, G64 in NEDD8. The strongly different nature of these patterns suggests that they can deeply influence the protein recognition. The space-filling representation shows the interaction of the FP pair of N4BP1 with ubiquitin and highlights the possible role of the molecular surface including positions 4, 12 and 64, together with the residue in position 14, in the discrimination between the two molecules (Figure 11B). In fact, the inspection of this molecular surface in the N4BP1/ubiquitin complex shows that this cluster is in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

close proximity with the FP motif, suggesting that, even though not directly involved in the interaction with ubiquitin (as predictable by the lack of CSPs for these residues), this molecular surface could be involved in maintaining the complex stability. Particularly, the electronic cloud of the aromatic ring of F865 may interact by stacking with the aromatic cloud of the ring of F4 of ubiquitin, flanked by 12T, 14T and E64, with the latter being the unique polar group in the cluster. On the other hand, the corresponding pattern in NEDD8 offers a charged surface characterized by the polar residues 4K, 12E and 14E, while lacks any aromatic electronic cloud that could interact with the FP pair (Figure 11C). In summary, we can suggest that the exposed side chains in this region represent a peculiar distribution of hindered (threonines), aromatic (phenylalanine), polar charges (lysine and glutamic acid) and flexibility sources (glycine) which are able to strongly influence, in different ways, the protein recognition with a determined substrate. Intriguingly, residues 4, 12, 14 and 64 have been recently shown to be molecular determinants for distinguishing between ubiquitin and NEDD8 by Ubiquitin-specific peptidase 2 (USP2) [32], thus confirming their key role in the discrimination mechanism.

While no data are available for KHNYN, to date most of the information regarding N4BP1 function is related to the activity of the HECT ligases Itch and Nedd4. Indeed, Nedd4 has been shown to post-translationally regulate the intracellular localization of N4BP1 in the nucleolus of primary cells and at promyelocytic leukaemia nuclear bodies. The interaction with Itch, on the contrary, does not seem to be followed by N4BP1 ubiquitination, but rather N4BP1 counteracts the activity of the Itch E3-ligase by directly competing with Itch substrates, such as c-Jun and the members of the p53 family, p63 and p73, thus suggesting a role as promoter of tumour progression [7]. Accordingly, by interacting with the de-ubiquitinating enzyme Cezanne, N4BP1 acts also as a potent negative regulator of NF- B signalling in neuroblastoma cells, a role that points to N4BP1 as a candidate target for immunotherapy [33].

N4BP1 is subjected to several post-translational modifications, such as ubiquitination and sumoylation. While the latter affects protein localization, the biological role of N4BP1 poly-ubiquitination is still an unexplored field. We show here that the UBD of N4BP1, when transiently expressed in cells, promotes its own poly-ubiquitination and that an FP motif is required for both covalent and non-covalent interactions with ubiquitin. Analogously, we have previously shown that full-length KHNYN is polyubiquitinated in cells and that this modification is abrogated following the deletion of the CUBAN domain [11]. In addition, we have shown here that the CUBAN domain itself is ubiquitinated when transiently expressed in cells. Intriguingly, many proteins containing UBD domains undergo mono-ubiquitination in a process, called coupled mono-ubiquitination, that requires the interaction between a functional UBD and an E3-ubiquitin ligase [31,34]. Thanks to the presence of the CUBAN domain, which binds to both free and Cullin-conjugated NEDD8, KHNYN can recruit distinct multi-subunit E3 ubiquitin ligase complexes in their active form, an interaction that could in turn promotes KHNYN ubiquitination. We expect that N4BP1, which does not share this feature, be mostly regulated by the interaction with HECT-ligases, as also suggested by published data. Interestingly, several examples of coupled mono-ubiquitination are found in the components of intracellular trafficking machineries. KHNYN and N4BP1 could therefore represent novel examples of proteins,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

regulated by coupled-monoubiquitination involving a new class of ubiquitin binding domain, whose functions, still far from being exhaustively characterized, seem to be primarily related with RNA metabolism processes [8].

Concluding, we believe that our data will help in shedding light on the molecular details governing the recognition of ubiquitin and NEDD8 and how the discrimination between these two highly related molecules is achieved. Moreover, they reveal a complex web of functional interactions, linking distinct classes of ubiquitin ligases, ubiquitin binding proteins and proteins involved in RNA metabolism, that future research will help to unveil.

Materials and methods

DNA constructs and site-directed mutagenesis. The cDNA constructs coding for the C-terminal ends of hN4BP1 (813-896) and hKHNYN (598-678) were previously described. The minimal ubiquitin binding region of N4BP1, including amino acids 845-896, was cloned in pGex2TK vector using primers R2326 (forward) CAAGGATCCGCTCAGAGATCTTCTGCA and R2473 (reverse) CTCTCGAGTCAATCCAACACCATGGCAG. The T-Rex flag-NEDD8 cell line was generated as previously described [11]. Site-directed mutagenesis of CUBAN W647A and N4BP1 UBD P866A were performed using the QuikChange® Site-Directed Mutagenesis Kit (Invitrogen, Frederick, MD, USA) according to the manufacturer's instructions.

Protein expression and purification. The C-terminal fragment of hN4BP1 (aa 845-896) was cloned in pGex2T vector and the recombinant protein prepared as previously described [12]. The bacterial expression was performed according to the protocol for NMR sample preparation described by Weber and collaborators [13].

Cell extracts, immunoprecipitation and in vivo ubiquitination assay. Whole-cell extracts were prepared from mammalian cells by lysis in 25mM Tris pH 7,5, 125mM NaCl, 1% glycerol, 1mM $MgCl_2$, 1mM orthovanadate, 5mM NaF, 0,5% NP-40, 0,5% Triton X- 100, complete proteases inhibitor cocktail (Sigma), complete phosphatase inhibitor (Sigma) 30 min on ice, before clarification by centrifugation. To detect protein in cell lysates, protein samples were separated by SDS-PAGE and transferred onto nitrocellulose membrane. Proteins were detected by immunoblotting and visualized by treating the blots with ECL (Millipore). For the pull-down experiments, equimolar amounts of GST-fusion proteins corresponding to the selected clones and GST alone, linked to glutathione-Sepharose 4B beads, were incubated 1,5 hrs with 1 mg of cell extract. Beads were washed three times in washing buffer (25mM Tris pH 7,5, 125mM NaCl, 1% glycerol, 1mM $MgCl_2$, 1mM orthovanadate, 5mM NaF, 0,5% NP-40, proteases inhibitor cocktail, phosphatase inhibitor cocktail) and bound proteins were electrophoresed on polyacrylamide gel, transferred to PVDF or nitrocellulose membranes, and detected with specific antibodies. For the coimmunoprecipitation

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

assay cells were plated in 100-mm dishes and grown to 60–70% confluence in culture medium supplemented with 10% FBS. Twenty hours post-transfection cells were solubilized in lysis buffer. An equal amount of each protein lysate was incubated with anti-Flag M2 affinity gel beads for 2 hours at 4°C and then washed three times with washing buffer supplemented with fresh inhibitors. The protein lysates and the immune complexes were analysed by Western blot analyses with rabbit anti-GFP and anti-ubiquitin antibodies.

Antibodies and reagents. Polyclonal anti-flag was from Sigma (St Louis, MO, USA), anti-tubulin and anti-ubiquitin P4D1 from Santa Cruz Biotechnology (Santa Cruz, CA, USA), anti-NEDD8 was from Abcam (Boston, MA, USA), antirabbit and anti-mouse peroxidase-conjugated were from Jackson Immunoresearch (West Grove, PA, USA).

Cell culture. Hek293 and Flp-In T-Rex-293-Flag-NEDD8 were grown on DMEM (Dulbecco's modified Eagle's medium; GIBCO) supplemented with 10% FBS (fetal bovine serum; GIBCO) and 100 units/ml penicillin, 100 µg/ml streptomycin (Invitrogen). Stably transfected Flp-In T-Rex-293-Flag-NEDD8 cells were generated as [11]. The cells were stimulated with 1 µg/ml tetracycline overnight to induce expression of the desired fusion proteins.

NMR spectroscopy. The experimental conditions were similar to those previously reported [11]. 1H and 15N NMR spectra were obtained at 298 K on a Bruker Advance instrument operating at 700.132 MHz. 2D and 3D experiments (15N –TOCSY and 3D 15N –NOESY spectra with mixing times TOCSY from 0.040 to 0.080s and for NOESY from 0.080 to 0.3s) [14,15] (Braunschweiler and Ernst, 1983; Wüthrich, 1986) were necessary to obtain the spin systems assignments and, subsequently, the sequential assignments. In order to study the intermolecular interactions, the chemical shift perturbation (CSP) method was applied by the 15N−1H HSQC NMR. Spectra of the 15N labelled N4BP1 domain alone and in presence of NEDD8 or ubiquitin at different molar ratios were run [16]. This pulse sequence was modified with a gradient water suppression [17]. To monitor the changes in ^{15}N labelled amide groups in the different conditions, TOPSPIN 3.1 and NMRView software packages were used for data processing. The analysis of the CSPs was performed as described in references [18,19].

Circular Dichroism. The circular dichroism spectroscopy (CD) measurements were performed on a Jabsco CD spectrophotometer using a 0.1 cm path length cell cuvette. The concentration of N4BP1 was 50 µM in H_2O. The CD spectra were obtained in the far UV spectral region from 195 to 260 nm, using a step resolution of 0.2 nm and a speed scan for spectra collection of 20 nm/min. Spectra were obtained are the average of four scans at room temperature. The estimate of secondary structure elements content was obtained using the web server for CD data prediction as described in the text.

Interaction of N4BP1 in complex with ubiquitin and NEDD8 performed by Circular Dichroism CD. CD studies were performed at a protein concentration of 50µM and adopting the scheme described below. In the

first step, the CD spectra of N4BP1 and ubiquitin alone were obtained and expressed as rotatory power. Then, the CD spectrum of N4BP1 in the presence of an equimolar amount of ubiquitin (or NEDD8 for comparison as reported below) was performed. In order to identify the possible conformational changes due to protein-protein interactions, the CD spectra, expressed as rotator power, obtained with the components alone, were subtracted point-by-point from the CD spectrum of the protein complex. Specifically, from the mixture N4BP1+ubiquitin, the first subtraction (mixture-ubiquitin) allows the evaluation of the conformational changes induced on N4BP1 in complex with ubiquitin, respect to N4BP1 alone. Subsequently, by subtracting the CD spectrum of N4BP1 alone from the one of the mixture N4BP1-ubiquitin (mixture-N4BP1), we obtained the conformational changes induced on ubiquitin in complex with N4BP1. Finally, in order to better detect the tiny changes in the CD profiles, the traces of the obtained differences transformed in mean residue ellipticity were vertically amplified. An identical procedure was applied to the study of the interaction of N4BP1 upon binding with NEDD8. In all these experiments great attention was devoted to maintaining the same protein concentration, buffering system and ionic strength.

Acknowledgments: We thank COST Action PROTEOSTASIS BM1307 members for supporting this work and Simona Polo for technical support. The work in the group of GC was supported by the Italian Association for Cancer Research (AIRC) project IG 2013 N14135.

Authors contributions: Conceptualization, M.P. and E.S.; Data curation, R.N., F.F., W.M., A.M., M.P. and E.S.; Formal analysis, L.C, M.P. E.S.; Funding acquisition, G.C: and M.P.; Investigation, R.N., F.F., W.M., A.M., M.P. and E.S.; Methodology, F.F., W.M., A.M. and E.S.; Supervision, M.P. and E.S.; Visualization, M.P. and E.S.; Writing – original draft, Elena Santonico; Writing – original draft, M.P. and E.S.; Writing – review & editing, L.C., G.C., M.P. and E.S.All the authors approved the final version of the manuscript.

Conflict of interest: The authors declare that they have no conflicts of interest.

References

[1] K.N. Swatek, D. Komander, Ubiquitin modifications, Nature Publishing Group. 26 (2016) 399–422. doi:10.1038/cr.2016.39.

[2] J.H. Hurley, S. Lee, G. Prag, Ubiquitin-binding domains, Biochem. J. 399 (2006) 361–372. doi:10.1042/BJ20061138.

[3] K. Husnjak, I. Dikic, Ubiquitin-Binding Proteins: Decoders of Ubiquitin-Mediated Cellular Functions, Annu. Rev. Biochem. 81 (2012) 291–322. doi:10.1146/annurev-biochem-051810-094654.

[4] K. Sokratous, A. Hadjisavvas, E.P. Diamandis, K. Kyriacou, The role of ubiquitin-binding domains in human pathophysiology, Critical Reviews in Clinical Laboratory Sciences. 51 (2014) 280–290. doi:10.3109/10408363.2014.915287.

[5] F.G. Whitby, G. Xia, C.M. Pickart, C.P. Hill, Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes, J. Biol. Chem. 273 (1998) 34983–34991.

[6] R. Murillas, K.S. Simms, S. Hatakeyama, A.M. Weissman, M.R. Kuehn, Identification of developmentally expressed proteins that functionally interact with Nedd4 ubiquitin ligase, J. Biol.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

Chem. 277 (2002) 2897–2907. doi:10.1074/jbc.M110047200.
[7] A. Oberst, M. Malatesta, R.I. Aqeilan, M. Rossi, P. Salomoni, R. Murillas, et al., The Nedd4-binding partner 1 (N4BP1) protein is an inhibitor of the E3 ligase Itch, Proc. Natl. Acad. Sci. U.S.a. 104 (2007) 11280–11285. doi:10.1073/pnas.0701773104.
[8] V. Anantharaman, L. Aravind, The NYN domains: novel predicted RNAses with a PIN domain-like fold, RNA Biol. 3 (2006) 18–27.
[9] P. Sharma, R. Murillas, H. Zhang, M.R. Kuehn, N4BP1 is a newly identified nucleolar protein that undergoes SUMO-regulated polyubiquitylation and proteasomal turnover at promyelocytic leukemia nuclear bodies, J. Cell. Sci. 123 (2010) 1227–1234. doi:10.1242/jcs.060160.
[10] B.J. Fenner, M. Scannell, J.H.M. Prehn, Identification of polyubiquitin binding proteins involved in NF-κB signaling using protein arrays, BBA - Proteins and Proteomics. 1794 (2009) 1010–1016. doi:10.1016/j.bbapap.2009.02.013.
[11] L. Castagnoli, W. Mandaliti, R. Nepravishta, E. Valentini, A. Mattioni, R. Procopio, et al., Selectivity of the CUBAN domain in the recognition of ubiquitin and NEDD8, Febs J. 98 (2019) 1–37. doi:10.1111/febs.14752.
[12] E. Santonico, S. Panni, M. Falconi, L. Castagnoli, G. Cesareni, Binding to DPF-motif by the POB1 EH domain is responsible for POB1-Eps15 interaction, BMC Biochem. 8 (2007) 29. doi:10.1186/1471-2091-8-29.
[13] D.J. Weber, A.G. Gittis, G.P. Mullen, C. Abeygunawardana, E.E. Lattman, A.S. Mildvan, NMR docking of a substrate into the X-ray structure of staphylococcal nuclease, Proteins. 13 (1992) 275–287. doi:10.1002/prot.340130402.
[14] A. Bax, D. Davis, MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy, (1985) 1–6.
[15] D. Marion, K. Wüthrich, Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of 1H-1H spin-spin coupling constants in proteins, Biochem. Biophys. Res. Commun. 113 (1983) 967–974.
[16] G. Bodenhausen, J.D. Riben, Natural abundance Nitrogen-15 NMR by enhanced heteronuclear spectroscopy, (2001) 1–5.
[17] B.D. Nguyen, X. Meng, K.J. Donovan, A.J. Shaka, SOGGY: solvent-optimized double gradient spectroscopy for water suppression. A comparison with some existing techniques, J. Magn. Reson. 184 (2007) 263–274. doi:10.1016/j.jmr.2006.10.014.
[18] E.R.P. Zuiderweg, Mapping protein-protein interactions in solution by NMR spectroscopy, Biochemistry. 41 (2002) 1–7.
[19] M.P. Williamson, Using chemical shift perturbation to characterise ligand binding, Prog Nucl Magn Reson Spectrosc. 73 (2013) 1–16. doi:10.1016/j.pnmrs.2013.02.001.
[20] S.C. Shih, G. Prag, S.A. Francis, M.A. Sutanto, J.H. Hurley, L. Hicke, A ubiquitin-binding motif required for intramolecular monoubiquitylation, the CUE domain, Embo J. 22 (2003) 1273–1281. doi:10.1093/emboj/cdg140.
[21] G. Prag, S. Misra, E.A. Jones, R. Ghirlando, B.A. Davies, B.F. Horazdovsky, et al., Mechanism of Ubiquitin Recognition by the CUE Domain of Vps9p, Cell. (2003) 1–12.
[22] R.S. Kang, C.M. Daniels, S.A. Francis, S.C. Shih, W.J. Salerno, L. Hicke, et al., Solution structure of a CUE-ubiquitin complex reveals a conserved mode of ubiquitin binding, Cell. 113 (2003) 621–630.
[23] D.S. Wishart, B.D. Sykes, The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data, J. Biomol. NMR. 4 (1994) 171–180.
[24] A.T. Brünger, Assessment of phase accuracy by cross validation: the free R value. Methods and applications, Acta Crystallogr. D Biol. Crystallogr. 49 (1993) 24–36. doi:10.1107/S0907444992007352.
[25] O. JG, P.V. Pedone, The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode, (2004) 1–11.
[26] C. Louis-Jeune, M.A. Andrade-Navarro, C. Perez-Iratxeta, Prediction of protein secondary structure from circular dichroism using theoretically derived spectra, Proteins. 80 (2012) 374–381. doi:10.1002/prot.23188.
[27] A. Lobley, L. Whitmore, B.A. Wallace, DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, Bioinformatics. 18 (2002) 211–212.
[28] N. Sreerama, R.W. Woody, Estimation of Protein Secondary Structure from Circular Dichroism

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set, Analytical Biochemistry. 287 (2000) 252–260. doi:10.1006/abio.2000.4880.
[29] V. Muñoz, L. Serrano, Local versus nonlocal interactions in protein folding and stability--an experimentalist's point of view, Fold Des. 1 (1996) R71–7. doi:10.1016/S1359-0278(96)00036-3.
[30] S.J. de Vries, M. van Dijk, A.M.J.J. Bonvin, The HADDOCK web server for data-driven biomolecular docking, Nat Protoc. 5 (2010) 883–897. doi:10.1038/nprot.2010.32.
[31] S. Polo, S. Sigismund, M. Faretta, M. Guidi, M.R. Capua, G. Bossi, et al., A single motif responsible for ubiquitin recognition and monoubiquitination in endocytic proteins, Nature. 416 (2002) 451–455. doi:10.1038/416451a.
[32] Y.-C. Shin, J.-H. Chen, S.-C. Chang, The molecular determinants for distinguishing between ubiquitin and NEDD8 by USP2, Sci Rep. 7 (2017) 159. doi:10.1038/s41598-017-02322-x.
[33] L. Spel, J. Nieuwenhuis, R. Haarsma, E. Stickel, O.B. Bleijerveld, M. Altelaar, et al., Nedd4 Binding Protein 1 (N4BP1) and TNFAIP3 Interacting Protein 1 (TNIP1) control MHC-1 display in neuroblastoma, Cancer Res. (2018) canres.0545.2018–21. doi:10.1158/0008-5472.CAN-18-0545.
[34] K. Haglund, H. Stenmark, Working out coupled monoubiquitination, Nat. Cell Biol. 8 (2006) 1218–1219. doi:10.1038/ncb1106-1218.
[35] G. Cornilescu, J. Marquardt, M. Ottiger, A. Bax, Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase, (1998) 1–2.

Figure Legends

Figure 1. The ubiquitin-binding domain of N4BP1 resembles a CUE domain. (A) Domain organization of human N4BP1 (UniProt code O75113). The sequence of the carboxyl-terminal end of N4BP1 identified by phage display and spanning residues 813-896 is shown. (B) Comparative sequence analysis of the UBD of N4BP1 human with CUE domain containing proteins, performed with SMART search using as query sequence the amino acid region underlined in (A). Coloured backgrounds highlight the FP pair with the invariant Proline (yellow) and the C-terminal hydrophobic motif (green).

Figure 2. A Pro-Phe motif in N4BP1 is involved in the interaction with ubiquitin and ubiquitinated proteins. (A) The amino acid fragment spanning residues 813-896 of N4BP1 and the Pro866Ala mutant (PA) were expressed as GFP fusions in Hek293 cells. Twenty hours post-transfection, cells were harvested and lysates were incubated with GST-Ubiquitin, GST-NEDD8 or GST alone. Samples were analysed by SDS-PAGE with anti-GFP antibody. (B) Cell extracts from Hek293 cells were incubated with the GST fusions of N4BP1 (813-896aa), N4BP1-PA, or with the GST alone. Samples were analysed by SDS-PAGE with ubiquitin antibody. Ponceau staining confirmed that equal amounts of purified proteins were loaded.

Figure 3. Structure in solution of the UBD domain of N4BP1. (A) ^{1}H-^{15}N HSQC NMR spectrum of N4BP1 obtained as reported in Materials and methods. (B) Amide proton region of the ^{1}H-^{1}H NOESY spectrum of N4BP1. (C) Finger print region of the ^{1}H-^{1}H NOESY spectrum of N4BP1. (D) (*Upper panel*). The sequence of the structured region is shown with the relative numbering here adopted. The numbering of residues referred to the full-length protein is shown in italic at the bottom. (*Left*) Cluster of the top 10 structures (out of 200) obtained by simulated annealing from NMR data of N4BP1. *(Right)* Cartoon representation showing the secondary structure of N4BP1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

Figure 4. Conformational studies of N4BP1 in complex with either ubiquitin or NEDD8, performed by circular dichroism (CD). (A) CD spectra of N4BP1 and ubiquitin alone expressed as rotatory power are shown in *black* and *red* traces respectively. The *orange* trace reports the CD spectrum of the 1:1 molar mix of N4BP1 and ubiquitin at the concentration of 30uM. The *green* trace ($N4BP1_{complex}$) reports the CD spectrum of N4BP1 in the complex obtained after subtracting point-by-point, from the CD of the complex *(orange)*, the CD spectrum of ubiquitin alone *(red)*. Similarly, the *blue* trace reports the CD spectrum of ubiquitin in the complex obtained after subtraction of the CD spectrum of N4BP1 alone *(black)*. (B) The differences shown in (A) are amplified and here reported as mean residue ellipticity. The *black* trace (a) represents the amplification of the point-by-point difference in the CD spectra between ubiquitin in complex with N4BP1 and ubiquitin alone. The *red* trace (b) represents the amplification of the point-by-point difference in the CD spectra between N4BP1 in complex with ubiquitin and N4BP1 alone. (C) CD spectra of N4BP1 and NEDD8 alone are shown in *black* and *red* traces respectively.. The *yellow* trace reports the CD spectrum of the 1:1 molar mix of N4BP1 and NEDD8 at the concentration of 30uM. The *blue* and *brown* traces report the CD spectra respectively of N4BP1, after subtracting the spectrum of NEDD8 from the spectrum of the mixture, and of NEDD8 after subtracting the spectrum of N4BP1 from the spectrum of the mixture. These differences are reported in *green* ($N4BP1_{diff}$) and *dashed red* ($NEDD8_{diff}$) traces respectively.

Figure 5. Chemical shift perturbation spectra of ^{15}N labelled N4BP1 in the presence of ubiquitin at 1:0.5 (A) and 1:1 (B) molar ratios.

Figure 6. Chemical shift perturbation spectra of ^{15}N labelled ubiquitin in the presence of N4BP1 at the molar ratio ubiquitin: N4BP1 of 1:0.5 (A) and 1:1 (B).

Figure 7. Conformational model of the N4BP1/Ubiquitin complex, obtained by combining the structure of N4BP1, the structure of Ubiquitin (BMRB 6457 and PDB 1D3Z) and the reported CSPs values [35] *(upper panels)*. Colours indicate the nature of the residues influenced by the interaction. Black: hydrophobic residues (A), Red: polar residues (B); Yellow: neutral residues (C). Sticks represent side chains of the amino acids whose ^{15}N resonances are perturbed upon interaction with ubiquitin (*lower panels*).

Figure 8. The distal residues K27, I28 and Q30 in helix-2 show significantly perturbed ^{15}N resonances upon interaction of N4BP1 with ubiquitin. Structural model showing residues in helix-2 (K27, I28 and Q30) that are perturbed in the N4BP1/ubiquitin complex. The side chains of the FP pair in the loop1 and the K27, I28 and Q30 residues in helix-2 are shown in sticks (see details in the text).

Figure 9. (A) Multiple sequence alignment between KHNYN, N4BP1 and members of the CUE domain protein family. Amino acid ranges: hAMFR (aa 457-498), hAUP1 (aa 362-404), hKHNYN (aa 634-678), hN4BP1 (aa 853-896), hTOLLIP (aa 229-271) and ScVps9p (aa 408-451). The FP/FW and the di-leucine

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

motifs are highlighted in grey boxes. Consensus symbols in the alignment: '.' indicates conservation between groups of weakly similar properties; ':' indicates conservation between groups of strongly similar properties; '*' indicates positions which have a single, fully conserved residue. Coloured boxes show the FP pair *(yellow)* and the φxx(I/L/V)L motif *(brown)*. (B,C,D) Cartoon representations of the CUE domain of AMFR (B, PdB 2EJS), the UBD of N4BP1 (C) and the CUBAN domain of KHNYN (D, PdB 2N7K). The FP/FW and the C-terminal hydrophobic motifs of N4BP1, AMFR and KHNYN are shown in sticks coloured as in (A).

Figure 10. (A) The GST fusions of KHNYN (aa 598-678) and N4BP1 (847-896), respectively including the wild-type carboxyl-terminal ends and the corresponding mutants in the FW (W647A) and FP (P866A) motifs were analysed by pull-down from a cell extract from T-Rex-Flag-NEDD8 cells expressing flag-tagged NEDD8 after incubation with 100nM doxycycline for 18h. After washing, the beads were analysed by SDS-PAGE with anti-Flag, anti-CUL1, anti-CUL2 and anti-ubiquitin. (B) The C-terminal regions of KHNYN and N4BP1 mediate the intra-molecular ubiquitination of the domains. The C-terminal end of KHNYN (aa 598-678) and N4BP1 (aa 813-896) were expressed as GFP fusions in HeLa cells and immunoprecipitated with anti-GFP antibody. Beads were washed and analysed by SDS-PAGE with anti-GFP and anti-ubiquitin antibodies. Two different acquisition times are shown. (C) The mutation W647A in KHNYN does not affect the ubiquitination of the C-terminal end. The KHNYN C-terminal end, both wild type and mutated in the FW motif (W647A), was transfected in HeLa cells and analysed as described in (B). Red bands represent western blotting signals that reached the saturation point. The immunoprecipitates were analysed by SDS-PAGE with anti-GFP and anti-ubiquitin antibodies. (D) The mutation P866A abrogates the polyubiquitination of N4BP1. The N4BP1 C-terminal end, both wild type and mutated in the FP motif (P866A), was transfected in HeLa cells and analysed as described in (C).

Figure 11. (A) Amino acid sequence alignment of ubiquitin and NEDD8 orthologs performed with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo). Consensus symbols used in the alignment: '.' indicates conservation between groups of weakly similar properties; ':' indicates conservation between groups of strongly similar properties; '*' indicates fully conserved residue. Conserved/divergent positions are highlighted in coloured boxes (*brown* for Ubiquitin, *violet* for NEDD8). (B) Complex between the UBD of N4BP1 and ubiquitin. The conserved/divergent patch in ubiquitin *(brown)* is represented by F4, T12, T14 and E64 (shown in bold in the sequence alignment). (C) The corresponding pattern in NEDD8 *(violet)* includes residues K4, E12, E14 and G64. The magnification of both patterns is shown. Protein models have been obtained with PyMol, using of the HADDOCK model coordinates.

Table 1. Chemical shift assignments of resonances of the ^{1}H and ^{15}N NMR spectra.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

Figures

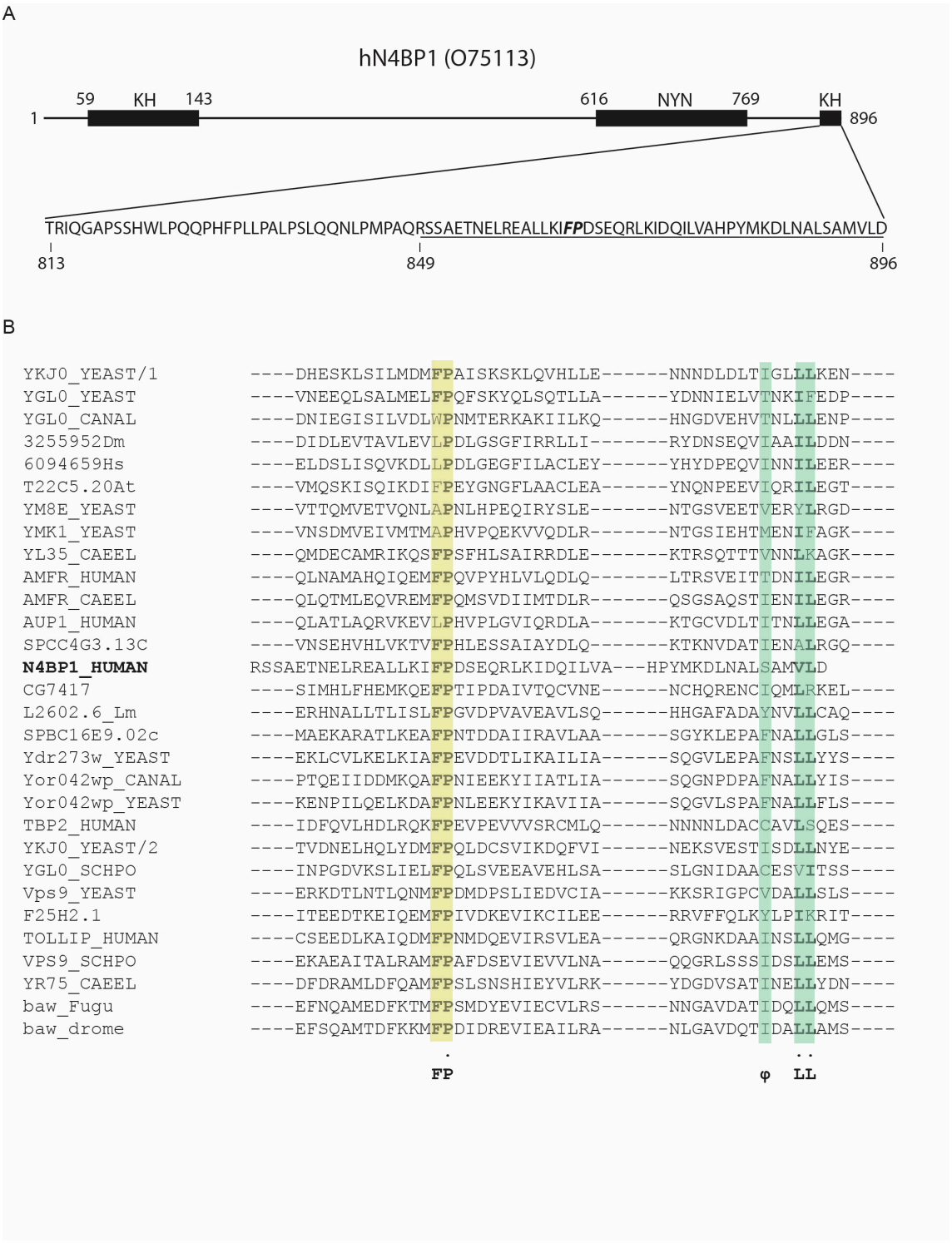


Figure 1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

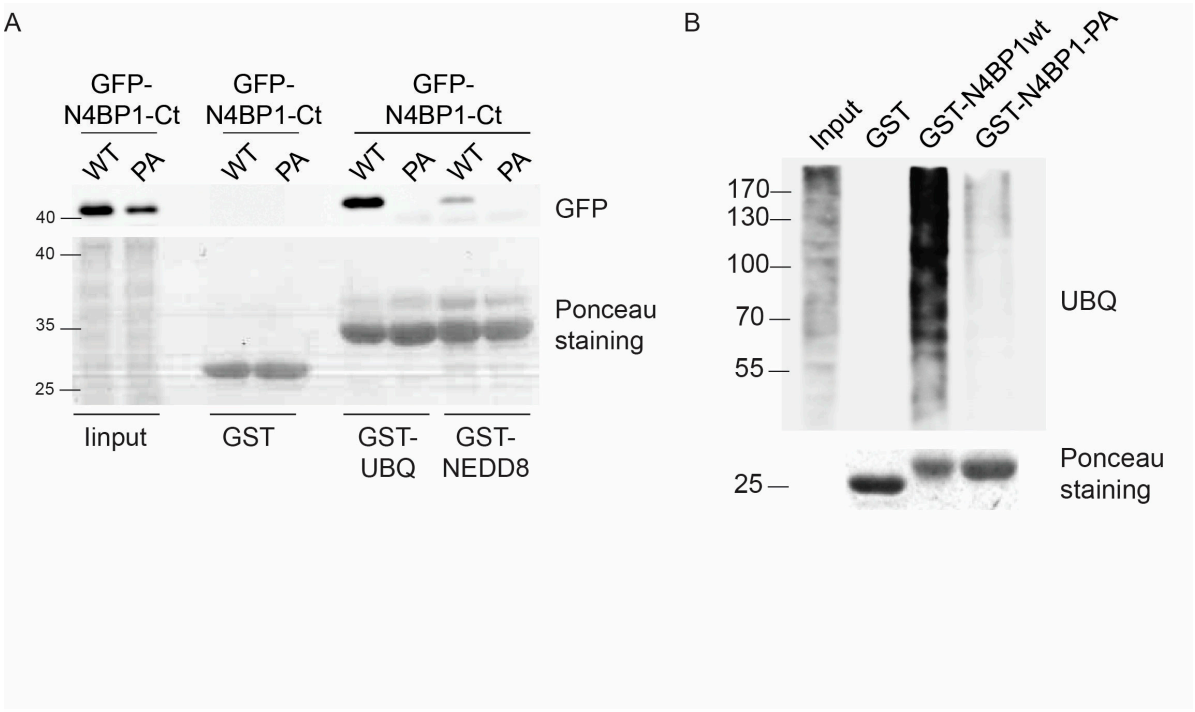


Figure 2

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

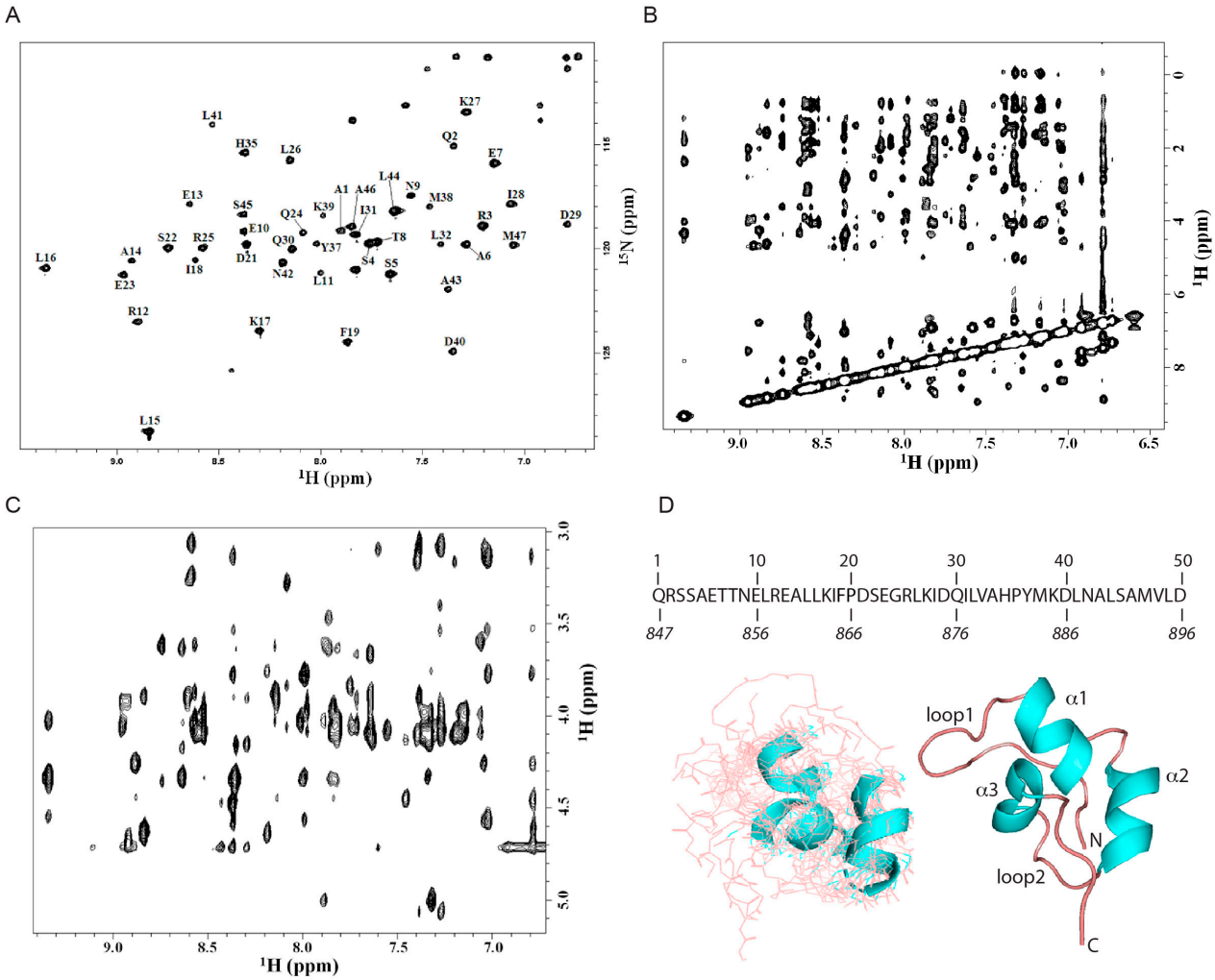


Figure 3

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

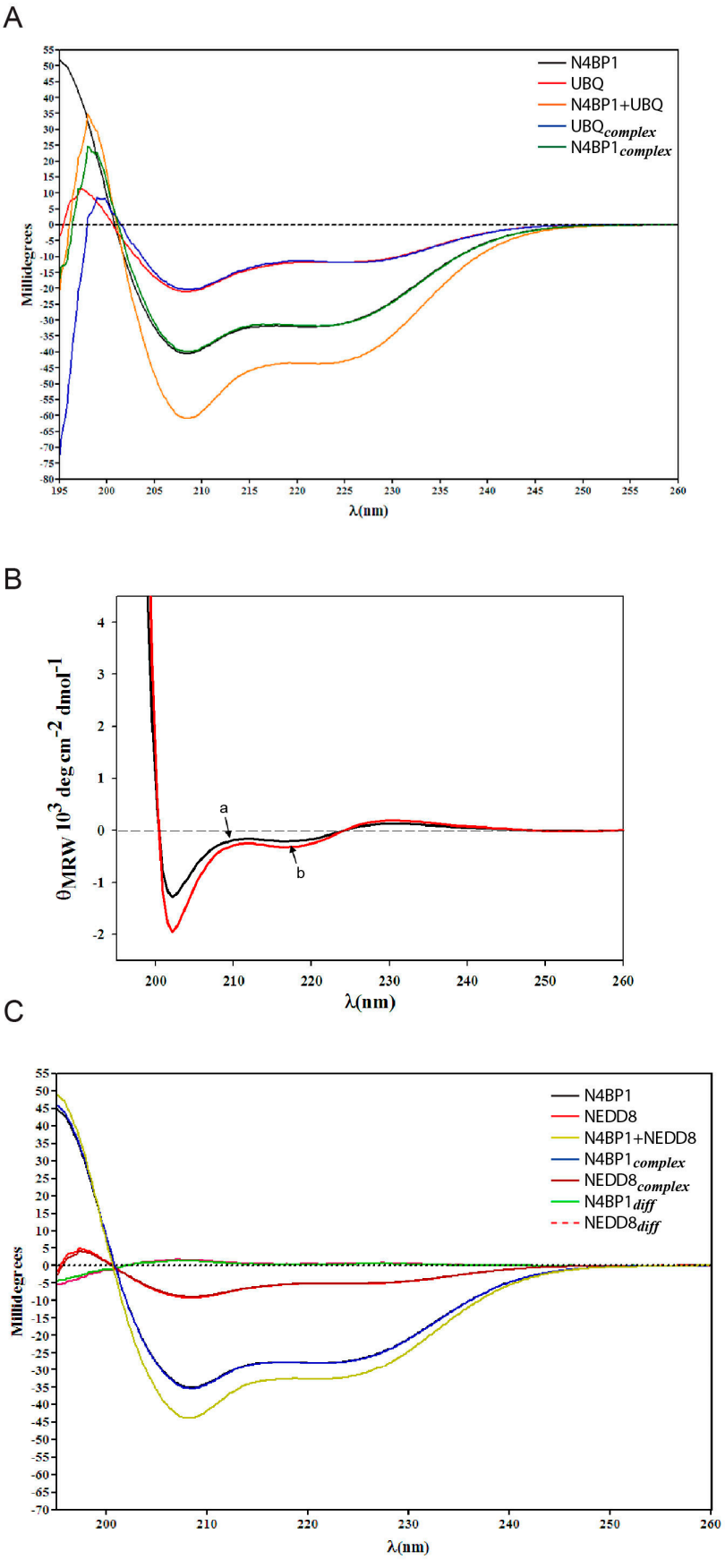


Figure 4

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

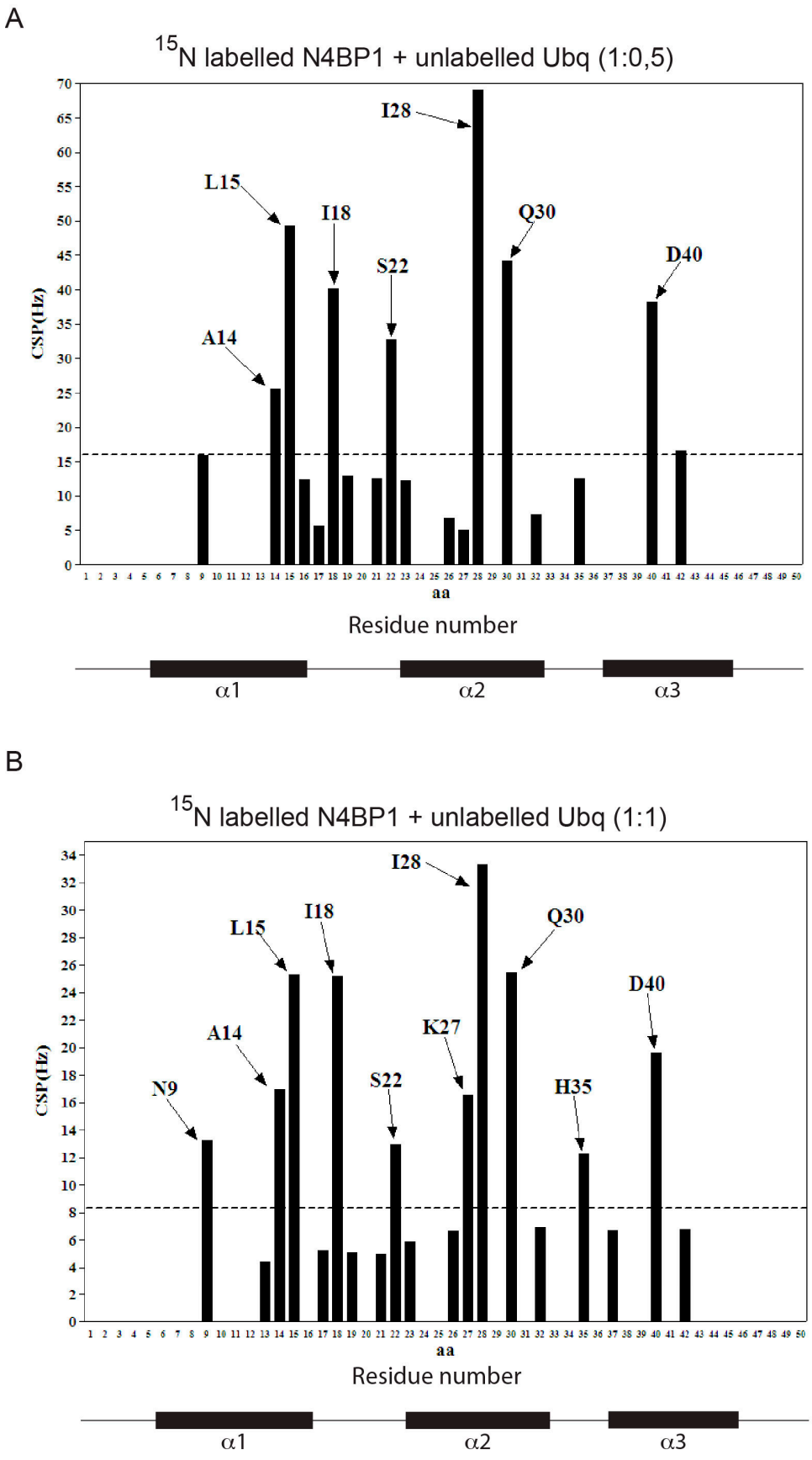
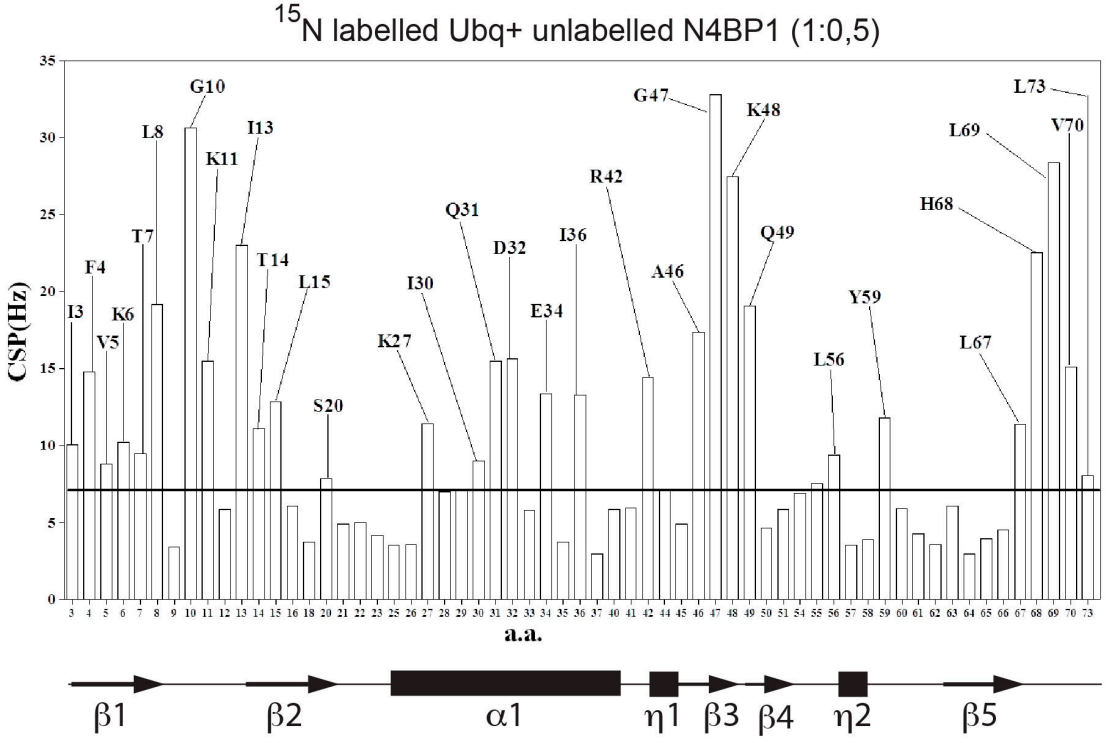


Figure 5

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

A



B

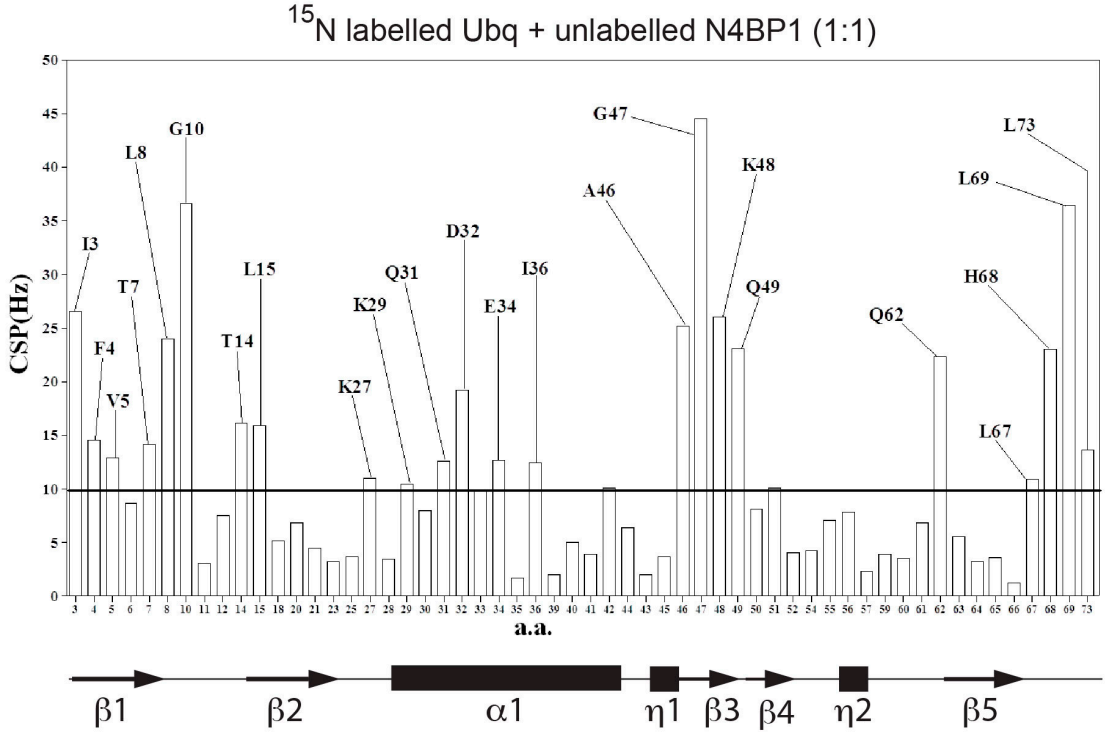


Figure 6

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

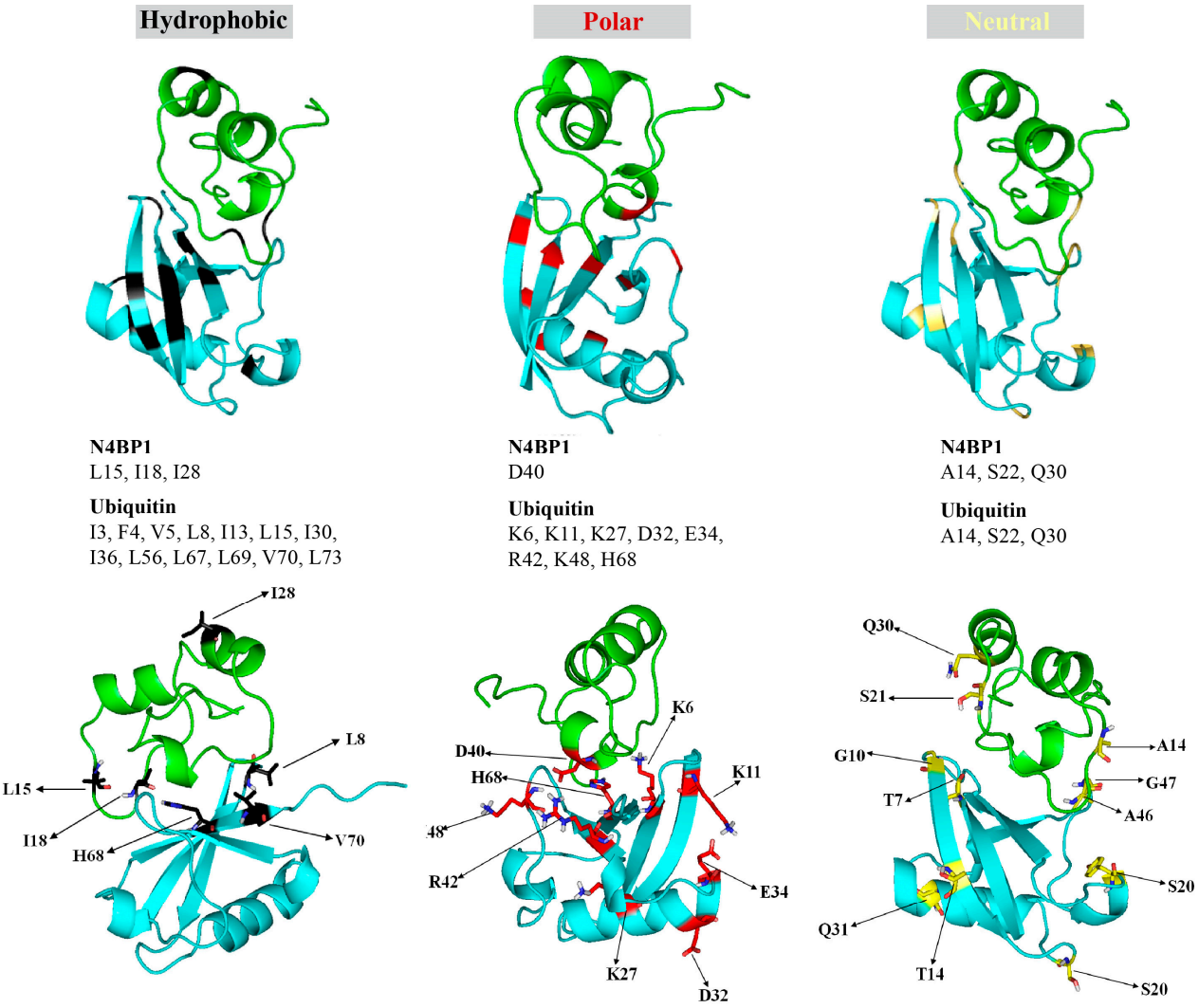


Figure 7

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

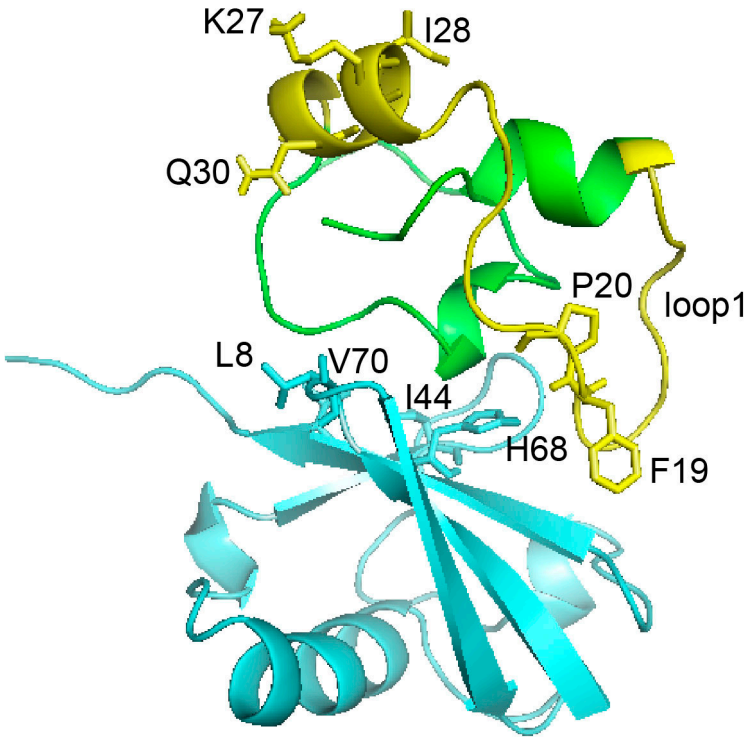


Figure 8

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

A

```
AMFR_Hs       457-LNAMAHQIQEMFPQVPYHL-VLQDLQLT-RSVEITTD--NILEGRI-498
AUP1_Hs       362-LATLAQRVKEVLPHVPLGV-IQRDLAKT-GCVDLTIT--NLLEGAV-404
KHNYN_Hs      634-ETERLRRQLLEVFWGQ--DHKVDFILQREPYCRDINQLSEALLSLNF-678
N4BP1_Hs      853-ETNELREALLKIFPDSEQRLKIDQILVAHPYMKDLNALSAMVLD----896
Tollip_Hs     229-CSEEDLKAIQDMFPNMDQEV-IRSVLEAQRGNKDAAIN--SLLQMG--271
Vps9p_Sc      408-ERKDTLNTLQNMFPDMDPSL-IEDVCIAKKSRIGPCVD--ALLSLSE-451
                          . : .::        :                    :*.
```

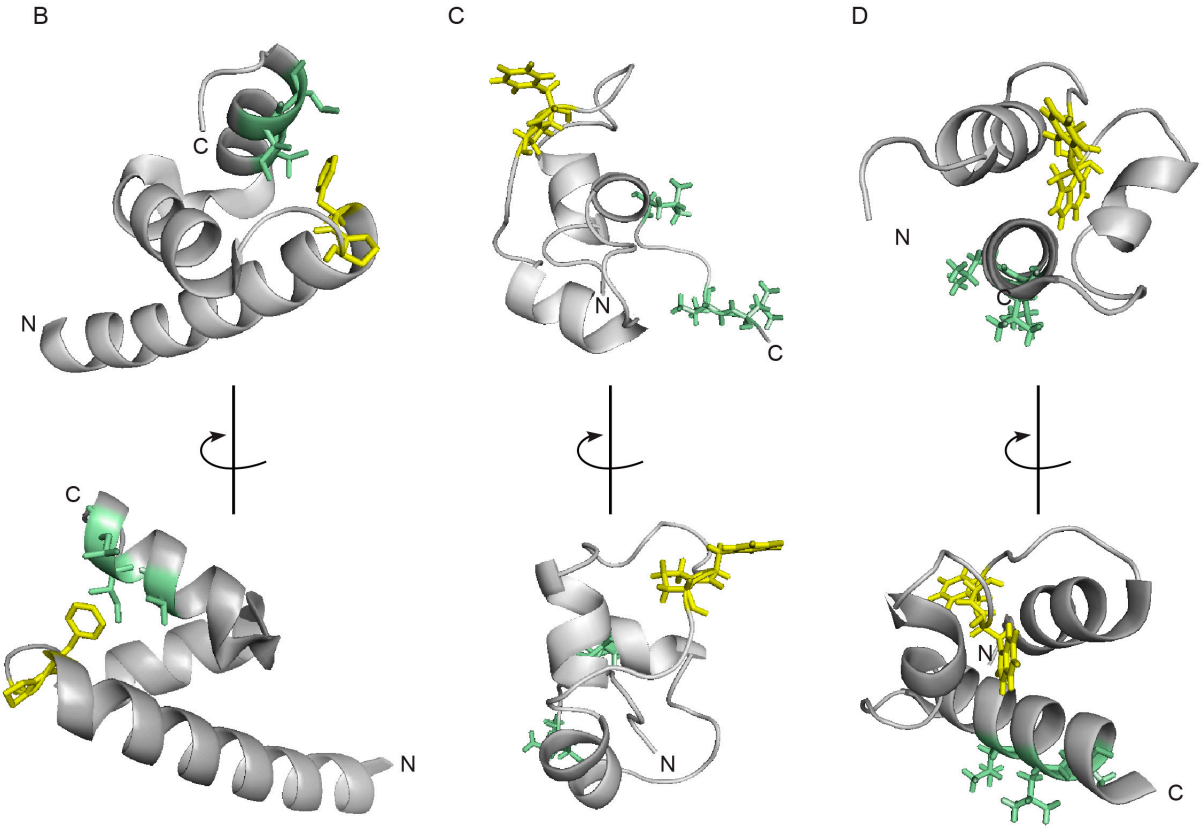


Figure 9

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

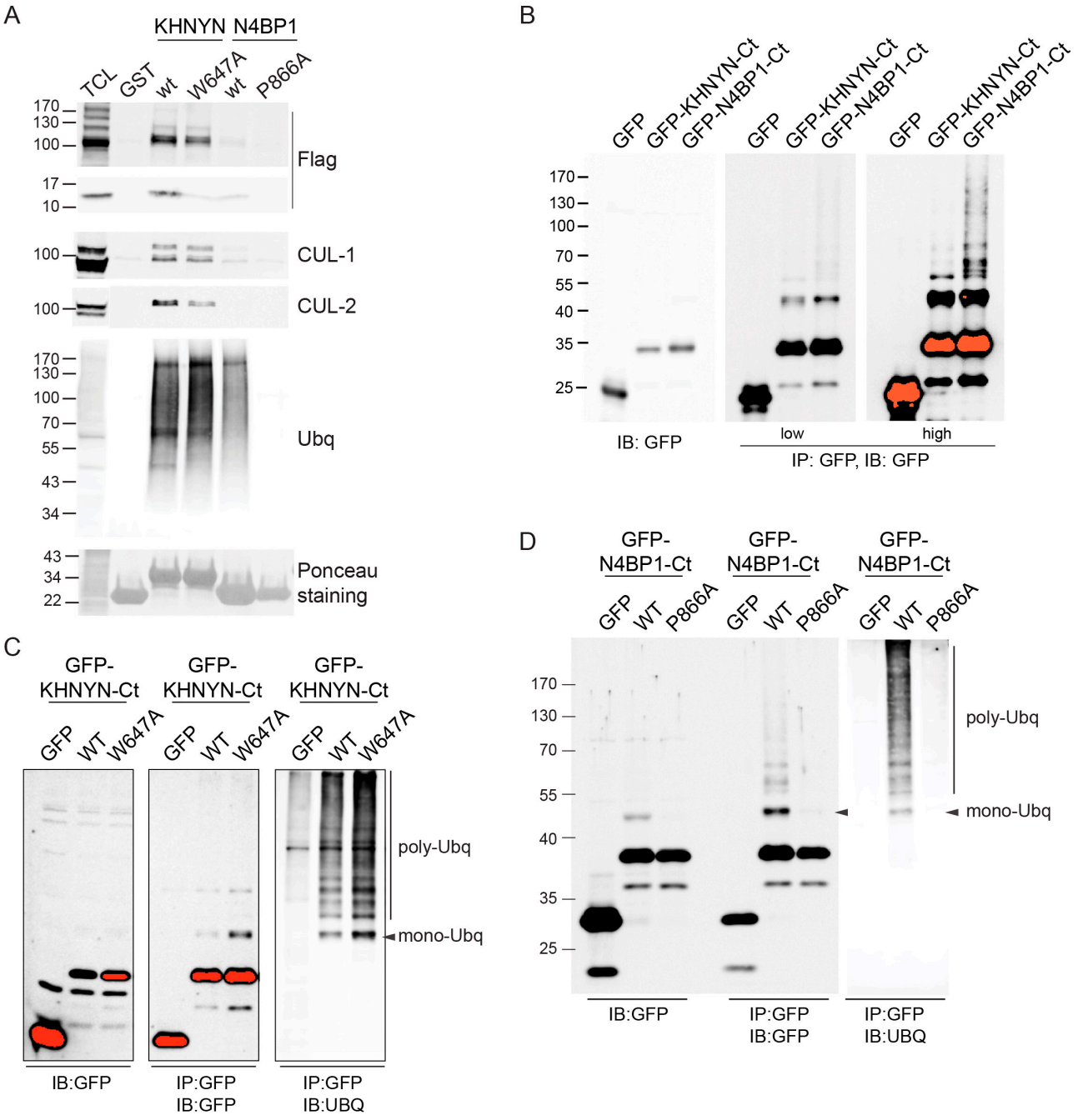


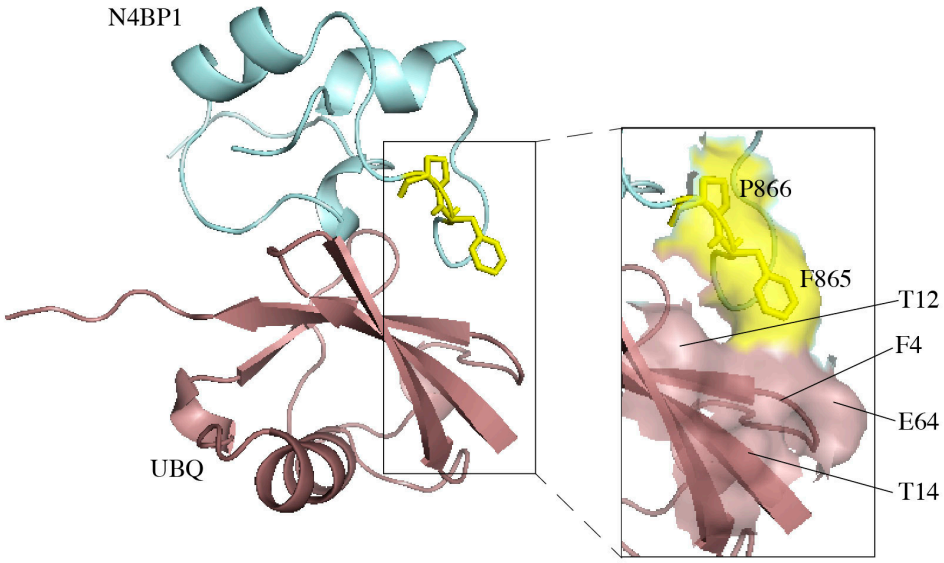
Figure 10

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

A

```
                                4          12 14                                                                64
Ubq(H.sapiens)          MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG
Ubq(R.norvegicus)       MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG
Ubq(D.melanogaster)     MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG
Ubq(C.elegans)          MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG
Ubq(B.taurus)           MQIFVKTLTGKTITLEVEPSDTIENVKGKIQEKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG
NEDD8(C.elegans)        MLIKVKTLTGKEIELDIEPNDRVERIKEKVEEKEGIPPPQQRLIFAGKQMNDDKTAADYKVLGGSVLHLVLALRGG
NEDD8(D.melanogaster)   MLIKVKTLTGKEIEIDIEPTDKVDRIKERVEEKEGIPPQQQRLIFSGKQMNDDKTAADYKVQGGSVLHLVLALRGG
NEDD8(H.sapiens)        MLIKVKTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQQRLIYSGKQMNDEKTAADYKILGGSVLHLVLALRGG
NEDD8(B.taurus)         MLIKVKTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQQRLIYSGKQMNDEKTAADYKILGGSVLHLVLALRGG
NEDD8(R.norvegicus)     MLIKVKTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQQRLIYSGKQMNDEKTAADYKILGGSVLHLVLALRGG
                        * * ******* * :::**.* ::.:* ::::****** *****::***::* :* :**::   *.***** ****
```

B



C

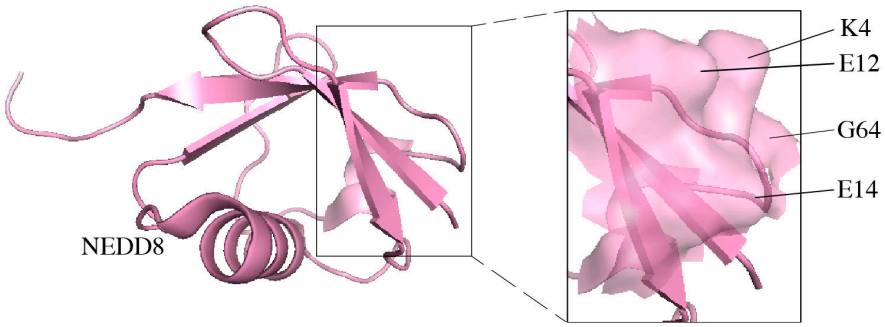


Figure 11

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 February 2019 doi:10.20944/preprints201902.0213.v1

Table1: ^{1}H and ^{15}N chemical shift assignments

	N	HN	Hα	Hβ1,2	Hγ	Hδ
A1	119,11	7,90				
Q2	115,05	7,34	4,35	1,96		
R3	118,85	7,20	4,00	1,68		
S4	119,71	7,76	4,70	3,85		
S5	121,19	7,66	4,15	3,95		
A6	119,78	7,29	4,06	1,48		
E7	115,90	7,15	4,00	_		
T8	119,64	7,72	4,07	4,05	1,20	
N9	117,40	7,56	4,10	2,70		
E10	119,06	8,43	3,97	1,88		
L11	121,01	8,00	4,37	1,70		
R12	123,40	8,91	4,60	1,61		
E13	117,83	8,64	3,80	1,90		
A14	121,51	8,96	4,60	_		
L15	128,80	8,82	4,60	2,04		
L16	120,89	9,30	4,60	_		
K17	123,89	8,31	4,24	1,59		
I18	120,30	8,60	4,13	_	1,00	
F19	124,40	7,80	4,38	3,06		
D21	119,80	8,43	4,38	2,74		
S22	119,90	8,73	_	3,79		
E23	121,33	9,10	4,60	2,08		
Q24	119,20	8,08	4,10	2,07		
R25	119,90	8,59	4,13	1,85		
L26	115,70	8,13	4,30	_		
K27	113,42	7,26	4,06	1,91		
I28	118,24	7,17	3,79	1,45	0,83	0,04
D29	118,87	6,87	4,70	2,93		
Q30	119,40	8,11	3,90	2,00		
I31	119,30	7,83	4,10	_	0,08 (Hγ2)	-0,08
L32	119,75	7,41	3,31	_		
V33	115,36	8,37	4,57	2,90		
H35	115,36	8,37	3,14	3,14		
Y37	119,73	8,01	5,10	3,01		
M38	117,95	7,47	4,10	2,70		
K39	118,38	8,04	4,08	_		
D40	124,90	7,46	4,36	2,66		
L41	113,90	8,55	4,05	_		
N42	120,64	8,20	4,70	2,86		
A43	121,92	7,37	4,67	1,40		
L44	118,16	7,64	4,15	_		
S45	118,32	8,37	4,58	3,97		0,9 (Hδ1,2)
A46	119,77	7,29	4,05	1,40		
M47	119,80	7,05	4,27	2,04	2,52	

Table 1