UNIVERSITA' DEGLI STUDI DI VERONA

Department of Biotechnology

Graduate School of Natural Sciences and Engineering

Doctoral Program in Biotechnology

Cycle XXX

Investigation of molecular recognition properties of ubiquitin, polyubiquitin and a disease-associated mutant

S.S.D. CHIM/06

Coordinator: Prof. Matteo Ballottari

Tutor: Prof. Michael Assfalg

PhD student: ANDREA BORTOT

Biomolecular interaction studies on the Ubiquitin system

Andrea Bortot

Tesi di Dottorato

Abstract

The small protein ubiquitin acts as a versatile cellular signal that controls a wide range of biological processes. The specificity of ubiquitin signalling is achieved by alternative conjugation signals and interactions with ubiquitin-binding proteins. Despite tremendous advancements in our understanding of ubiquitin function, the molecular details of recognition are still not fully elucidated. In this regard, solution NMR spectroscopy studies show promise to shed light into transient molecular interactions and conformational dynamics governing ubiquitin-mediated signalling.

The traditional methods of studying proteins are implemented within dilute solutions with less than 10 g/L of total protein concentration. This low concentration allows to obtain good signals but may not adequately represent a biological environment. One distinctive feature of cellular systems is that the cytoplasm is deeply crowded with macromolecules (50-400 g/L) which affect several protein attributes. Macromolecular crowding can result in non-specific interactions between the protein of interest and the target protein. The broad aim of our study is to understand the effects of macromolecular crowding on ubiquitin recognition. We focused on the ubiquitin-UBA interaction, investigating the perturbations induced by the presence of a synthetic crowding agent in comparison with dilute solution by NMR. We analysed differences in binding affinity, structure and dynamics of the complex dissolved in the different media.

Protein-protein interactions are a prime target for drug development and chemical biology research. Mechanisms of protein recognition have been extensively studied for single-domain proteins, but are less well characterized for dynamic multidomain systems. PolyUb represent an important multidomain system that requires recognition by structurally diverse ubiquitin-interacting proteins. Thus, the development of chemical species able to selectively recognize polyUb has become a subject of strong interest. Clearly, nanoparticles (NPs) present several advantages for protein recognition, including a large surface available for interaction. In our project, we aimed to explore NP systems for the development of polyUb-specific receptors. We investigated the binding specificity of chemically diverse NPs towards structurally

distinct polyUb. Solution NMR spectroscopy was chosen as the central experimental technique due to its ability to provide site-resolved information on reversibly binding protein-NP pairs. Our results constitute the basis for an improved understanding of polyUb recognition by artificial receptors and for the development of NP-based therapeutic strategies.

Given the central role of the Ub network in cellular physiology, misregulation is often associated with diseases, including cancer, immune disorders, and neurodegeneration. Due to a possible participation of the frameshift Ub mutant Ubb^{+1} in the molecular events leading to neurotoxicity and neurodegeneration in Alzheimer's disease, there is large interest in elucidating the structural details of this aberrant protein and the consequent functional differences with respect to the wild-type protein. In our work, we investigated structural and dynamic features of Ubb^{+1} using NMR methods that are particularly suited to explore protein molecules containing flexible domains such as the C-terminal extension of Ubb^{+1}.

Sommario

La proteina ubiquitina agisce come un segnale cellulare versatile in grado di controllare un'ampia gamma di processi biologici. La specificità del segnale dell'ubiquitina è ottenuta mediante segnali di coniugazione alternativi e interazioni con proteine leganti l'ubiquitina. Nonostante gli enormi progressi nella comprensione della funzione dell'ubiquitina, i dettagli del riconoscimento molecolare non sono ancora del tutto chiariti. A questo proposito, gli studi di spettroscopia NMR in soluzione hanno fornito promettenti informazioni su interazioni molecolari transitorie e dinamiche conformazionali che governano la segnalazione mediata da ubiquitina.

I metodi tradizionali di studio delle proteine vengono effettuati in soluzioni diluite, le quali presentano meno di 10 g / L di concentrazione proteica totale. Questa bassa concentrazione consente di ottenere buoni segnali ma potrebbe non rappresentare adeguatamente un ambiente biologico. Una caratteristica distintiva dei sistemi cellulari è dovuta al citoplasma molto affollato di macromolecole (50-400 g/L) che influisce su diverse caratteristiche delle proteine. L'affollamento macromolecolare può determinare interazioni non specifiche tra la proteina di interesse e la proteina bersaglio. L'obiettivo generale del nostro studio è comprendere gli effetti dell'affollamento macromolecolare sul riconoscimento dell'ubiquitina. Ci siamo concentrati sull'interazione ubiquitina-UBA, studiando, mediante spettroscopia NMR, le perturbazioni indotte dalla presenza di un agente affollante sintetico rispetto alla soluzione normalmente diluita. Abbiamo analizzato le differenze nell'affinità di legame, nella struttura e nella dinamica del complesso immerso nelle diverse soluzioni.

Le interazioni proteina-proteina sono un obiettivo primario per lo sviluppo di farmaci e nella ricerca chimico-biologica. I meccanismi di riconoscimento proteico sono stati ampiamente studiati per le proteine a singolo dominio, ma sono meno caratterizzati per sistemi dinamici multidominio. Le catene di poliubiquitina (polyUb) rappresentano un importante sistema multidominio, le quali richiedono di essere riconosciute da parte di proteine strutturalmente diverse. Pertanto, lo sviluppo di specie chimiche in grado di riconoscere selettivamente le polyUb è diventato un argomento di forte interesse.

Chiaramente, le nanoparticelle (NP) presentano numerosi vantaggi per il riconoscimento delle proteine, inclusa una grande superficie disponibile per l'interazione. Nel nostro progetto intendiamo esplorare i sistemi NP per lo sviluppo di recettori specifici per le polyUb. Abbiamo quindi studiato la specificità di legame di NP chimicamente diverse verso polyUb strutturalmente distinte. La spettroscopia NMR in soluzione è stata scelta come tecnica sperimentale centrale grazie alla sua capacità di fornire informazioni sito specifiche sulle coppie, a interazione reversibile, proteina-NP. I nostri risultati costituiscono la base per una migliore comprensione del riconoscimento delle polyUb da parte di recettori artificiali e per lo sviluppo di strategie terapeutiche basate su NP.

Dato il ruolo centrale della rete Ub nella fisiologia cellulare, la sua errata regolazione è spesso associata a malattie, tra cui cancro, disordini immunitari e neurodegenerativi. A causa di una possibile partecipazione del mutante Ub di frameshift, noto come Ubb^{+1}, agli eventi molecolari che portano a neurotossicità e a processi neurodegenerativi nella malattia di Alzheimer, vi è un grande interesse nello studiare i dettagli strutturali di questa proteina aberrante e le conseguenti differenze funzionali rispetto alla proteina nativa. Nel nostro lavoro abbiamo studiato le caratteristiche strutturali e dinamiche della proteina Ubb^{+1} usando metodi NMR, i quali risultano particolarmente adatti per esplorare molecole proteiche contenenti domini flessibili come l'estensione C-terminale di Ubb^{+1}.

INDEX

ACRONYMS

AD Alzheimer's disease
APP Amyloid precursor protein
AAZTA Heptadentate ligand 6-amino-6-methylperhydro-1,4 diazepinetetraacetic acid
CARA Computer Aided Resonance Assignment
CSP Chemical shift perturbation analysis
DLS Dynamic Light Scattering
DS Down Syndrome
FICOLL (Poly(sucrose-co-epichlorhydrin))
HSQC Heteronuclear single-quantum coherence
HHR23A Human isoformA of Rad23
IPTG Isopropyl-β-D-1-thiogalactopyranoside
L Liter
LB Luria broth
LNP Lipid nanoparticle
LUV Large unilamellar vesicle
M Molar
MHz Megahertz
min Minute
mL Milliliter
mm Millimeter
mM Millimolar
mol Mole
NMR Nuclear Magnetic Resonance
NOE Nuclear Overhauser Effect
NP Nanoparticle
PP Parkinson's disease
P:L Protein/ligand ratio
PolyUb Polyubiquitin
PPI Protein-protein interaction
ppm Chemical shift in parts per million
R_1 Longitudinal relaxation rate
R_2 Transverse relaxation rate
SDS-PAGE Sodium dodecyl sulphate polyacrilamide gel Electrophoresis
SUV Small unilamellar vesicles
T_1 Longitudinal relaxation time
T_2 Transverse relaxation time
Ub Ubiquitin
UBA Ub-associated
UBL Ub-like
UPP Ubiquitin-Proteasom Pathway
UPS Ubiquitin Proteasome system
µL Microliter
1D Mono-dimensional
2D Bi-dimensional

CHAPTER 1:
Introduction

1) Structural overview of ubiquitin

Protein modifications via phosphorylation, acetylation, methylation and glycosylation are well recognized as biologically significant signals for mediating cellular response to external factors. In the late 1970's, Hershko, Ciechanover, and Rose found that a heat-stable protein, ubiquitin (Ub), covalently attaches to proteins in rabbit reticulocyte lysates and signals for proteolysis[1]. Biochemical studies thereafter have determined that the covalent linkage of substrate proteins to Ub, termed ubiquitination, signals for a multitude of cellular outcome[1].

Ubiquitin is an 8.5 kDa regulatory protein found in almost all tissues in many eukaryotic organisms. It contains a highly conserved sequence of 76 amino acids that is identical in a wide variety of sources including humans, fish, and insects[2]. Ub is a highly stable protein that adopts a compact β-grasp fold (Figure 1), consisting of 3.5 turns of an amphipathic α-helix and a short 3_{10}-helix packed against a five-stran β-sheet with seven reverse turns. The surface of Ub is complex, with multiple functionalities, which explains its high degree of amino acid sequence conservation. The first loop containing Leu-8 is able to adopt different conformations important for interaction with distinct Ub binding proteins. Another region, consisting of Ile-44, Leu-8, Val-70 and His-68, called the Ile-44 hydrophobic patch, interacts with the proteasome and other Ub binding proteins[3]. In addition to a core of sixteen-seventeen hydrophobic residues, there is extensive intramolecular hydrogen bonding. Altogether the properties of tight packing, a large hydrophobic core and extensive hydrogen bonding apparently confer structural stability, explaining its heat stable properties.

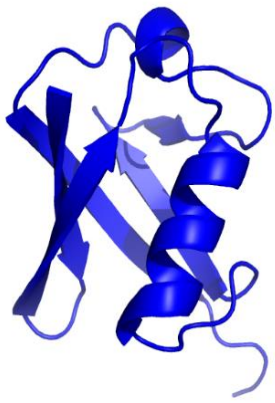


Figure 1. Solution structure of Ubiquitin. Pdb code 1d3z.

Because Ub itself possesses 7 surface Lys residues, it can undergo iterative ubiquitination to result in the formation of polyubiquitin chains (polyUb) of varying

length and Lys linkages. All Ub-mediated cellular events involve the covalent linkage of the C-terminal Gly (Gly76) of Ub to the ε-amine of a specific Lysine residue on the surface of the substrate protein. The remarkable versatility and complexity of Ub signaling is acquired from these variations. Thus, the outcome of ubiquitination depends on whether a protein is attached to a single Ub or a polyUb chain and the specific lysine involved in forming the chains[4].

1.1) Polyubiquitin chains

Modification of substrate proteins with monoUb or polyUb signals for various cellular outcomes depends on the nature of the modification. PolyUb chains function as signalling molecules in the regulation of a host of cellular processes, ranging from progression through the cell cycle, to transcriptional activation, antigen processing, and vesicular trafficking of proteins[5–8]. The diversity in signalling has been attributed to the different conformations adopted by polyUb chains as a result of various lysine linkages and chain lengths.

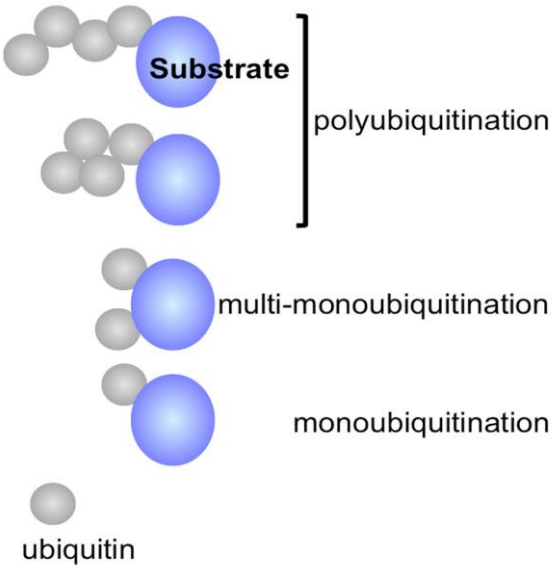


Figure 2. Representation of the diversity of ubiquitinated products [6]

One Ub is generally attached through an isopeptide bond involving the free carboxyl group in its C-terminal glycine residue and the ε-amino group in the side chain of a lysine residue in the target protein. In addition to a single Ub modification (monoubiquitination) or modification by one Ub at multiple sites of the same substrate (multi-monoubiquitination), substrates can be modified by Ub chains (polyubiquitination) (Figure 2).

There are seven potential lysine residues (K6, K11, K27, K29, K33,K48 and K63) on Ub that can participate in chain linkage formation[9]. Five Ub lysine residues (K6, K11, K29, K48 and K63) are known to be sites of initiation and these linkages exist *in vivo* as anchored and unanchored species. Thus, the different Ub-Ub linkages form distinct conformations and consequently utilize a distinct set of downstream interacting proteins that "interpret" the diverse Ub signals[10]. The functional outcome of polyubiquitinaton of a target protein depends on the length of the polyUb tag and the

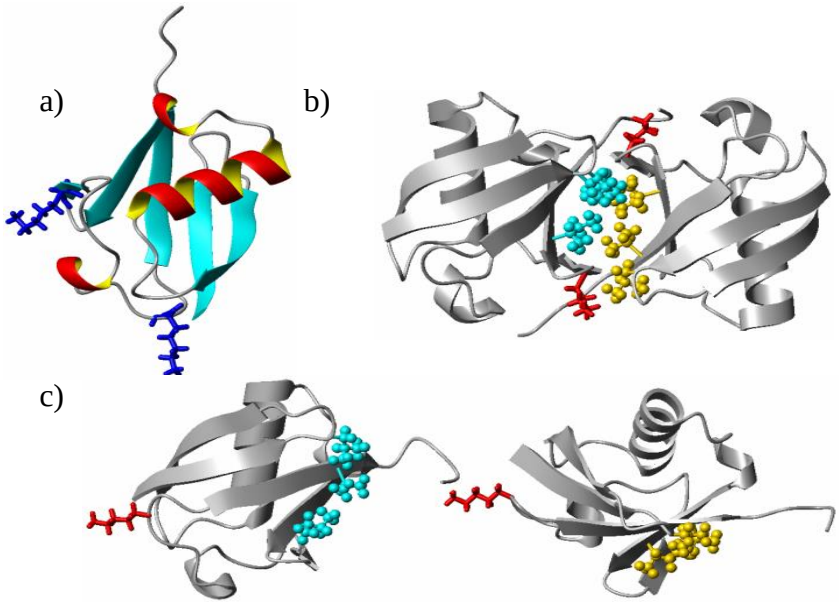


Figure 3: Di-ubiquitin linked via Lys48 versus Lys63 a) Ubiquitin structure,sidechains of Lys48 and Lys63 shown as blue sticks. b) crystal structure of Lys48-linked Ub2 with sidechains of Leu8, Ile44, Val70 shown as cyan ball and stick model on distal Ub and yellow on the proximal Ub. Red sticks designate position of sidechains of Gly76 and Lys48. c) solution structure of Lys63-linked Ub2 with sidechains of Leu8, Ile44, Val70 shown as cyan ball and stick model on distal Ub and yellow on the proximal Ub. Red sticks designate position of sidechains of Gly76 and Lys63.

lysine residue involved in the Ub_Ub linkage[11]. The type and number of polyUb chains that are conjugated to a target are highly regulated to generate distinct signals that affect different physiological processes.

Residue K48 is a major site of chain initiation and K48 linkages are highly abundant, being the predominant signal for proteins destined for degradation by the proteasome (Figure 3). The first polyUb structure, determined in 1992, was Lys48-linked Ub2 crystallized at pH 4.5. The Ub moieties adopted a closed conformation involving direct contacts between the hydrophobic patches (Leu8, Ile44, Val70) of the Ubs. The other principal and relatively abundant polyUb chain has K63 linkages (Figure 3). The solution structure at pH 6.8 reveals an extended conformation where the hydrophobic patches on each Ub do not interact and are solvent exposed. Ub_2 linked via Lys63 adopts a more extended conformation than Ub_2 linked via Lys48[12]. K63 linkages do not seem to play a role in protein turnover and have been implicated in receptor endocytosis and sorting, translation, DNA damage repair, the stress response and signalling through the TRAF pathway of NF-kB.

It is believed that the specificity of the recognition signal carried by a particular polyUb chain is determined by the unique conformations that a particular chain can adopt, which in turn are dictated by the linkage type[13].

This tagging process leads to their recognition by the 26S proteasome, a very large multicatalytic protease complex that degrades ubiquitinated proteins to small peptides[14]. The rapid degradation of ubiquitinated proteins is catalyzed by the 26S proteasome (Figure 4). This structure is found in the nucleus and the cytosol of all cells and constitutes approximately 1 to 2% of cell mass[15]. The 26S complex is composed of a central barrel-shaped 20S proteasome with a 19S regulatory particle at either or both of its ends[16]. After the ubiquinated protein binds to the 19S component, the polyUb chain is cleaved off the substrate and disassembled. The protein is unfolded somehow by the six ATPases in the base of the particle[17]. Linearization of the folded protein is essential for it to be translocated through the gated entry channel into the 20S particle. After the substrate enters the 20S's central chamber, the polypeptide is cleaved by its six proteolytic sites on the inner face of the changer, forming small peptides that range from three to 25 residues in length [18].

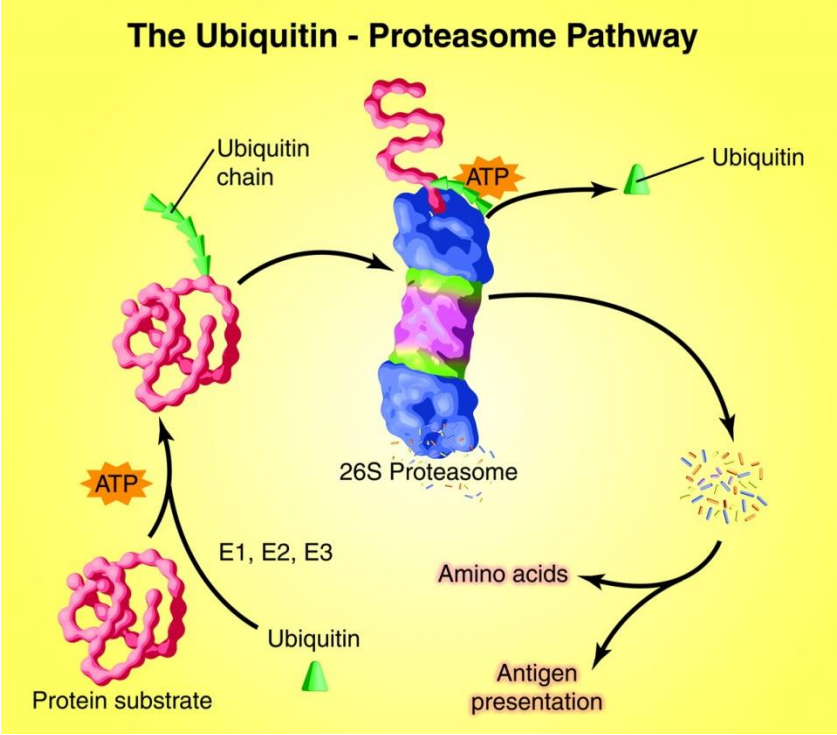


Figure 4. The ubiquitin (Ub)-proteasome pathway (UPP) of protein degradation. Ub is conjugated to proteins that are destined for degradation by an ATP-dependent process that involves three enzymes. A chain of five Ub molecules attached to the protein substrate is sufficient for the complex to be recognized by the 26S proteasome. In addition to ATP-dependent reactions, Ub is removed and the protein is linearized and injected into the central core of the proteasome, where it is digested to peptides. The peptides are degraded to amino acids by peptidases in the cytoplasm or used in antigen presentation.[6]

1.2) Enzymatic Conjugation in vivo and in vitro

The covalent attachment of Ub to the substrate protein is an obligatory step in all Ub-mediated events. The process of Ub conjugation (ubiquitination) occurs via tightly regulated enzymatic steps catalyzed by a series of Ub activating (E1), conjugating (E2) and ligase (E3) enzymes[1,19,20] (Figure 5). In the first step, the conserved C-terminal glycine of Ub is essential for activation by E1. This charged C-terminal residue eventually becomes conjugated to the lysyl ε-amino group of target proteins to form isopeptide linkages and subsequent conjugates. The E1 for Ub contains two active sites and activates the C-terminus of Ub via a two-step, intra-molecular, ATP-dependent reaction. Initially, a tightly enzyme-bound Ub adenylate (with PPi from ATP) is formed.

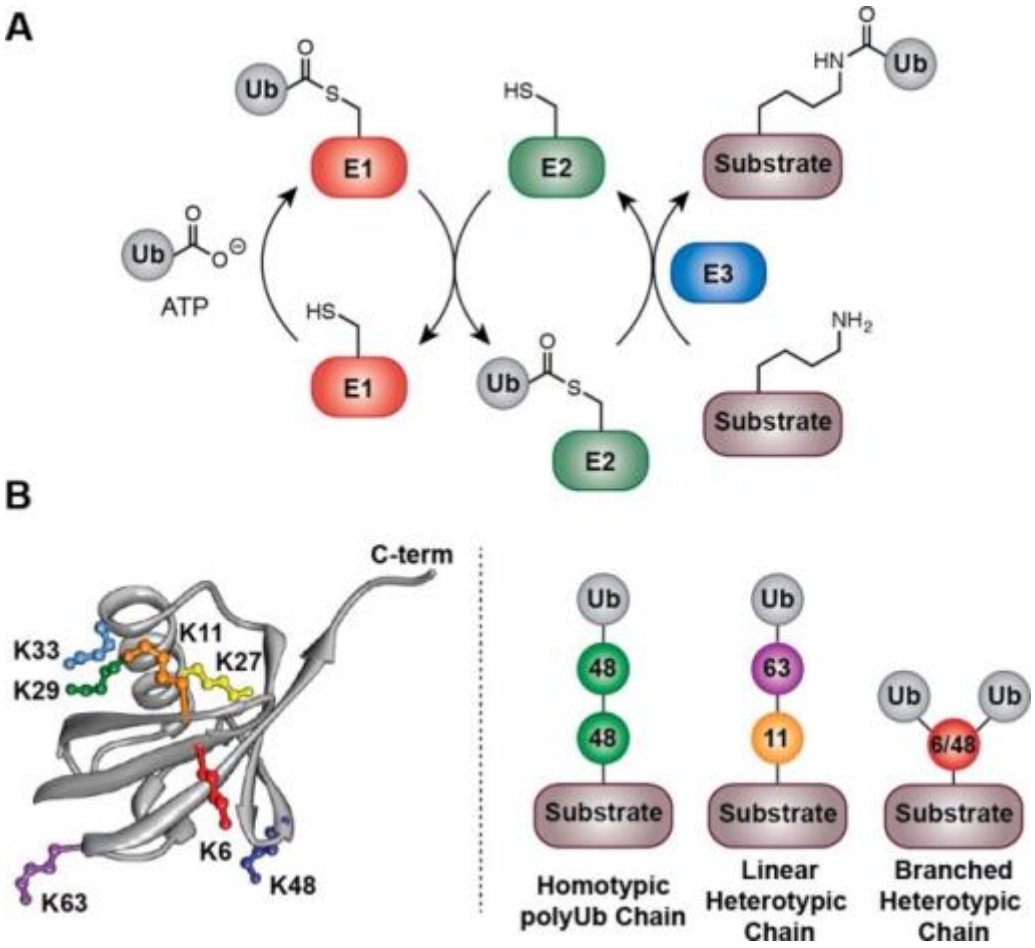


Figure 5: Protein ubiquitylation.(A) Cascade of E1, E2, and E3 enzymes thatcatalyze the formation of an isopeptide bond between a substrate proteinand ubiquitin (Ub). (B) Structure of Ub (PDB entry 1ubq) showing the sevenlysines (K6, K11, K27, K29, K33, K48, and K63) and the types of polymericUb (polyUb) chains that form due to the presence of these residues.The lines between Ub subunits in the chains denote an isopeptide linkage,and the numbers indicate the lysine used to link subunits together.In the case of a branched heterotypic chain, a single subunit is modifiedwith two or more Ub molecules via two or more isopeptide linkages. Biochemistry. 2014 Aug 5;53(30):4979-4989

This intermediate is then converted to form AMP and a covalent enzyme-ubiquitin thioester. Activation of a second Ub molecule gives a complex with one equivalent each of Ub thioester and tightly bound Ub adenylate per subunit of enzyme. Activated Ub is then transferred to an active site Cys residue of a Ub-carrier protein, E2. These enzymes perform the second step in conjugation reactions by forming a thioester linkage with the

C-terminal glycine. E2 enzymes function alone and in conjunction with E3 ligases to catalyze the attachment of Ub to acceptor lysine residues of target proteins to form isopeptide bonds. Mass spectrometric analysis of Ub conjugates in *S. cerevisiae* has determined that polyUb chains are found in relative abundance Lys48 > Lys63 and Lys11 >> Lys33, Lys29, Lys27, and Lys6 [21]. An essential step necessary to allow in vitro biochemical and structural studies of polyUb is the production of their chains in high quantities and purity. PolyUb chains have been obtained either by chemical synthesis of suitable functionalized precursors or by enzymatic assembly. The specific enzymes for the formation of most Ub chain types are known and well characterized, allowing an enzymatic procedure. A rather convenient enzymatic method for the production of K48-linked Ub chains of selected length was reported in 1997 by Pickart and coworkers[22]. The method was later extended to K63-linked chains, and detailed protocols were published[23,24]. It relies on the ability of an E2 Ub conjugating enzyme to form a specific type of isopeptide bond between two Ub molecules. An E1 enzyme activates Ub for the reaction, and the specific E2 catalyzes the conjugation. ATP is consumed to form the charged E1. Specificity for a given chain type is achieved by the choice of a specific E2 (or E2/E3 couple). Two Ub reactants are used: one reversibly capped at its C terminus and the other blocked at its lysine site. The authors defined as the "proximal end" of a chain an unconjugated G76 residue because it serves for the attachment of a proximal Ub residue.

Accordingly, the "distal end" of a chain is the unconjugated lysine residue (Figure 6A). The Ub mutant blocked at the proximal site (also referred to as proximally blocked Ub) is Ub-D77 because the presence of an extra amino acid after G76 impairs conjugation through the C terminus of a proximal Ub. The distally blocked Ub is obtained by mutating to cysteine or arginine the lysine involved in the formation of the specific chain type by the selected E2, so that the covalent binding of a distal Ub is impaired. For example, in the case of K48-linked chains, recombinant E2-25K, able to specifically form K48-linked chains, is used to catalyze the reaction between equal amounts of Ub-D77 and Ub-K48R mutants (Figure 6B). Thus, diUb is formed in almost equimolar concentration with respect to the monoUb reactants. The reaction will not proceed with the formation of longer chains unless the doubly blocked diUb is

conveniently deprotected. A purification step (often cation exchange chromatography) will separate the product from unreacted Ub molecules.

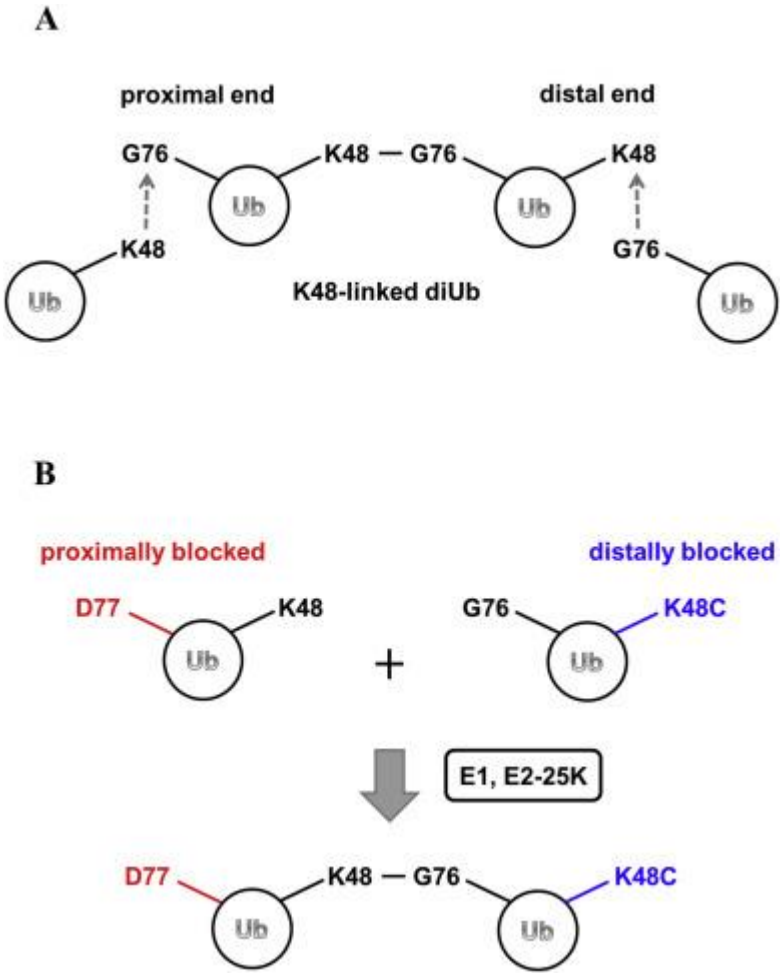


Figure 6. (A) Scheme of K48-linked diUb. Chain elongation can occur at the "proximal end" on binding of a proximal Ub or at the "distal end"on covalent conjugation of a distal Ub. (B) Production of K48-linked diUb using Ub mutants. (*Faggiano et al., 2016)*

2) Ubiquitin Associated Domains

The variety of cellular events that involve Ub-mediated regulation suggests the existence of several proteins in the cell that interact with polyUb chains. Indeed, a number of different Ub interacting proteins that possess both an Ub binding domain and a variable effector domain have been found. A complex network of Ub-associated (UBA) and Ub-like (UBL) domains have been implicated in the Ubiquitin Proteasome

Pathway (UPP), however the functional roles of these proteins require further characterization. Indeed, a number of different Ub interacting proteins that possess both an Ub binding domain and a variable effector domain have been found[16]. The various Ub binding motifs include UBA (ubiquitin associated). The UBA domain was originally identified by sequence analysis as a domain present in enzymes of the Ub-proteasomal pathway and has been subsequently found in several other proteins involved in Ub-mediated signalling pathways [11]. At least 15 proteins in fission yeast have been found to contain putative UBA domains[25].

The UBA domain is a short sequence motif of 45 amino acid residues that occurs frequently in proteins found in all eukaryotes. UBA motif can bind directly to Ub and/or poly-Ub, leading to an inhibition of the degradation of target substrates through the proteasome[26]. The UBA construct used in this study contains a full-length UBA2 domain from HHR23A (human isoform A of Rad23). HHR23A is one of the human homologues of the yeast Rad23 and Rhp23. All of the Rad23 homologues share a common domain structure, including a ubiquitin-like domain at the N-terminus and two copies of a highly conserved domain termed the Ub associated domain (UBA) located in the middle [UBA(1)] and at the C-terminus [UBA(2)] of the protein[27]. The solution structure of HHR23A revealed that the domain forms a compact three helix bundle with an unusually large hydrophobic surface patch (Figure 7).

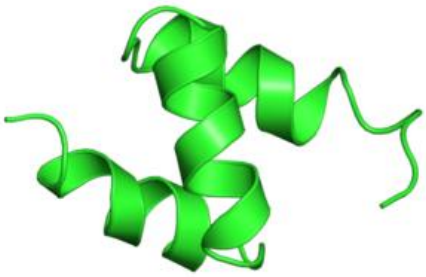


Figure 7. Solution structure of UBA2. Pdb code 1F4I.

UBA has been frequently used as a model domain to investigate biomolecular recognition by Ub and polyUb[28]. In particular, UBA was found to bind differently to

K48- an K63linked polyUb, enabling to explain the molecular determinants of substrate specificity by polyUb [29] (Figure 8).

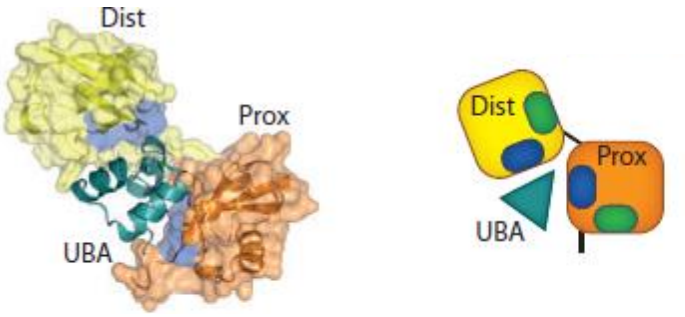


Figure 8. NMR model of the hHR23 UBA domain bound to Lys48-linked diubiquitin, pdb code 1zo6

3) Protein degradation and inhibition by the frameshift ubiquitin mutant Ubb^{+1}

The selective degradation of cellular proteins is a precisely regulated process required for various cellular functions including apoptosis, biogenesis of organelles, cell cycle and division, DNA transcription, modulation of cell surface receptors, ion channels and the secretory pathway, stress response to extracellular modulators and viral infection, and likely many other still undiscovered pathways. Responsibility for the breakdown of many short-lived regulatory proteins, enzymes, and structural proteins falls upon the ubiquitin/proteasome system. PolyUb chains act as a universal recognition signal that targets proteins for proteasomal degradation[30].

Regulation of the levels of cellular proteins is critical for maintaining cellular homeostasis. For example, most neurodegenerative diseases are characterized by abnormal protein folding, processing, and/or aggregation[31 32]. Specifically, impaired degradation leads to Alzheimer's and Parkinson's diseases[33], while overactive protein degradation has been linked to Cystic Fibrosis[34]. Also, the development of several types of cancer has been attributed to defects in the ubiquitin-proteasome pathway, which is integral to the destruction of cyclins that mediate cell cycle progression; such defects can lead to uncontrolled cell division and tumorogenesis[5].

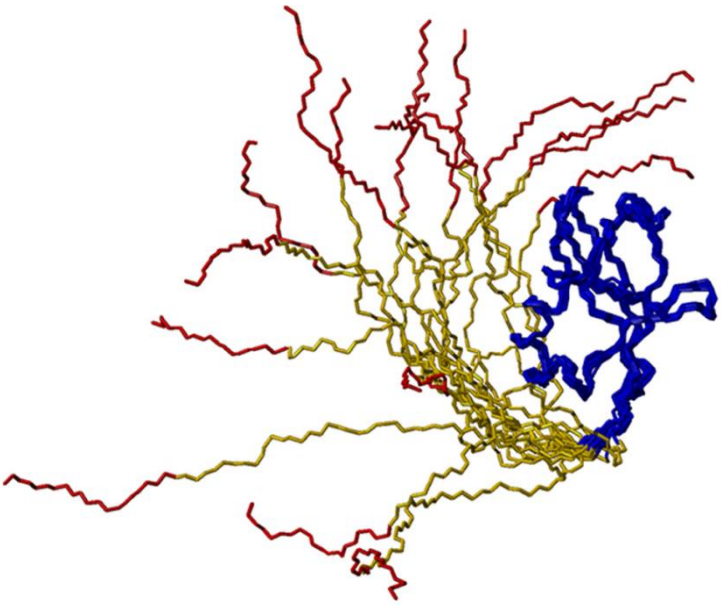


Figure 9. NMR structure of Ubb^{+1}. The 20 lowest energy structures were superimposed using the backbone atoms in the Ub region (residues 1–74). *Blue* (residues 1–74) and *yellow-red* (residues 75–95) *lines* indicate the Ub and C-terminal tail regions, respectively. The *yellow* (residues 75–88) and *red* (residues 89–95) *lines* indicate the residual structured and unstructured regions, respectively(Sunggeon Ko et al.,2010)

In Alzheimer's disease (AD) and Down syndrome (DS) patients, intracellular and extracellular deposits of proteins in tangles, neuropil threads, and neuritic plaques are correlated with neuronal dysfunction leading to dementia[35]. Ub is the first identified member of a large family of highly conserved eukaryotic proteins, which all share a similar structure and a carboxy-terminal diglycine motif[36]. Ub affects protein stability and protein sorting, small Ub like modifier molecules control nuclear transport, and interferon-stimulated gene-15 is involved in inflammatory and immune responses[36].

Ub immunopositive inclusions have been found in Alzheimer's disease (AD), Parkinson's disease (PD), Lewy body disease, polyglutamine diseases such as Huntington's disease and spinocerebellar ataxia, amyotrophic lateral sclerosis, progressive supranuclear palsy, Pick's disease and frontotemporal dementia[37]. An alteration in the UPS is likely to affect cellular functioning because the system not only has a pivotal role in protein quality control but also regulates numerous other cellular functions through the degradation of specific cell surface, cytoplasmic and nuclear proteins[35].

Ubb^{+1} was first identified as a frameshift mutant of the Ub B protein in the brains of neurodegenerative disease patients and is composed of a Ub moiety (75 residues) with a

19-residue C-terminal extension (Figure 9). The genes from which Ubb^{+1} mRNAs are transcribed contain several GAGAG motifs, and dinucleotide deletions (_GA) from within the GAGAG motif result in an abnormal C-terminal sequence.

The generation of this mutant Ub protein is unusual - the mutation is found in the messenger RNA, but not in the DNA sequence of the Ub-B gene. The mutant Ub results from a dinucleotide deletion near the 39 end of the mRNA transcript which shifts the reading frame for translation. The dinucleotide deletion event in the mRNA has been termed ‘‘molecular misreading’’, though the mechanism by which the deletion occurs remains elusive[38]. Misreading is a rare event, which was initially proposed to become more common with ageing; as adult neurones are post-mitotic they are likely to be particularly vulnerable to any deleterious effects of molecular misreading[38]. The products of misreading known as ‘+1 proteins’ potentially contribute to the initiation of neuropathological events; for example, APP+1, a mutant form of amyloid precursor protein (APP) has been implicated in the pathogenesis of AD[39]. Normally these aberrant proteins are removed by the Ub-proteasome system (UPS), which executes the proteolytic degradation of aberrant proteins via a Ub-tagging mechanism.

Ubb^{+1} inhibits the 26S proteasome in a dose-dependent manner, resulting in the accumulation of aberrant proteins[40] (Figure 10). The aberrant C terminus of Ubb^{+1} prevents the protein from covalently modifying other proteins as its Gly76 is no longer present, although Ubb^{+1} itself can still act as an acceptor of Ub and become covalently ubiquitinated with wild-type Ub. This effectively allows Ubb^{+1} to ‘cap’ poly-Ub chains making them unanchored, that is, nonsubstrate linked. These molecules are reportedly refractory to the deubiquitinating enzyme system. Consequently, when Ubb^{+1}-anchored polyUb is targeted to the 26 S proteasome, it acts as a functional antagonist, inhibiting the activity of the proteasome and leading to neurotoxicity[40].

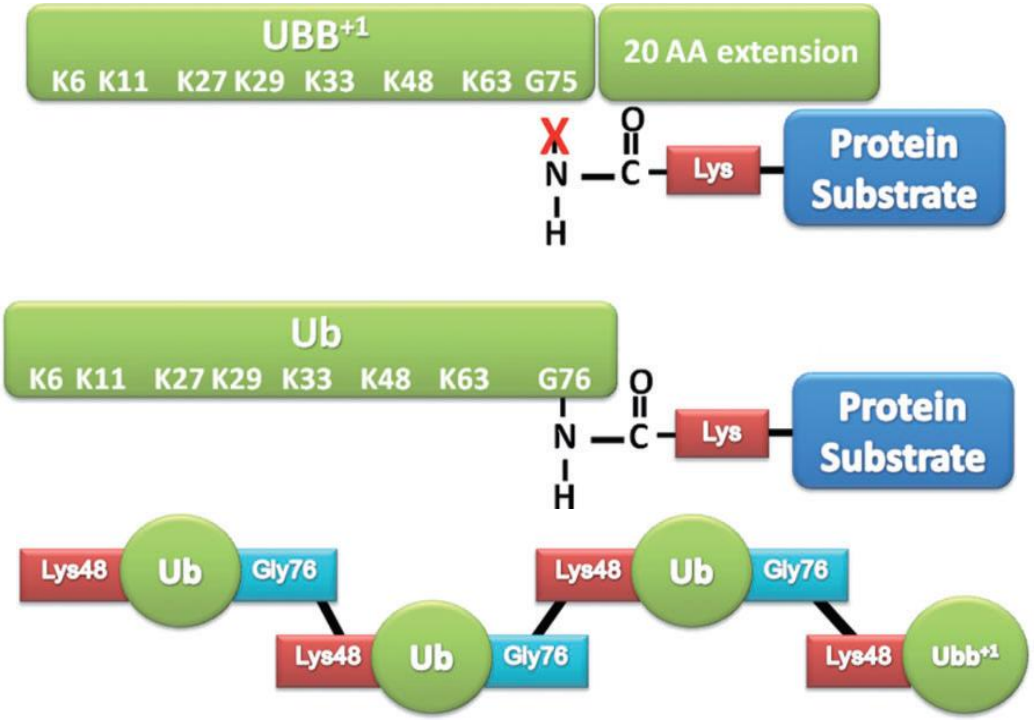


Figure 10. Ubb^{+1} is a frameshift ubiquitin mutant. (**A**) Top, the Ubb^{+1} frameshift ubiquitin (Ub) mutant protein lacks the C-terminal Gly76 residue which allows wild-type Ub to conjugate to substrate proteins (bottom). (**B**) Ubb^{+1} can still act as an acceptor for the conjugation of wild-type Ub (in this example using Lys48 and forming a homotypic chain) and allows Ubb^{+1} to effectively 'cap' free poly-Ub chains.

4) Synthetic cellular environment: macromolecular crowding

The intracellular environment is complex and difficult to study directly. The cellular environment significantly modulates the behavior of macromolecules, affecting their structure, dynamics, and stability[41]. It is widely recognized that the crowded conditions found in cellular environments can significantly impact the equilibria and kinetics of biochemical processes. Usually, *in vitro* protein studies are performed in dilute aqueous solutions, with a total protein concentration less than 10 g/L. The high concentrations of macromolecules found in the cell result in non-specific interactions between the protein of interest and the crowder components[42]. In reality, the characteristics of a cellular environments are macromolecular crowding, local viscosity, compartmentalization, and confinement[43]. The cytoplasmic medium is deeply crowded with concentrations of

macromolecules up to 400 g/L[4344]. Such high concentration of macromolecules (reduced available space), referred to as crowding, inevitably influences the probability of intermolecular encounters and shifts the binding equilibria (Figure 11). Studies based on simplified cell mimics could provide important insights into protein chemistry in native-like environments. The physical and chemical properties of molecular surfaces determine the nature of the interactions and dictate the attributes of molecular recognition[45]. Actually, weak nonspecific interactions could support or hinder functional intracellular communication. Furthermore, transient interactions with cytoplasmic components significantly and differently affect protein mobility[46]. To explain the effects of macromolecular crowding, two different types of interactions can be defined: hard and soft interactions. Hard interactions are referred to as a volume exclusion, in which the crowder agent is considered as inert: the crowder agent occupies a solution space that, in the absence of crowder, is available for the protein[47]. The solution volume to which the centre of mass of the protein does not have access due to the presence of the crowder is defined as covolume[48].

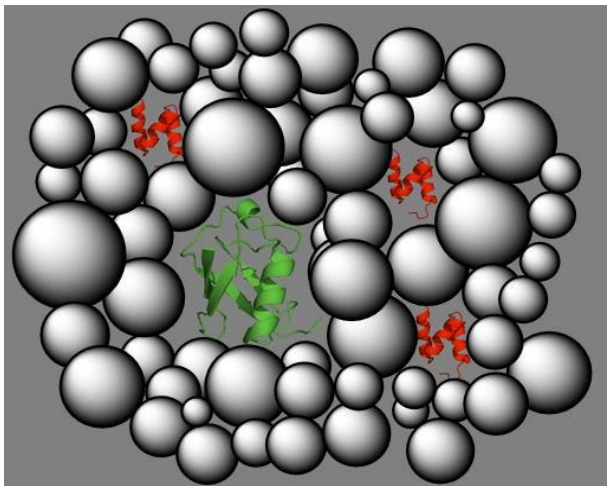


Figure 11. Representation of a crowding environment. Ubiquitin is shown in green and UBA2 in red.

Synthetic crowding agents such as Ficoll behave as relatively inert cosolutes and are convenient systems to investigate the effects of hard repulsion. Ficoll® is an important branched polysaccharide polymer with a variety of applications in biology and biophysics, it is highly soluble, it exhibits an average molecular mass of ~70 kDa and forms, at high concentrations, network-like structures of different viscosities[49](Figure 12). Soft interactions refer to the chemical interactions between the biomolecular crowding agent and the protein, considering the chemical nature of the molecules involved rather than simple steric hindrance[50].

An experimental study of transient interactions in the extremely complex and heterogeneous intracellular environment is a very challenging task. NMR spectroscopy is a particularly attractive tool for studying a protein's behavior in cells, because it provides information at the residue level. The success of in cell NMR however is still hampered by serious technical problems, including the difficulty to produce or internalize sufficient amounts of isotope-labelled proteins in living cells. Therefore, the use of cell-mimicking model systems still constitutes a preferred tool to decipher mechanistic aspects of protein function.

In part of this PhD thesis, we use cell mimics to provide novel insights into ubiquitin-effector protein interactions.

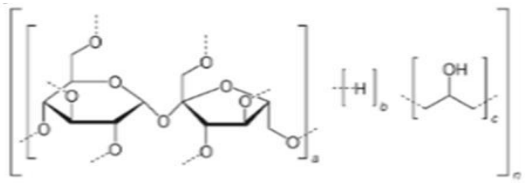


FIGURE 12. Poly(sucrose-co-epichlorhydrin) (Ficoll (74 kDa)) a macromolecular cosolutes (crowders) that mimic the intracellular milieu.

5) Biomolecular NMR spectroscopy

5.1) Protein NMR spectra

Nuclear magnetic resonance, or NMR is a phenomenon that occurs when the nuclei of certain atoms are immersed in a static magnetic field and exposed to a second oscillating magnetic field. In a general NMR experiment the sample nuclei are oriented by a strong magnetic field (in the order units of Tesla), absorb radiation at characteristic radiofrequencies (about 50-1000 MHz) and undergo an energetic transition between two states. The behaviour of the magnetization decay (FID, free induction decay), necessary to reach the equilibrium state, is converted from the acquired time-domain data to the frequency-domain signal by Fourier Transform. Nuclei of the same element in different environments give rise to distinct spectral lines because they adsorb radiation at

characteristic frequencies. The parameters that can be measured from the resulting spectra give information on molecular structure, conformation and dynamics. A very wide range of different elements have nuclei suitable for NMR spectroscopy. Hydrogen (^{1}H, the proton) is the most sensitive nucleus detected by NMR and it is by far the most important nucleus for the study of biological molecules. Other nuclei especially relevant for NMR analysis of proteins are nitrogen (^{15}N isotope) and carbon (^{13}C isotope) which unfortunately represent only 0.4% and 1%, respectively, of the naturally occurring isotopes. The main problem in NMR study of proteins is that these big molecules contain thousands of protons that determine the overlapping of signals in the ^{1}H spectrum. The most powerful approach to overcome this problem is the isotope labelling with ^{15}N and ^{13}C which, together with multidimensional heteronuclear NMR, which are crucial to extend NMR analysis to larger molecules. The isotope labelling, together with the availability of amount of proteins in the order of milligrams, has been possible thanks to the development of molecular biology techniques and to the ability of setting up over-expression systems.

Within the arsenal of NMR experiments, ^{1}H,^{15}N-HSQC spectra are an invaluable tool for protein analysis. Monitoring the perturbations of chemical shifts (chemical shift perturbation, CSP), linewidths and/or intensity changes of HN and 15N resonances upon binding is commonly used to map residues involved in binding sites and/or conformational rearrangements. We use Heteronuclear Single Quantum Coherence (1H-15N HSQC) experiments, which are the simplest and most useful two dimensional (2D) pulse sequence, to study chemical shift variations of protein signals in a residue-specific manner. The HSQC experiment is used in the field of protein NMR, especially for correlating the ^{15}N nucleus with its attached proton in the amidic groups of protein backbone, by exploiting the coupling between the two nuclei (JHN= 92 Hz). The ^{1}H-^{15}N HSQC represents the fingerprint of a protein, because the number and the position of each signal is specific for each protein sample.

5.2) Binding and exchange

The variation of resonance frequency of a nuclear spin, in the presence of an external magnetic field, allows the study of protein attributes by NMR spectroscopy. Chemical shift changes are sensitive indicators of interactions among biomolecules in simple buffer and in the presence of the crowder agents. The chemical shift (the resonance frequency position) of a given nucleus reports on its local chemical environment, the signal intensity depends on the number of nuclei resonating at a given frequency. Chemical shift values are quoted as parts per millions, or ppm. The chemical shift is a fundamental parameter in protein NMR and determined the NMR sensitivity as it gives separately detectable signals for the hundreds of protons that can therefore be distinguished and assigned. High NMR sensitivity of the NMR chemical to structure and environment is very much evident in proteins. Chemical shift dispersion arises because different residues experience different microenvironments: internal residues in globular proteins are shielded from the solvent and more in contact with other residues. Chemical shift dispersion is also observed among protons within the same residue, due to different local electron distribution. Protein NMR signals are expected to be perturbed by the presence of NPs or other ligands dispersed in the solution, depending on the nature of the protein–NP interaction and the physico-chemical properties of the particles.

Chemical exchange refers to any dynamic process in which a nucleus exchanges between two or more environments in which its NMR parameters (chemical shift, scalar or dipolar coupling, relaxation) differ. This has an important effect on the NMR spectra of an exchanging system. For example we consider the HSQC spectrum of a protein as ligand is titrated in. When the exchange rate is slow on the chemical shift timescale, or in other words when koff is significantly slower than the difference in Hz between the chemical shifts of free and bound protein, then as ligand is titrated in, the free signal gradually disappears and the bound signal appears, the intensities of the two peaks reflecting the concentrations of free and bound protein. On the other hand, when exchange is fast, i.e. when koff is much greater than the chemical shift difference, then the signals will move smoothly from their position in the free spectrum to those in the

bound spectrum, with the frequency of the signal at any titration point being the weighted average of free and bound shifts (Figure 13).

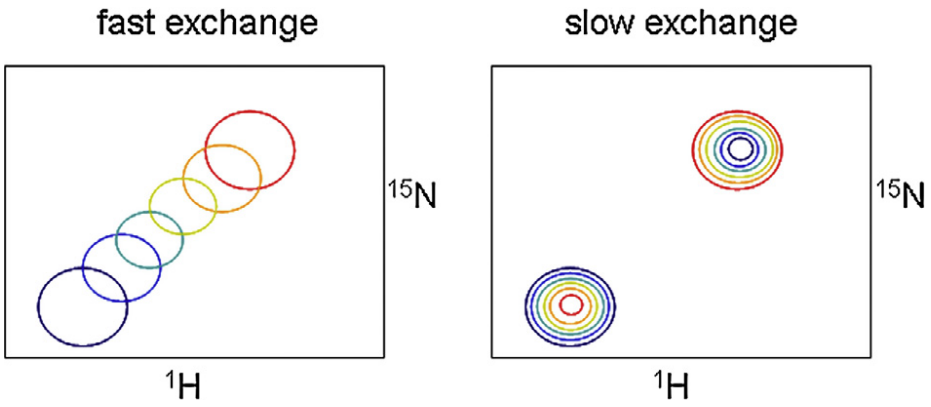


Figure 13. The dependence of two-dimensional NMR peak shape on exchange rate.
(Left) Fast exchange: peaks move smoothly from free (blue) to bound (red). In the limit of very fast exchange, peaks have the same shape throughout. As they move out of this limit, peaks may become broader when in equilibrium between free and bound, and then sharpen up again close to saturation. (Right) Slow exchange: the free peak (blue) decreases in intensity as the bound peak (red) increases (Mike P. Williamson, 2013).

5.3) Protein dynamics and spin relaxation

Protein dynamics can affect a wide range of functions, such as ligand binding, protein folding, aggregation, misfolding, effectiveness of small molecules inhibitors, etc. The term "protein dynamics" refers to time dependent fluctuations in structure. For obtaining dynamic information, we use the measurement of ^{15}N relaxation rates, since they can be related to global or local motions according to well-established theories[49]. Relaxation is the phenomenon by which the equilibrium is regained, i.e. it is relaxation which drives the spin to the equilibrium state along the direction of the applied magnetic field. It is possible to obtain information about the physical properties of a molecule and the sample (i.e. crowding) by studying the relaxation phenomena. Usually, NMR allows to investigate a wide range of time scales (picoseconds to milliseconds) based on the measurement of ^{15}N T_1 (longitudinal relaxation time constant) and ^{15}N T_2 (transverse relaxation time constant)[51]. If the sample is allowed to

be unperturbed for a long time in the magnetic field, it reaches a state of thermal equilibrium. This implies that all coherences are absent (M_0) and that the population are given by the Boltzmann distribution, at the temperature of the molecular environment:

$$\frac{N_{upper}}{N_{lower}} = e^{-\frac{\Delta E}{KT}}$$

Where N is the population in each state, T is the absolute temperature in Kelvin and K is the Boltzmann constant. Then, a radiofrequency (RF) pulse perturbs this equilibrium state, thus exciting the spin system and creating a coherence. Longitudinal relaxation time (T_1) is the time in which it is restored the equilibrium state (z-magnetization is returned to its equilibrium value), hence the spin gives back to the surrounding lattice the energy obtained from the RF pulse. Thus, R_1 ($1/T_1$) is the rate at which z magnetization is replaced and depends on the probability that the local field has a component oscillating at the Larmor frequency. Fluctuating fields parallel to the main magnetic field are responsible for the adiabatic contributions to the relaxation, generating variations in the total field along the z direction. Transverse relaxation time (T_2) defines the rate with which the magnetization, after being perturbed, decays in the plane perpendicular to the direction of the static magnetic field B_0. The main cause of transverse relaxation is loss of the phase coherence of spins obtained after a pulse resulting from local magnetic field fluctuations (mostly in the xy plane) affecting the Larmor frequency of individual spins. In fact, the local field varies from spin to spin, so the precession frequency is slightly different for each spins. The rate at which the coherence is lost is called R_2 ($1/T_2$).[52 53]

The extreme narrowing limit is when, at very short rotational correlation times, T_1 and T_2 are equal. When the correlation time increases, T_1 passes through a minimum and then increases. Instead, transverse relaxation constant continues decreasing. R_2 is also affected by conformational exchange phenomena in addition to fast dynamics.

The Nuclear Overhauser Enhancement or Effect (NOE) is a further important relaxation phenomenon in liquid-state NMR for the characterization of the structure and

dynamics of biomacromolecules. The heteronuclear NOE (hnNOE) results from through-space magnetization transfer via dipolar coupling between different types of nuclei. Experimentally, this typically involves transfer from ^{1}H to a directly attached ^{15}N or ^{13}C "heteronucleus." The hnNOE ranges from−4 to 1 for $\{^{1}H\}–^{15}N$ and from 1 to 3 for $\{^{1}H\}–\ ^{13}C$ and both are reduced in the presence of internal flexibility via $J(\omega H+\omega X)−J(\omega H−\omega X)$. The $\{^{1}H\}–^{15}N$ hnNOE can fall below zero due to the negative sign of the 15N gyromagnetic ratio γ.

5.4) Paramagnetic NMR

The association of paramagnetic complexes with a target protein or ligand, can be conveniently probed by monitoring paramagnetic-induced perturbations of NMR signals due to the altered spin relaxation properties of the nuclei around the paramagnetic metal center[54]. Paramagnetic metals have a small ad positive susceptibility to magnetic fields. These materials are slightly attracted by magnetic fields and do not retain the magnetic properties when the external field is removed. Paramagnetic properties are due to the presence of unpaired electrons, and to the realignment of the electron orbit caused by the external magnetic field.

Paramagnetic metals have large magnetic moments and relaxation times varying from 10^{-13}–10^{-8} sec (depending on the atomic number of the metal). The proximity of unpaired electrons to a nuclear spin generates local magnetic fields that are a potential source of paramagnetic relaxation enhancements. Interactions between the unpaired electron spin on a metal centre and the nuclear spins of surrounding atoms can shorten nuclear T_1 and T_2 relaxation times by many orders of magnitude[55]. The magnetic moment of the electron is very much greater than that of the proton, so the local field generated by an electron is correspondingly much greater.

In the proximity of the paramagnetic metal, longitudinal and transverse protein nuclear relaxation rates (R_1 and R_2 respectively) are the sum of a diamagnetic, $R_{1,2dia}$, and paramagnetic, $\Delta R_{1,2para}$ contribution. As shown in following equation, the relaxation rate enhancement effects depend on the metal-to nucleus distance, and they tend to vanish rapidly.

$$\Delta R_{1,2\,para} = \frac{K_{1,2}}{r^6}$$

Where K1 and K2 are dipolar constants specific for longitudinal and transverse relaxation rates, respectively, and depend on operating frequencies and the correlation time.

Effect on transversal relaxation rates can be indirectly obtained from changes in the NMR peak intensities of spectra of the paramagnetic sample, Ipara, compared with those of a diamagnetic reference, Idia according to the general relationship given by Equation below.

The first term of the product accounts for the intensity change produced by relaxation enhancement during the acquisition time of the NMR experiment (signal linewidth) and the second term refers to the effect operating during the time t of the pulse sequence in which the magnetization of the observed nucleus is in the transverse plane. [54]

$$\frac{I_{para}}{I_{dia}} = \frac{R_{2dia}}{R_{2dia} + \Delta R_{2para}}\, e^{-\Delta R_{2para} t}$$

In the presence of paramagnetic metal ions the nmr lines became broader on decreasing the distance between the metal and the observed nucleus. Therefore, generally there is a sphere around the metal ion in which proton NMR lines are to broad to be detected. The size of this shell depends on the nuclear relaxing capability of the metal ion, which in turn depends on the number of the unpaired electron, on the electron relaxation time, and on the rotational correlation time of the molecule.

In this PhD project, we use NMR spectroscopy to calculate R_1, R_2, NOE and paramagnetic relaxation enhancement of 15N-enriched protein Ubiquitin and Ubb^{+1}.

6) Nanoparticle receptors

Modulation of protein-protein interactions is a long-term goal in chemical biology. Selective binding of proteins by artificial receptors is a method of choice to interfere in cellular processes with encouraging applications in diagnostics and therapeutics. NP-based receptors are promising candidates for targeting protein surfaces, however this topic remains largely unexplored, particularly with respect to dynamic multidomain proteins. The rapid development of novel nanoparticulate materials for applications in many areas of bioscience necessitates an improved characterization of the nano-bio interfaces[56]. Rationalizing how nanoparticles (NPs) interact with biological systems constitutes an essential step toward the identification of biocompatible and bioadverse interactions[57]. In a biological medium, NPs may interact with biomolecules such as proteins, nucleic acids, lipids, and metabolites due to their nano-size and large surface-to-mass ratio. Of particular importance is the adsorption of proteins on the NP surface, as they profoundly influence NP biodistribution and bioreactivity. accumulating evidence indicates that proteins may display preferential orientations with respect to the NP surface, mediated by specific noncovalent chemical interactions. Therefore, understanding how NPs impact protein-based recognition processes will require knowledge about the presentation of functional biomolecular motifs at the NP-biomolecule interface.

Modulation of interactions among biomacromolecules using nanoparticle(NP)-based receptors capable of biomolecular recognition offers new possibilities for applications in therapeutics, diagnostics, and sensing. Indeed, specific noncovalent binding among biomolecules is critical in all cellular activities, and pathological conditions almost invariably involve some degree of protein miscommunication[58]. Therefore, protein-protein interactions (PPI) are a prime target for drug development and chemical biology research[59]. The development of small organic molecule modulators of PPI encounters difficulties owing to the fact that interaction surfaces between protein and ligand are small (300-1000 Å) compared to those involved in PPI (1500-3000 Å)[59]. Dynamic proteins are particularly challenging for the design of artificial receptors. PolyUb are multidomain proteins, found in dynamic conformational

ensembles, able to regulate a plethora of life-essential pathways. NPs are attractive receptors for such targets due to the large surface available for interaction and for countless modification possibilities.

In our project we aim at identifying the molecular determinants of polyUb recognition by selected NPs. We focus on lipid vesicles, silica nanoparticles and micelles as highly versatile candidate receptors. Lipid vesicles were used as model systems of lipid-based NPs to establish the conditions for obtaining lipid-protein recognition. Lipid-based membranes have attracted considerable attention due to their potential application as tools to probe cellular and molecular interactions and as bioactive coatings for biosensor or medical implant applications[60 61 62].

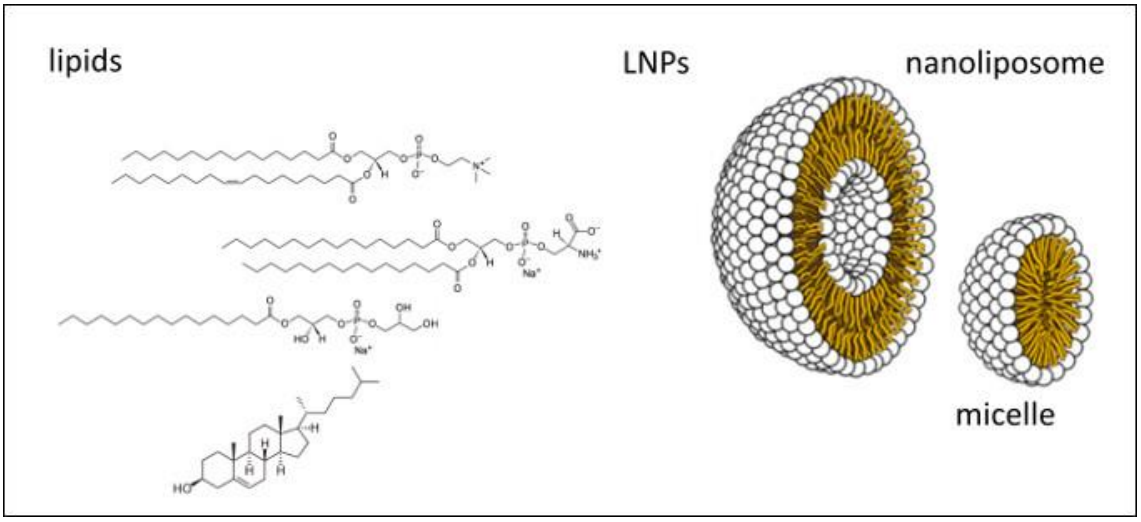


Figure 14. Examples of available lipids to be used for the production of LNPs and morphology of LNPs.

Liposomes are the most used membrane mimetic system, distinguishing variety of vesicle types exists, differing in the degree of multilayering, size, lipid composition, net charge, and behaviour in transition phase. Liposomes are artificial spherical vesicles formed by self – assembly of lipids, consisting of a closed lipid bilayer which encompasses an aqueous core[63]. On the basis of their internal organization they can be divided into different classes. Among them, the most frequently used are the small unilamellar vesicles (SUVs) (diameter range from 15nm to 50nm) and large unilamellar

vesicles (LUVs) (diameter range from 50nm to 200nm). A vast literature describes the preparation and characterization of lipid vesicles[64 65]. We are currently using the thin-film hydration method for the production of LNP in aqueous solution,[66] followed by downsizing by sonication (to obtain 30nm unilamellar vesicles) and extrusion (100 nm vesicles). The lipid head groups may carry positive or negative net charges, or neutral dipoles. Liposomes used in the clinic generally contain phosphatidylcholine and cholesterol, which were thus incorporated in our formulations. The fluidity of the lipid bilayer can be modulated by changing phosphatidylcholine/cholesterol molar ratios and by varying the content of saturated/unsaturated fatty acyl chains. Charged liposomes are obtained by replacing part of the neutral lipids with anionic (e.g. phosphatidylglycerol) or cationic (e.g.DOTAP) molecules. In our work, we investigated the binding specificity of chemically diverse vesicles towards structurally distinct polyUb using site-resolved solution NMR experiments for observation of protein signals, together with ancillary biophysical techniques.

Silica nanoparticles (SNP) are attracting considerable interest for several biomedical applications due to their low toxicity, biocompatibility and scalable synthetic availability. The many possible surface modifications of silica nanoparticles allow precise control of surface chemistry to modulate drug or chemical loading, nanoparticle dispersion, blood circulation, and site specific targeting. The ability to combine these properties makes silica nanoparticles a desirable platform for biomedical imaging, assaying, therapeutic delivery, monitoring, and ablative therapies[57]. We used SNP to probe the interaction of polyUb with particles of similar size as SUV but different surface properties.

As a micellar system, we selected an amphiphlic molecule incorporating the heptadentate ligand 6-amino-6-methylperhydro-1,4 diazepinetetraacetic acid (AAZTA), which is able to chelate gadolinium ions. This complex displays excellent relaxation enhancement properties and it shows good thermodynamic and kinetic stability. *Gianolio et al.* have conjugated the Gd-AAZTA complex with a long aliphatic chain (C17). The presence of the aliphatic chain induces the formation of micelles (average diameter 5.5 nm) already at submillimolar concentrations (0.108 mM).

Unraveling the basis of molecular recognition in the polyUb/NP system will pave the way to the development of polyUb-targeted NP-based receptors aimed at interfering with fundamental cellular communication pathways.

7) Aim of the thesis project and brief summary of the research articles

The aim of this doctoral project is to provide new atomic-level insights into the determinants of molecular recognition of the protein ubiquitin. We focused on the transient interactions of monoubiquitin with a representative domain of a receptor protein under macromolecular crowding, we explored the structural, dynamic and interaction properties of an aberrant mutant involved in neurodegenerative diseases, and we investigated the adsorption of polyUb chain isomers to nanoparticle receptors. In our work, we made extensive use of solution NMR spectroscopy in order to study transient interactions and conformational dynamics at atomic resolution.

The traditional methods of studying proteins are, in general, implemented within diluted solutions with less than 10 g/L of total protein concentration. This low concentration allows to obtain good signals but more biological information is needed[44]. One of the distinctive features of cellular systems is that the cytoplasmic medium is deeply crowded with significant concentrations of macromolecules (50-400 g/L) which affect several protein attributes. This high concentration of macromolecules in the cell interior is known as crowding[67]. The presence of high concentrations of crowding agents in the biological systems can result in non-specific interactions between the protein of interest and the target protein. The broad aim of the study is to understand the effects of crowding on protein-protein interactions. In particular, we focused on the Ub-UBA association and investigated the perturbations induced by the presence of an 'inert' synthetic crowding agent (Ficoll 70) in comparison with dilute solution. We wished to determine the differences in binding affinity, structure and dynamics of the complex dissolved in the different media. The Ub-UBA system is relatively small and therefore suitable for detailed NMR investigation, our primary method of investigation. The results obtained in the study (Research Article 1)[42] showed that the stereospecific complex of ubiquitin and the ubiquitin-associated domain (UBA) is minimally

perturbed by the crowding agent Ficoll. However, in addition to the primary canonical recognition patch on ubiquitin[12], secondary patches were identified, indicating that in cell-mimicking crowded solution, UBA contacts ubiquitin at multiple sites.

Specific noncovalent binding among biomolecules is critical in all cellular activities[58]. Pathological conditions almost invariably involve some degree of protein miscommunication. Therefore, protein-protein interactions (PPIs) are a prime target for drug development and chemical biology research[59]. The development of small organic molecule modulators of PPIs encounters several difficulties owing to the fact that interaction surfaces between protein and ligand are small compared to those involved in PPIs[59]. Nanoparticles (NPs) offer an additional intriguing possibility to target proteins, as they display large surfaces that can be decorated with an ample choice of functionalities[68].

Mechanisms of protein recognition have been extensively studied for single-domain proteins, but are less well characterized for dynamic multidomain systems. PolyUb represent an important multidomain system that requires recognition by structurally diverse ubiquitin-interacting proteins.[7] Fundamental cellular pathways are regulated by eight structurally distinct polyUb types, two of which have been extensively studied. Lysine(K)48-linked polyUb target proteins for proteasomal degradation, whereas K63-linked polyUb have multiple non-degradative roles[13].

In our project, we aim to explore NP systems for the development of polyUb chain-specific receptors. We selected lipid vesicles as highly versatile systems to be used as a model platform for understanding the determinants of recognition. We investigated the binding specificity of chemically diverse NPs towards structurally distinct polyUb. The results of this project (Research Article 2) constitute the basis for an improved understanding of polyUb recognition by artificial receptors and for the development of NP-based therapeutic strategies.

Dysfunction of Ub-related enzymes and pathways has been linked to the pathogenesis of severe human diseases, including Alzheimer's (AD), Parkinson's (PD), Huntington's (HD) diseases, Amyotrophic Lateral Sclerosis (ALS), cancer, metabolic

syndromes, and genetic disorder. In particular, a mutated version of Ub, named Ubb^{+1}, was found specifically accumulated in neurofibrillary tangles, neuropil threads and dystrophic neurites in brain tissues of AD patients. Ubb^{+1} maintains the well-structured globular domain of Ub and its lysine side chains can act as acceptors for polyUb chain linkage. However, as a result of the lack of the C- terminal Gly76, Ubb^{+1} does not modify substrate proteins and instead terminates the elongation of polyUb chains. The resulting Ubb^{+1}-capped polyUb chains ($polyUbb^{+1}$) were shown to inhibit the proteasome, and were recalcitrant to disassembly mediated by deubiquitinating enzymes.

Due to a possible participation of Ubb^{+1} in the molecular events leading to neurotoxicity and neurodegeneration in AD, there is large interest in elucidating the structural details of this frameshift mutant Ub protein and the consequent functional differences with respect to the wild-type protein. In our work (Research Article 3), we investigated structural and dynamic features of Ubb^{+1} using NMR spectroscopy methods that are particularly suited to explore protein molecules containing flexible domains such as the C-terminal extension of Ubb^{+1}. Calorimetry measurements were carried out to evaluate the contribution of the tail to the protein's thermal stability in both its monomeric and dimeric forms. Finally, the ability of Ubb^{+1} to interact with different biomolecules, in particular the UBA2 domain of the human homologue of the yeast DNA repair protein RAD23 (HHR23A) and membrane mimics, was explored and characterized.

References

1. Pickart CM. Ubiquitin enters the new millennium: Meeting review. *Mol Cell*. 2001;8(3):499-504. doi:10.1016/S1097-2765(01)00347-1.

2. Goldstein G, Scheid M, Hammerling U, Schlesinger DH, Niall HD, Boyse EA. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci*. 1975;72(1):11-15. doi:10.1073/pnas.72.1.11.

3. Pickart CM. Echanisms nderlying ubiquitination. *Annu Rev Biochem*. 2001;70(1):503-533. doi:10.1146/annurev.biochem.70.1.503.

4. Pickart CM. Back to the Future with Ubiquitin. *Cell*. 2004;116(2):181-190. doi:10.1016/S0092-8674(03)01074-2.

5. Hersko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem*. 1998;67:425-479.

6. Callis J. The Ubiquitination Machinery of the Ubiquitin System. *Arab B*. 2014;12:e0174. doi:10.1199/tab.0174.

7. Ye Y, Blaser G, Horrocks MH, et al. Ubiquitin chain conformation regulates recognition and activity of interacting proteins. *Nature*. 2012;492(7428):266-270. doi:10.1038/nature11722.

8. Alfano C, Faggiano S, Pastore A. The Ball and Chain of Polyubiquitin Structures. *Trends Biochem Sci*. 2016;41(4):371-385. doi:10.1016/j.tibs.2016.01.006.

9. Mastellos DC, Yancopoulou D, Kokkinos P, et al. HHS Public Access. 2016;45(4):423-440. doi:10.1111/eci.12419.Compstatin.

10. Pickart CM, Fushman D. Polyubiquitin chains: Polymeric protein signals. *Curr Opin Chem Biol*. 2004;8(6):610-616. doi:10.1016/j.cbpa.2004.09.009.

11. Hicke L. Protein Regulation By Monoubiquitin. *Nat Rev Mol Cell Biol*. 2001;2(March):195-201. doi:10.1038/35056583.

12. Varadan R, Assfalg M, Haririnia A, Raasi S, Pickart C, Fushman D. Solution Conformation of Lys63-linked Di-ubiquitin Chain Provides Clues to Functional Diversity of Polyubiquitin Signaling. *J Biol Chem*. 2004;279(8):7055-7063. doi:10.1074/jbc.M309184200.

13. Komander D, Rape M. The Ubiquitin Code. *Annu Rev Biochem*. 2012;81(1):203-229. doi:10.1146/annurev-biochem-060310-170328.

14. Baumeister W, Walz J, Zühl F, Seemüller E. The proteasome: Paradigm of a self-compartmentalizing protease. *Cell*. 1998;92(3):367-380. doi:10.1016/S0092-8674(00)80929-0.

15. Voges D, Zwickl P, Baumeister W. the 26S P Roteasome : a Molecular Machine Designed for Controlled. *Annu Rev Biochem*. 1999;68:1015-1068. doi:10.1146/annurev.biochem.68.1.1015.

16. Lecker SH. Protein Degradation by the Ubiquitin-Proteasome Pathway in Normal and Disease States. *J Am Soc Nephrol*. 2006;17(7):1807-1819. doi:10.1681/ASN.2006010083.

17. Kisselev AF, Akopian TN, Woo KM, Goldberg AL. The Sizes of Peptides Generated from Protein by Mammalian 26 and 20S Proteasomes. *J Biol Chem*. 1999;274(6):3363-3371. doi:10.1074/jbc.274.6.3363.

18. Benaroudj N, Zwickl P, Seemüller E, Baumeister W, Goldberg AL. ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. *Mol Cell*. 2003;11(1):69-78. doi:10.1016/S1097-2765(02)00775-X.

19. Beal R, Deveraux Q, Xia G, Rechsteiner M, Pickart C. Surface hydrophobic residues of multiubiquitin chains essential for proteolytic targeting. *Proc Natl Acad Sci*. 1996;93(2):861-866. doi:10.1073/pnas.93.2.861.

20. Wang Y, Li C, Pielak GJ. Effects of Proteins on Protein Diffusion - Journal of the American Chemical Society (ACS Publications). *J Am Chem Soc*. 2010;(16):9392-9397. doi:10.1021/ja102296k.

21. Peng J, Schwartz D, Elias JE, et al. A proteomics approach to understanding protein ubiquitination. TL - 21. *Nat Biotechnol*. 2003;21 VN-r(8):921-926. doi:10.1038/nbt849.

22. Piotrowski J, Beal R, Hoffman L, Wilkinson KD, Cohen RE, Pickart CM. Inhibition of the 26 S proteasome by polyubiquitin chains synthesized to have defined lengths. *J Biol Chem*. 1997;272(38):23712-23721. doi:10.1074/jbc.272.38.23712.

23. Hampton RY. Fusion-based strategies to identify genes involved in degradation of a specific substrate. *Methods Enzymol*. 2005;399(5):310-323. doi:10.1016/S0076-6879(05)99002-2.

24. Ubiquitin chain synthesis. (2005). 2018;(2005):2018.

25. Medina B, Paraskevopoulos K, Boehringer J, et al. The ubiquitin-associated (UBA) 1 domain of Schizosaccharomyces pombe Rhp23 is essential for the recognition of ubiquitin-proteasome system substrates both in vitro and in vivo. *J Biol Chem*. 2012;287(50):42344-42351. doi:10.1074/jbc.M112.419838.

26. Mueller TD, Feigon J. Solution structures of UBA domains reveal a conserved hydrophobic surface for protein-protein interactions. *J Mol Biol*. 2002;319(5):1243-1255. doi:10.1016/S0022-2836(02)00302-9.

27. Gardner EM, Maravi ME, Rietmeijer C, Arthur J, Burman WJ, Health DP. NIH Public *References* Access. 2009;6(3):145-155. doi:10.1038/nn.2120.Red-shifted.

28. Raasi S, Varadan R, Fushman D, Pickart CM. Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat Struct Mol Biol*. 2005;12(8):708-714. doi:10.1038/nsmb962.

29. Varadan R, Assfalg M, Raasi S, Pickart C, Fushman D. Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Mol Cell*. 2005;18(6):687-698. doi:10.1016/j.molcel.2005.05.013.

30. Pickart CM. Ubiquitin in chains. *Trends Biochem Sci*. 2000;25(11):544-548. doi:10.1016/S0968-0004(00)01681-9.

31. Taylor JP. Toxic Proteins in Neurodegenerative Disease. *Science (80-).* 2002;296(5575):1991-1995. doi:10.1126/science.1067122.

32. Michalik A, Van Broeckhoven C. Pathogenesis of polyglutamine disorders: aggregation revisited. *Hum Mol Genet*. 2003;12(suppl 2):R173-R186. doi:10.1093/hmg/ddg295.

33. Liu Y, Lansbury P. The UCH-L1 Gene Encodes Two Opposing Enzymatic Activities that Affect a-Synuclein Degradation and Parkinson's Disease Susceptibility. *Cell*. 2002;111:209-218. doi:S0092867402010127.

34. Ward CL, Omura S, Kopito RR. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell*. 1995;83(1):121-127. doi:10.1016/0092-8674(95)90240-6.

35. Duyckaerts C. Looking for the link between plaques and tangles. *Neurobiol Aging*. 2004;25(6):735-739. doi:10.1016/j.neurobiolaging.2003.12.014.

36. Isaacson MK, Ploegh HL. Ubiquitination, Ubiquitin-like Modifiers, and Deubiquitination in Viral Infection. *Cell Host Microbe*. 2009;5(6):559-570. doi:10.1016/j.chom.2009.05.012.

37. Fischer DF, De Vos R a I, Van Dijk R, et al. Disease-specific accumulation of mutant ubiquitin as a marker for proteasomal dysfunction in the brain. *FASEB J*. 2003;17(14):2014-2024. doi:10.1096/fj.03-0205com.

38. Leeuwen F, Hol E, H. Burbach P. *Mutations in RNA: A First Example of Molecular Misreading in Alzheimer's Disease*. Vol 21.; 1998. doi:10.1016/S0166-2236(98)01280-6.

39. George DR, Whitehouse PJ, D'Alton S, et al. Pathological Cascade. *Arch Med Res*. 2012;12(3):173-189. doi:10.1146/annurev-neuro-061010-113613.Amyloid.

40. Ko S, Kang GB, Song SM, et al. Structural basis of E2-25K/UBB+1 interaction leading to proteasome inhibition and neurotoxicity. *J Biol Chem*. 2010;285(46):36070-36080. doi:10.1074/jbc.M110.145219.

41. Pickart CM, Cohen RE. Proteasomes and their kin: proteases in the machine age. *Nat Rev Mol Cell Biol*. 2004;5(3):177-187. doi:10.1038/nrm1336.

42. Castañeda C, Liu J, Chaturvedi A, Nowicka U, Cropp TA, Fushman D. Nonenzymatic assembly of natural polyubiquitin chains of any linkage composition and isotopic labeling scheme. *J Am Chem Soc*. 2011;133(44):17855-17868. doi:10.1021/ja207220g.

43. Theillet F-X, Binolfi A, Frembgen-Kesner T, et al. Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem Rev*. 2014;114(13):6661-6714. doi:10.1021/cr400695p.

44. Charlton LM, Barnes CO, Orans J, Young GB, Pielak GJ, Li C. Residue-Level Interrogation of Macromolecular Crowding Effects on Protein Stability Residue-Level Interrogation of

Macromolecular Crowding Effects on Protein Stability. 2008;(13):6826-6830. doi:10.1021/ja8005995.

45. Keskin O, Gursoy A, Ma B, Nussinov R. Principles of protein-protein interactions: What are the preferred ways for proteins to interact? *Chem Rev*. 2008;108(4):1225-1244. doi:10.1021/cr040409x.

46. Wang Q, Zhuravleva A, Gierasch LM. Exploring weak, transient protein-protein interactions in crowded in vivo environments by in-cell nuclear magnetic resonance spectroscopy. *Biochemistry*. 2011;50(43):9225-9236. doi:10.1021/bi201287e.

47. Zhou H-X, Rivas G, Minton AP. Macromolecular crowding and confinemment: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys*. 2008;37:375-397. doi:10.1146/annurev.biophys.37.032807.125817.Macromolecular.

48. Li C, Pielak GJ. Using NMR to distinguish viscosity effects from nonspecific protein binding under crowded conditions. *J Am Chem Soc*. 2009;131(4):1368-1369. doi:10.1021/ja808428d.

49. Georgalis Y, Philipp M, Aleksandrova R, Krüger JK. Light scattering studies on Ficoll PM70 solutions reveal two distinct diffusive modes. *J Colloid Interface Sci*. 2012;386(1):141-147. doi:10.1016/j.jcis.2012.07.062.

50. Miklos AC, Li C, Sharaf NG, Pielak GJ. Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry*. 2010;49(33):6984-6991. doi:10.1021/bi100727y.

51. Kay LE, Torchia DA, Bax A. Backbone Dynamics of Proteins As Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry*. 1989;28(23):8972-8979. doi:10.1021/bi00449a003.

52. Tekely P. Malcolm H. Levitt. Spin dynamics: basics of nuclear magnetic resonance. John Wiley & Sons, Chichester, UK, 2001, 686 pp. Price £34.95. ISBN 0-471-48921-2. *Magn Reson Chem*. 2002;40(12):800-800. doi:10.1002/mrc.1092.

53. Hore PJ. *Nuclear Magnetic Resonance*. Oxford University Press, USA; 2015.

54. D'Onofrio M, Gianolio E, Ceccon A, et al. High relaxivity supramolecular adducts between human-liver fatty-acid-binding protein and amphiphilic GdIII complexes: Structural basis for the design of intracellular targeting MRI probes. *Chem - A Eur J*. 2012;18(32):9919-9928. doi:10.1002/chem.201103778.

55. Metalloproteins E. *Encyclopedia of Metalloproteins*; 2013. doi:10.1007/978-1-4614-1533-6.

56. Nel AE, Mädler L, Velegol D, et al. Understanding biophysicochemical interactions at the nano-bio interface. *Nat Mater*. 2009;8(7):543-557. doi:10.1038/nmat2442.

57. Liberman a., Mendez N, Trogler WC, Kummel a. C. Synthesis and surface functionalization of silica nanoparticles for nanomedicine. *Surf Sci Rep*. 2014;69(2-3):132-158. doi:10.1016/j.surfrep.2014.07.001.Synthesis.

58. Blackstone NW. Molecular Biology of the Cell. Fourth Edition. By Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts , and Peter Walter. *Q Rev Biol*. 2003;78(1):91-92. doi:10.1086/377840.

59. Wells JA, McClendon CL. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature*. 2007;450(7172):1001-1009. doi:10.1038/nature06526.

60. Barroso RP, Riske KA, Henriques VB, Lamy MT. Ionization and structural changes of the DMPG vesicle along its anomalous gel-fluid phase transition: A study with different lipid concentrations. *Langmuir*. 2010;26(17):13805-13814. doi:10.1021/la101784w.

61. Mashaghi S, Jadidi T, Koenderink G, Mashaghi A. *Lipid Nanotechnology*. Vol 14.; 2013. doi:10.3390/ijms14024242.

62. Caracciolo G. Liposome-protein corona in a physiological environment: Challenges and opportunities for targeted delivery of nanomedicines. *Nanomedicine Nanotechnology, Biol Med*. 2015;11(3):543-557. doi:10.1016/j.nano.2014.11.003.

63. Lian T, Ho RJ. Trends and developments in liposome drug delivery systems. *J Pharm Sci*. 2001;90(6):667-680. doi:10.1002/jps.1023.

64. Feix JB. Liposomes: A Practical Approach. Second Edition. Practical Approach Series, Volume 264. Edited by Vladimir Torchilin and Volkmar Weissig. *Q Rev Biol*. 2004;79(2):205-206. doi:10.1086/423065.

65. Karn PR, Cho W, Park HJ, Park JS, Hwang SJ. Characterization and stability studies of a novel liposomal cyclosporin a prepared using the supercritical fluid method: Comparison with the modified conventional Bangham method. *Int J Nanomedicine*. 2013;8:365-377. doi:10.2147/IJN.S39025.

66. Akbarzadeh A, Rezaei-Sadabady R, Davaran S, et al. Liposome: classification, preparation, and applications. *Nanoscale Res Lett*. 2013;8(1):102. doi:10.1186/1556-276X-8-102.

67. Hong J, Gierasch LM. Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J Am Chem Soc*. 2010;132(30):10445-10452. doi:10.1021/ja103166y.

68. Walker M, Kublin JG, Zunt JR. Engineering the Nanoparticle-Protein Interface: Applications and Possibilities. *Curr Opin Chem Biol*. 2009;42(1):115-125. doi:10.1086/498510.Parasitic.

CHAPTER 2:
Material and Methods

2.1) Materials for protein expression

2.1.1) Plasmids

The plasmid pET3a is a 4640bp expression vector based on the strong T7 promoter which is recognized by T7 RNA polymerase. T7 RNA polymerase is so selective and active that, when fully induced, almost all of resources of the cell are converted to target gene expression: the desired product can often account for more than 50% of the total cell protein. We use this plasmid for the expression of wild-type ubiquitin, D77, K48R and K63R ubiquitin.

The Figure 15 shows the vector map of pET3a and the elements characterizing this vector.

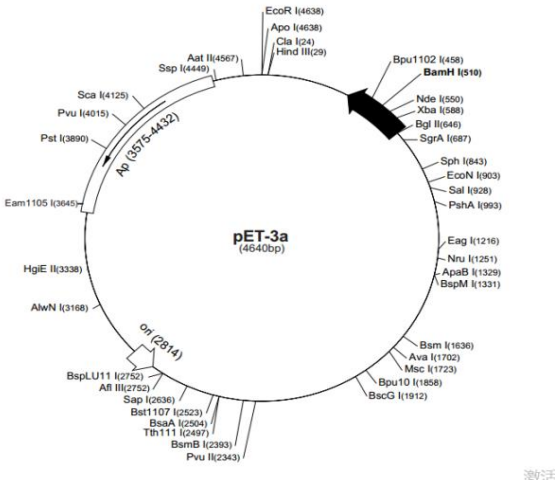


Figure 15. pET3a vector map

2.1.2) Bacterial strains

The plasmid pET3a were inserted in the Rosetta(DE3) host strains. They are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *Rosetta*. These strains supply tRNAs for 7 rare codones (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) on a compatible chloramphenicol-resistant plasmid.

2.1.4) Heat shock transformation

The following protocol has been used to transform *Rosetta* competent cells through heat shock:

• 50µl of competent cells stored at -80°C were placed in ice for thawing

• 30-50ng of DNA were added and the culture was incubated in ice for 30 minutes

• the heat shock was carried out by placing the culture in a water bath at 42°C for 45 seconds without shaking

• the culture was immediately placed in ice for 2 minutes

• 500µl of LB medium without antibiotic was added directly to the cells and then the culture was incubated at 37°C for 1 hour with shaking (200-250rpm)

• the cells were plated on selective LB agar and incubated overnight at 37°C

2.2) Protein sample preparation

2.2.1) Glycerol stock

After transformation, bacterial colonies with plasmid coding for mutant protein were grown in 5ml LB medium with antibiotic at 37°C for 6 hours. A cell solution with a final concentration of 30% glycerol was prepared, vortexed and cooled in liquid nitrogen. Glycerol stocks were stored at -80°C.

2.2.2) Culture growing media

The following media have been used for *Rosetta* growth. Sterilisation was achieved in autoclave and media for bacterial growth were prepared as follows if not otherwise specified. Selective media were prepared adding chloramphenicol, ampicillin or kanamycin.

Luria-Bertani (LB) medium

Compound*	**Quantity**
Triptone	10g
Yeast extract	5g
NaCl	10g

pH was adjusted to 7.0. After sterilisation,antibiotic was added.

The M9 medium is used to achieve uniform ^{15}N isotope incorporation in proteins. This medium is prepared as follows:

For 1 L final volume:

Compound	**Quantity**
Sterile water	780ml
M9 salts	200ml
$MgSO_4$ (stock solution 1M)	2ml
$ZnSO_4$ (stock solution 50mM)	2ml
$CaCl_2$ (stock solution 0,1M)	1ml
$FeCl_3$ (stock solution 10mM)	100µl
Glucose (stock solution 20%)	40ml
NH_4Cl*	**1g**

M9 salts (5x)

For 1L final volume:

Compound	**Quantity**
Na_2HPO_4	33,9g
KH_2PO_4	15g
NaCl	2,5g

The solution was sterilized by autoclave.

2.2.3) Protein expression and cell lysis

The expression of unlabelled proteins was achieved by adding 10ml of the starter culture to 1 litre of LB medium, then incubated in a shaker at 37°C till an OD600 value of 0.6 was reached. Protein expression was induced adding IPTG at a final concentration of 0.5mM and leaving the culture in the shaker at the best temperature and time found in the preliminary protein expression evaluation. After a centrifugation step at 8000g for 15 minutes, the pellet was resuspended in 2.5ml/g pellet of lysis buffer and sonicated at the maximum sonicator power for 7 cycles of 1 minute(except for the

enzyme where we add more lysozyme), alternating the sample sonication time to an equal period on ice. The lysate was then centrifuged for 10 minutes at 10000 g at 4°C and the supernatant solution was collected. The supernatant, containing the soluble protein fraction, was then loaded on the first purification column. M9 medium was used for expression of 15N labelled proteins and the same purification protocol was performed.

Lysis buffer Ub, K63R, K48R, D77, GB1, UBA2:
Tris-HCl 50mM, pH 7.6
Triton X-100 0.02% (v/v)
PMSF 1mM
Lysozyme 0.4mg/ml
DNasi 1 (20µg/ml)
MgCl2 10mM

Lysis buffer Ubc13, E1, E2:
Tris-HCl 20mM, pH 7.4
NaCl 150mM
Triton X-100 0.02% (v/v)
PMSF 1mM
Lysozyme 0.4mg/ml
DNasi 1 (20µg/ml)
MgCl2 10mM

Due to the high sensibility of the enzyme Mms2, for the lysis buffer we used a different protocol with less stringent condition in respect to the previous purification. The buffer is composed as follow:

Lysis buffer Mms2:
Tris-HCl 20mM, pH 8.0
NaCl 50mM

inibitori di Proteasi (50µl)
0.1% Nonidet P-40
Lysozyme 0.4mg/ml
DNasi 1 20µg/ml
MgCl2 10mM

2.3) Protein and enzyme purification

Ub was purified by precipitation of bacterial proteins through addition of a few drops of 70% perchloric acid until pH drops to 4.5 while stirring to cell lysate. Centrifuge suspension at 18,000 rpm for 20 minutes. The supernatant was dialyze against 50 mM Sodium Acetate, pH 4.5 overnight. The supernatant was loaded into a 15 mL SP column pre-equilibrated with 50mM Sodium Acetate pH 4.5. The column was washed by adding 1-2 column volumes of 50mM Sodium Acetate, pH 4.5. Ubiquitin was elutes with a 200mL linear gradient from 0-0.5M NaCl in 50mM Sodium Acetate pH 4.5. Finally, ubiquitin was dialyse into 10mM KPi pH 6.8 buffer.

UBA(2) domain of hHR23A was produced as GST-fusion protein as described by Raasi *et al.*(JMB, 2004, 341, 1367) and purified using Glutathione Sepharose (GE Healthcare) according to standard protocols. The GST tag was then removed by cleavage with thrombin followed by gel-filtration on a Superdex-75 (GE Healthcare). UBA2 was then dialyzed into final NMR buffer.

GB1 (streptococcal protein GB1 domain) protein was produced in *BL21(DE3) Escherichia coli* cells as His-tagged recombinant protein in fusion with C-terminal Green Fluorescent Protein (GFP). Protein purification was performed by IMAC chromatography using a Ni^{2+} charged-chelating sepharose (GE Healthcare). Next, GFP was cleaved from His-GB1 by TEV protease (Sigma Aldrich) and removed by size exclusion chromatography (Sephacryl S-100 HR, GE Healthcare).

E2-25k, Ubc13 were expressed as GSTfused proteins in the medium described above. Cell lysis was performed in PBS buffer, pH 7.4 complete with protease inhibitors, lysozyme and Dnase1. The lysate was centrifuged and the soluble extract was filtered and loaded onto Glutathione Sepharose beads (Molecular Probes, 10 ml bead suspension per liter culture) preequilibrated with PBS buffer, pH 7.4.

The unbound proteins were washed out with 6-8 bed volumes of PBS buffer, and the fusion protein was eluted with 50mM Tris pH 8, 10mM glutathione. GST-E2 was used as a fusion protein; however, the GST tags of Ubc13 were cleaved using thrombin.

E1, Mms2 were expressed as His-tagged recombinant protein. Protein purification was performed by IMAC chromatography using a Ni^{2+} charged-chelating sepharose (GE Healthcare). The elution was made with 50mM imidazole, Tris20mM NaCl0.5M buffer.

Protein samples were concentrated using centrifugal filter units (Millipore). All samples for NMR measurements were prepared in 10 mM potassium phosphate buffer at pH 6.8, also containing 0.02% NaN_3 and 8% D_2O.

2.3.1) Chromatography resins and columns

Ion exchange chromatography resin*:* it has been used the weak anion exchanger DE52 preswollen microgranular diethylaminoethyl (DEAE) cellulose (Whatman). 80g of resin were resuspended in 200ml of 0.2M Tris-acetate pH 7.8. After sedimentation the supernatant was removed and the resin washed twice with 200ml of 50mM ammonia acetate pH 4.5. The resin was then poured into a XK 26/40 *(GE Healthcare)* column and extensively washed with 50mM ammonia acetate pH 4.5, 0.02% NaN_3 as running buffer. The cleared cellular lysate was loaded directly into the column and the protein was eluted with several column volumes of running buffer. The chromatography was performed at a constant flow rate of 1ml/min using ÄKTAprime *(GE Healthcare)* and the proteins elution profile was followed monitoring the absorbance at 280nm (Abs_{280}). Fractions of 2ml were collected and analyzed by SDS-PAGE. After use the resin was regenerated according to the manufacturer's instructions.

Size exclusion chromatography resin*:* 500ml of Sephacryl S-100 High Resolution resin, designed for separating peptides and small proteins, were poured into a XK26/100 (*GE Healthcare)* column and extensively washed with 50mM Tris-HCl pH 7.2, 0.2M NaCl, 0.02% NaN_3 (running buffer) by using ÄKTAprime *(GE healthcare).* Fractions containing the protein of interest were concentrated to 5ml with an Amicon® or Centricon concentrator. The protein sample was then injected in the column and the chromatography was performed at a constant flow rate of 1 ml/min using ÄKTAprime

(GE Healthcare). The proteins elution profile was followed monitoring the absorbance at 280nm (Abs280). Fractions of 10ml were collected and analyzed by SDS-PAGE. After use the resin was regenerated according to the manufacturer's instructions.

Nickel affinity chromatography: Chelating Sepharose, when charged with Ni^{2+} ions, selectively binds to proteins if complex forming amino acid residues, in particular hystidines, are exposed on the protein surface. The matrix Chelating Sepharose Fast Flow *(Amersham Biosciences),* which is able to bind metal cations around neutral pH values, was packed into a XK16/20 column *(GE Healthcare).* First the resin was charged with Ni^{2+} using 1-2 volumes of a solution of 0.25M $NiSO_4$, then the equilibration was achieved with 5 volumes of 20mM Tris-HCl pH 7.5, 0.5 M NaCl, 10 mM imidazole, 0.02% NaN_3 (binding buffer). The protein elution was performed applying a step imidazole gradient from 10mM to 0.5M, exploiting the ability of this compound to compete with the (His) 6 tag for Ni^{2+} binding. Fractions of 2ml were collected and analyzed by SDS-PAGE.

Glutathione Sepharose chromatography resins: 20mL of Glutathione Sepharose resin (*GE Health care*) were packed in a XK 16/20 (*GE Healthcare*) column. The resin is first washed with 10-20 volumes of bidistilled water and successively with the proper buffer. The protein elution was performed adding 10mM Glutathion in the proper buffer, at pH 8.

His-tag removal: After affinity chromatography the protein was buffer exchanged by dialysis, with the buffer 20mM Tris, 150mM NaCl, 0.02% NaN_3 at pH 7.5. (His)6 tag was removed at 20° o/n, by adding thrombin at a final concentration of 4U/mg and $CaCl_2$ at a final concentration of 2.5mM. Successively, an aliquot of the sample, before and after the cleavage, was analyzed by SDS-PAGE to verify the removal of the tag.

2.3.2) Instruments

To concentrate the proteins the Amicon® *concentrator* (50ml) with regenerated cellulose membranes (YM series) of 3000Da cutoff and disposable Centricon of 3000Da cutoff were used.

Liquid chromatograpy was carried out using AKTA Prime Automated Liquid Chromatography System (GE Healthcare).

For *protein gel electrophoresis* the power supply EPS 301(Pharmacia Biotech) and the Mini- PROTEAN® 3 cell (Biorad) was used while for separation DNA fragments, electrophoresis was performed on HE33 Mini-Sub® (Hoefer Inc. USA).

Media sterilization was achieved using an Alfa 10 plus (PBI) autoclave.

Bacterial growth were performed using a FOC 225E refrigerated incubator (VELP scientifica) and bacterial lysis were carried out with the sonicator Sonoplus HD 2070 (Bandelin Electronic).

Centrifugation steps were carried out using centrifuges J2-HS (Beckman), Centrifuge 5804R and MiniSpin® plus (Eppendorf) and a Biofuge Fresco (Heraesus instruments).

2.3.3) SDS-PAGE

Protein purity was evaluated using Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS-PAGE). Every protein sample (10ul) was heated at 95°C for 5 minutes after the addition of 5 ul of reducing loading buffer. Samples were loaded into a separate well of a discontinuous 17% polyacrylamide gel together with 5 μL of molecular weight marker (GE Healthcare) loaded in a separate well internal referencing and to identify the proteins of interest after separation. Gels were run at 120 V for approximately one hour until the bromophenol blue exits the running gel. After running the gels were immersed in the staining solution for 20 minutes and destained overnight with bidistilled hot water.

Reducing loading buffer

Compound	**Quantity**
Tris-HCl pH 6.8	50mM
DTT	100mM
SDS	2%
glycerol	10%
bromophenol blue	0.1%

Polyacrylamide gel

Stacking gel solution

For a final 2ml:

Compound	**Quantity**
Acrylamide mix (30% stock solution)	330µl
Tris-HCl pH 6.8 (1M stock solution)	250µl
SDS (10% stock solution)	20µl
Ammonium persolfate (10% stock solution)	20µl
TEMED	2µl

Running gel solution

For a final 5ml at 17% of acrylammide:

Compound	**Quantity**
Acrylamide mix (30% stock solution)	2ml
Tris-HCl pH 6.8 (1M stock solution)	1ml
SDS (10% stock solution)	40µl
Ammonium persolfate (10% stock solution)	40µl
TEMED	2µl

Running buffer

Compound	**Quantity**
Tris-HCl pH 8.3	25mM
glycine	192mM
SDS	0.1%

Staining solution

For a final 1L:

Compound	Quantity
Ethanol	500ml
Acetic acid	200ml
Comassie Brilliant Blue	3g

2.4) Synthesis of polyUb chains

2.4.1) Notation and design of segmentally isotope-enriched polyUb chains

To overcome the spectroscopic equivalence of all ubiquitin monomers comprising the polyUb chain, chains with only one Ub unit isotope labeled (15N) in a given sample were synthesized. The following notations are used throughout to refer to Ub2 isotope labeled at different Ub units. Ub2s are referred to as 'Ub2-P' and 'Ub2-D' for Ub2 samples isotope labeled at the proximal and distal positions (with respect to a possible substrate) respectively. Ub2-Pchains were synthesized using isotope labeled D77 Ub (proximal Ub) and unlabelled K48R or K63R Ub (distal Ub), and Ub2-D chains were synthesized using unlabeled D77 Ub and isotope labeled K48R or K63R Ub.

2.4.2) Synthesis of Ub2 chains

Lys48- and Lys63-linked Ub2 were synthesized following the protocol in above. (Petrowski et al., 1997)

K63-D77 Mix	Initial concentration	Final concentration
PBDM8	5X	1X
ATP	0,1M	2mM
TCEP	0,5M	3mM
K63	10mM	1mM
D77	10mM	1mM
Ubc13	0,3mM	100µM
Mms2	0,3mM	60µM
E1	94µM	400nM
NaN3	20%	0,05%

K48-D77 mix	Initial concentration	Final concentration
PBDM8	5X	1X
ATP	0,1M	2mM
TCEP	0,1M	3mM
K48	9mM	0,9mM
D77	7,7mM	0,9mM
E2	1,1mM	100µM
E1	94µM	500nM
NaN3	20%	0,05%

Typically, 2ml reactions with a total of 30mg Ub were set up and incubated overnight at 37°C. The formation of Ub2 was confirmed by SDS polyacrylamide gel electrophoresis and the Ub2 was separated from unreacted monoUb and other components of the reaction mixture using cation exchange chromatography as follows. The pH of the mixture was dropped to ~4 by the addition of 2-3 drops of undiluted acetic acid. The reaction mixture was then applied to a SP Sepharose Fast Flow column (1ml) pre-equilibrated in 50mM ammonium acetate, pH 4.5. Components of the reaction mixture such as E1 and E2, did not bind the column, and the Ub2 was separated from monoUb using a salt gradient over 40 column volumes to a final salt concentration of 0.4M. 1-1.2 ml fractions were collected and checked on a 15% SDS gel. The fractions corresponding to Ub2 were pooled, concentrated in Amicon filtration units and exchanged into the desired buffer

2.4.3) Estimation of protein concentration

The absorbance at a wavelength of 280nm was detected using Nanodrop spectrometer. Protein concentration (c) was calculated from the Beer-Lambert equation:

$$Abs = \varepsilon \times l \times c$$

A theoretical extinction coefficient at a wavelength of 280nm measured in water (ε280nm) was estimated for each construct using the ProtParam Tool from the Expasy Tools server (http://www.expasy.ch/tools/protparam.html).

2.4.4) Buffer exchange

Prior to storage, purified samples were buffer exchanged into a buffer suitable for NMR Experiments, in our case we use KPi10mM pH 6.8.

After chromatography purified protein samples were concentrated to 0.5mM in the proper Buffer. Then they were frozen with liquid nitrogen and stored at -20°C. A protein concentration in the mM range is usually required for NMR purpose. However in this case a concentration 0.5mM (or lower) has been kept to avoid aggregation and precipitation.

2.4.5) Sample quality control

Unlabelled purified proteins were checked by ^{1}H mono-dimensional NMR experiment to asses the protein folding and presence of contaminants. For ^{15}N and/or ^{13}C isotope labelled proteins basic bi-dimensional heteronuclear NMR experiments were performed to confirm the isotope inclusion and to evaluate the goodness of the protein preparation.

2.5) Materials for liposome preparation and monitoring

2.5.1) Lipids and materials

1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) sodium salt (POPG), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dioleoyl-3 trimethylammonium-propane chloride salt (DOTAP), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), and cholesterol (Chl) were purchased from Avanti Polar. Ludox® TMA colloidal silica nanoparticles were purchased from Sigma-Aldrich. All other reagents necessary for this study were from Sigma-Aldrich.

2.5.2) Instruments

Large unilamellar vesicles (LUVs) were formed by pressure extrusion using *Avanti Polar Device "Mini-Extruder"* (Alabaster, AL). Polycarbonate membranes with a pore size of 400nm and 100nm were from Whatman (GE Healthcare).

Small unilamellar vesicles (SUVs) were prepared using *probe tip sonicator* (Misonix, NY). Liposomes were characterized in terms of their sizes and charge using dynamic light scattering with a *Zetasizer Nano ZS instrument* (Malvern Instruments, USA) operating at $\lambda = 633nm$.

2.5.3) Liposome preparation protocol

Properties of lipid formulations can vary depending on the composition (cationic, anionic, neutral lipid species) and dimension, however, the same preparation method can be used for all lipid vesicles regardless of composition. The general elements of the procedure used, involve three steps: preparation of lipid for hydration, hydration of lipid film and sizing to a homogeneous distribution of vesicles.

2.5.4) Preparation of Lipid for Hydration

Powdered pure phospholipids or mixtures were dissolved in a chloroform solution to assure a clearly homogeneous solution and dried as a thin film in round-bottom glass flask by rotary evaporation for 2 hours. However, often a small amount of methanol is also required to solubilize polar lipids, such as phospholipids. During this process the temperature was maintained at around 30°C.

2.5.5) Hydration of lipid film

Hydration of the dry lipid film was accomplished simply by adding 30mM phosphate solution (Na_2HPO_4/NaH_2PO_4), 0.1M NaCl, pH 7 to a final concentration of ~120 mM lipid. The film was then dispersed by gently mixing. After addition of the hydrating medium the temperature of the hydrating solution should be above the gel-liquid crystal transition temperature (Tc or Tm) of the lipid with the highest Tm before adding to the dry lipid. For that reason the lipid suspension was maintained above the DMPG and DMPC lipid phase transition temperature (23°C). The product of hydration is a large multilamellar vesicle (LMVs) analogous in structure to an onion, with each lipid bilayer separated by a water layer. Once a stable, hydrated LMV suspension has been produced, the particles can be downsized by a variety of techniques, including sonication or extrusion. All the next steps were performed above the Tm of the lipid unless otherwise specified.

2.5.6) Small unilamellar vesicles (SUVs) preparation

SUV were prepared using probe tip sonicator (Misonix, NY) without cooling at 40 Watt for 10 minutes. The sonication continued until the solution became clear, in order to assure that the vesicles had the smallest possible size. Sonication tips also tend to release titanium particles into the lipid suspension. After sonication SUV solutions were centrifuged to sediment small metal pieces released by the sonication tip.

2.5.7) Large unilamellar vesicles (LUVs) preparation

LUV were formed by pressure extrusion using Avanti Polar device. Through the extrusion technique a lipid suspension is forced through a polycarbonate filter with a defined pore size to yield particles having a diameter near the pore size of the filter used. In order to obtain LUVs with a mean diameter of 100nm repeated extrusions of the multilamellar liposomes were performed through polycarbonate membranes with a pore size of 1μm, followed by extrusion through 400nm and 100nm pore size membranes.

2.5.8) Liposome storage

Storage time depends on a number of factors including temperature, pH, medium, etc. POPG SUVs and LUVs were stored at 30°C above the phase transition temperature for few days. Small unilamellar vesicles are much more unstable than large unilamellar vesicles. Beyond 24 hours SUVs size increases due to aggregation events and other sonication process was performed in order to assure the original vesicle diameter.

2.5.9) Checking liposome size and stability

Liposomes were characterized in terms of their sizes and over time stability using dynamic light scattering with a Zetasizer Nano ZS instrument (Malvern Instruments, USA) backscattering (173°C) at λ = 633nm at 30°C. Samples were loaded into disposable cells and diluted to 100uM with 30mM phosphate buffer pH 7.0. Standard Operating Procedures (SOP) were performed repeating data acquisition 5 times.

2.6) Protein analysis

2.6.1) Differential scanning calorimetry (DSC)

Thermal denaturation data were acquired with a Nano DSC instrument (TA instruments Inc.) at a scan rate of 1 K min^{-1}. Samples contained 0.1 mM protein dissolved in 100 mM sodium phosphate, pH 7.4, 100mM NaCl. Before measurements, sample and reference solutions were properly degassed in an evacuated chamber for 10 minutes at room temperature and carefully loaded into the cells to avoid bubble formation. Calorimetric cells (operating volume 300µL) were kept under a pressure of 3 atmospheres. Exhaustive cleaning of the cells was undertaken before each experiment. A background scan collected with buffer in both cells was subtracted from each scan.

2.6.2) Fluorescence spectroscopy

Fluorescence measurements were performed on a FP-8200 spectrophotometer (JASCO) in Na_2HPO_4/ NaH_2PO_4 buffer (10mm, pH 6.8). Dansyl-DHPE spectra were recorded (emission 450–600 nm, 5 nm bandwidth; excitation 336nm, 2.5nm bandwidth). All measurements were performed in triplicate, accumulated, and smoothed. Fluorescence shift was measured from the barycentric mean fluorescence (lbcf), defined as [Equation]:

$$\lambda_{bcf} = \frac{\sum_\lambda \lambda I(\lambda)}{\sum_\lambda I(\lambda)}$$

Fixed liposome concentration of 10µM were titrated with an increasing amount of ubiquitin and poly-ubiquitin up to a final [Lipo]/[Ub] molar ration of 1:30.

2.7) NMR spectroscopy experiments

NMR experiments were acquired at 25 °C on a Bruker Avance III spectrometer, operating at ^1H Larmor frequency of 600.13MHz, equipped with a triple resonance TCI cryogenic probe. NMR data were processed with Topspin 3.2 (Bruker) or NMRpipe and analyzed with the software Sparky (T. D.Go ddard and D. G. Kneller, University of California, San Francisco). All samples for NMR measurements were prepared in

10mM potassium phosphate aqueous buffer at pH 6.8, also containing protease inhibitors (Sigma) and 8% D_2O.

One-dimensional ^{1}H-NMR experiments were acquired with a standard pulse sequence incorporating the excitation sculpting water suppression scheme. A total of 128 transients were acquired over a spectral width of 12019 Hz, using a recycle delay of 2s. The spectra were processed applying an exponential window function prior to the Fourier transformation.

Typical sensitivity enhanced ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) were conducted by acquiring 256 complex points in the T_1 dimension, 1024 in the T_2 dimension. A total number of 4-8 transients were acquired for each spectrum with an interscan delay of 1.2s. Standard sequence schemes with pulsed field gradients were used to achieve suppression of the solvent signal and cancellation of spectral artifacts. Data were processed using the Bruker software TOPSPIN 3.2 (Bruker, Karlsruhe) and analyzed with the program CARA (The Computer Aided Resonance Assignment Tutorial) by R. Keller (http://www.nmr.ch).

Protein–ligand NMR titration experiments were conducted with ^{15}N-labeled Ubiquitin od polyUb with an increasing amounts of Uba2, liposomes or nanoparticle, using different ligand/protein molar ratios, and ^{1}H,^{15}N HSQC spectra were recorded at each ligand addition.

Chemical shift perturbations (CSP) were reported as weighted average chemical shift differences ΔδHN for each backbone amide to account for differences in spectral widths between ^{1}H and ^{15}N resonances, and calculated as $CSP = [(\Delta\delta H)^2 + (\Delta\delta N/5)^2]^{0.5}$ where ΔδH and ΔδN correspond to the change in ^{15}N and the ^{1}H chemical shifts, respectively.

Backbone amide transverse relaxation ($^{15}NR_2$) experiments were measured in gradient-selected sensitivity-enhanced and interleaved mode (64 transients, interscan delay of 3.0 s). Backbone amide longitudinal relaxation (^{15}N-R_1) rates were measured with the same parameters. Relaxation constant values T_1 and T_2 were determined by fitting peak volumes obtained by CARA to a two-parameter single exponential function with the software RELAXFIT.

Steady-state $\{^1H\}^{15}N$ heteronuclear nuclear Overhauser effects (hetNOE) were measured with a 6 s recycle delay. hetNOE values were calculated taking the ratio of peak intensities in saturated and reference spectra.

The dissociation constant value for the interaction was obtained by fitting experimental binding isotherms obtained from ^1H-^{15}N HSQC-based titration experiments, assuming a one-site binding model and by use of the Matlab program Kdfit.

Transverse $_1$HN PRE rates, (^{1}HN-R_2p) were obtained as the difference in ^{1}HN-R_2 measured on samples containing or not containing 2 mM gadodiamide. In principle, the difference removes contributions to ^{1}HN-R_2 relaxation common to both states, including exchange contributions. The deviation of ^{1}HN-R_2p from a linear concentration dependence was estimated by:

$$\Delta = R_{2p}(r_2) - \overline{R}_{2p}$$

where

$$\overline{R}_{2p} = \frac{|R_{2p}(r_3) - R_{2p}(r_1)}{r_3 - r_1}(r_2 - r_1) + R_{2p}(r_1)$$

and r1 = 0, r2 = 0.5, r3 = 1, corresponding to the different molar ratios.

CHAPTER 3:
Publications

3.1 Publication 1:

"Identification of primary and secondary UBA footprints on the surface of ubiquitin in cell-mimicking crowded solution."

FEBS Lett. 2017;591(7):979-990

Identification of primary and secondary UBA footprints on the surface of ubiquitin in cell-mimicking crowded solution

Francesca Munari[1†], Andrea Bortot[1†], Serena Zanzoni[1], Mariapina D'Onofrio[1], David Fushman[2] and Michael Assfalg[1]

†These authors contributed equally to this work

ABSTRACT

Despite significant advancements in our understanding of ubiquitin-mediated signaling, the influence of the intracellular environment on formation of transient ubiquitin-partner complexes remains poorly explored. In our work, we introduced macromolecular crowding as a first level of complexity towards the imitation of a cellular environment in the study of such interactions. Using NMR spectroscopy, we found that the stereospecific complex of ubiquitin and ubiquitin-associated domain (UBA) was minimally perturbed by the crowding agent Ficoll. However, in addition to the primary canonical recognition patch on ubiquitin, secondary patches were identified, indicating that in cell-mimicking crowded solution, UBA contacts ubiquitin at multiple sites.

INTRODUCTION

Ubiquitin (Ub) is a prototypical small protein modifier regulating a vast number of fundamental cellular events[1,2]. The addition of a single Ub molecule to a target protein (monoubiquitination) can alter protein activity and localization, while the conjugation of distinct types of Ub chains (polyubiquitination) is implicated in a variety of processes such as proteasomal degradation, DNA repair, and immune signaling [2–5]. Differentially ubiquitinated substrates are recognized by an ample variety of Ub-binding proteins, which propagate the Ub signal eliciting specific cell responses [6,7]. Ub receptors contain modular elements, referred to as Ub-binding domains (UBD), able to specifically associate with the protein modifier. The binding of individual UBD to monoUb is generally weak but high-avidity interactions can be established by polyUb chains [8–11]. Intense work has been carried out during the past two decades to elucidate the structural determinants of Ub-UBD recognition, however further scrutiny is required to obtain a full understanding of the specificity of Ub-mediated signaling.

In spite of the complexity of Ub recognition by effector proteins, the simplicity of monomeric Ub's architecture is startling. Composed of only 76 amino acids, highly conserved throughout eukaryotes, Ub is among the smallest proteins found in a cell. The polypeptide chain adopts a compact globular β-grasp fold, exposing a surface area of less than 5,000 $Å^2$. A solvent-exposed hydrophobic area, centered around residues Leu8, Ile44, and Val70 (Ile44 patch[4,12]), stands out from the predominantly polar protein surface. The Ub backbone is generally considered rigid, with the exception of the C-terminus and of the β1- β2 loop. However, a certain degree of structural plasticity allows a mechanism of dynamic adaptation to operate during molecular recognition events [13]. Interactions with many ubiquitin-binding proteins involve the Ile44 patch [4], albeit distinct residues form the interaction surface with different protein partners. Furthermore, additional protein recognition sites have been identified, including the C-terminus, the α/β2 groove, the Ile36, the Asp58, and the Phe4 patches [11,14].

Given that a finely tuned surface chemistry dictates the determinants of Ub-UBD recognition, a definitive description of the corresponding modes of interaction requires evaluation of all factors potentially affecting the binding phenomena in the native

environment. While our current understanding of Ub-UBD associations relies on interaction studies performed on binary protein mixtures in buffered aqueous solution, it has become clear that the complex cellular environment can exert significant influence on biomolecular properties and interactions [15–17]. A prominent feature of the cellular interior is intense macromolecular crowding caused by elevated concentrations of large biological molecules that occupy a significant fraction (10-40%) of the cell volume [18]. Macromolecular crowders, or cosolutes, alter a protein's effective concentration, increase solution viscosity, originate excluded-volume effects, and may engage in unspecific or specific chemical interactions with the protein solutes [19–24]. The impact of tight molecular packing on the formation of Ub-UBD complexes has remained unexplored so far.

In our work, we introduced macromolecular crowding as a first level of complexity towards the imitation of a cellular environment in the study of Ub-UBD interactions. Steric repulsions and depletion forces are predicted to perturb conformational equilibria and diffusive dynamics, to shift equilibrium states, and change the probability of intermolecular collisions [19]. However, the extent to which these phenomena reshape the protein-protein interaction landscape, influencing the relative orientations of components in non-covalent assemblies and the specificity of transient associations, remains elusive. Here, we investigated the interaction between Ub and the Ub-associated C-terminal domain (UBA2) of the human homologue of the yeast DNA repair protein RAD23 (HHR23A) in the presence of a non-interacting hydrophilic polymeric crowding agent. UBA2 is a compact three-helix bundle displaying a Ub recognition site, encompassing the non-adjacent helices α1 and α3, that associates weakly (K_d ~0.4 mM) with the Ile44 patch of Ub [25–27]. The Ub-UBA2 interaction can be considered paradigmatic in the study of analogous Ub-UBD pairs.

We used NMR spectroscopy to explore the UBA2 interaction sites on the surface of Ub. NMR allows the exploration of interactions over a broad range of affinities, yielding atomic-resolution insights into dynamic biochemical equilibria[22]. By complementing traditional chemical shift perturbation mapping with solvent paramagnetic relaxation enhancement analysis [28,29], a robust identification of binding surfaces was obtained. In addition to the well-characterized Ile44 patch, secondary

patches were found, indicating that in cell-mimicking crowded solution, UBA2 contacts Ub at multiple sites. Identification of such transient specific associations adds further detail to our description of protein-protein interactions in the complex cellular interior.

EXPERIMENTAL SECTION

A description of materials and protein preparation procedures can be found in the Supporting Information.

NMR experiments

All experiments were recorded at 25 °C on a Bruker Avance III spectrometer, operating at ^{1}H Larmor frequency of 600.13 MHz, equipped with a triple resonance TCI cryogenic probe. NMR data were processed with Topspin 3.2 (Bruker) or NMRpipe [30], and analyzed with the software Sparky (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

^{15}N relaxation experiments were performed on samples containing 0.37 mM [^{15}N]-Ub in the absence of crowder, in 200g/L Ficoll, or in 200g/L sucrose solution. Longitudinal relaxation rates, ^{15}N-R_1, were measured using relaxation delays in the interval 0.01-1.26s for Ub in uncrowded solution and in the range 0.01-1.44s for Ub in 200g/L Ficoll or sucrose. Transverse relaxation rates, ^{15}N-R_2, were measured with a CPMG-based pulse program, using relaxation delays in the range 8-208ms for Ub in simple buffer, 8-176ms for Ub in 200g/L sucrose, and 8-104ms for Ub in 200g/L Ficoll.

Proton transverse relaxation rates, $^{1}H_N$-R_2, were measured on a 0.5 mM [^{15}N]Ub sample using the pulse program described by Iwahara et al. [31], kindly provided to us by the author. To remove $^{3}J_{HN\text{-}H\alpha}$ modulation of peak intensities, a selective H_N 180° pulse (REBURP of 2ms) centered at 8.2 ppm was used in the INEPT period. For each titration point (1:0, 1:0.5 and 1:1), measured in the presence or absence of gadodiamide, seven relaxation delays were acquired and the signal intensity decays were fitted to a single exponential function to obtain the corresponding rates. Delays between 11.7 and 52.5 (45.3)ms were used for samples without (with) gadodiamide. Residues affected by signal overlap or with insufficient signal-to-noise ratio were excluded from the analysis.

Data analysis

Amide chemical shift perturbation was calculated as: $CSP = [(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{0.5}$, where $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical shift changes measured in the ^1H and ^{15}N frequency dimensions, respectively.

Volume occupancy by Ficoll was calculated from a reported partial specific volume of $650 cm^3/g$ [32], a value in agreement with our experimental verification of volume change upon dissolution of a known mass of solute. The volume fraction occupied by the crowder, ϕ_C, in a 200g/L Ficoll solution was determined to be 13%, assuming no volume changes resulted from perturbations of solvent molecules.

Dissociation constant values were obtained by fitting experimental binding isotherms according to a one-site binding model using the Matlab program Kdfit [26]. Titration data were analyzed assuming that the observed CSP is a weighted average between the value corresponding to the free (CSP = 0) and ligand-bound ($CSP = CSP_{bound}$) states, so that $CSP = CSP_{bound} \times f_{bound}$, where f_{bound} is the relative population of the bound state of the molecule under observation, related to the dissociation constant K_d according to the following expression [26]:

$$f_{bound} = \left(C_P + C_L + K_d - \sqrt{(C_P + C_L + K_d)^2 - 4C_P C_L} \right) / (2C_P)$$

where C_P and C_L are the total concentrations of Ub and UBA2, respectively. The analytical concentration of the initial Ub solution was 500μM and that of the titrant stock solution was 10mM. Dilution-corrected values for C_P and C_L were used at successive titration steps. Reported values are the average of K_d values determined from seven binding isotherms and the corresponding standard deviation.

Relaxation rate values were obtained from fitting of signal intensity decays with a single exponential function. The rotational correlation time, τ_R, was estimated with the program ROTDIF [33] from the experimental ^{15}N-R_1 and ^{15}N-R_2 relaxation rate values, assuming a constant NOE value of 0.74 for all residues displaying secondary structure and using an isotropic model. Relative solution viscosities were estimated using the Stokes-Einstein-Debye equation: $\tau_R = 4\pi\eta r^3/(3kT)$, where η is the solution viscosity, r the hydrodynamic radius of Ub (assumed constant in all experimental conditions), k the

Boltzmann constant and T the temperature. Using the measured τ_R values, we determined the following ratios: $\eta^{Ficoll}/\eta^{sucrose}$ = 1.19, $\eta^{Ficoll}/\eta^{buffer}$ = 2.3 for 200g/L crowder solutions.

Transverse ^1H_N PRE rates, (^1H_N-R_{2p}) were obtained as the difference in ^1H_N-R_2 measured on samples containing or not containing 2mM gadodiamide. In principle, the difference removes contributions to ^1H_N-R_2 relaxation common to both states, including exchange contributions[31]. The deviation of ^1H_N-R_{2p} from a linear concentration dependence was estimated by:

$$\Delta = R_{2p}(r_2) - \overline{R}_{2p}$$

where

$$\overline{R}_{2p} = \frac{R_{2p}(r_3) - R_{2p}(r_1)}{r_3 - r_1}(r_2 - r_1) + R_{2p}(r_1)$$

and r_1 = 0, r_2 = 0.5, r_3 = 1, corresponding to the UBA2/Ub molar ratios.
The uncertainties of Δ, σ_Δ, were obtained by propagation of errors on relaxation rates [34].
Protein structures were visualized with PyMOL (The PyMOL Molecular Graphics System, Version 1.1r1, LLC).

RESULTS

The structures of Ub and of the Ub/UBA2 complex are retained in crowded solution

Macromolecular crowding of the interior of cells can be conveniently modelled by use of a variety of crowding agents, including biomacromolecules and synthetic polymers [35,36]. Depending on the chemical nature of cosolutes, macromolecular crowding effects may include both steric repulsions as well as nonspecific chemical interactions. Here, we selected the macromolecular crowder Ficoll (70 kDa), a nearly spherical and densely branched hydrophilic neutral polymer that is reported not to interact with most proteins [17,35] and in particular with Ub [37]. Experiments were conducted after dissolving the protein(s) in 200 g/L Ficoll solution, approximately

corresponding to the concentration of macromolecules in the average eukaryotic cytoplasm.

Our preliminary investigation concerned the possible effects of Ficoll on the structural and dynamic properties of Ub. We therefore recorded two-dimensional proton-nitrogen correlation (^{1}H,^{15}N-HSQC) NMR spectra of ^{15}N-enriched Ub, these experiments being exquisitely sensitive to binding events and structural changes. The spectral fingerprint of Ub in Ficoll solution, reporting separate signals for individual amino acid residues of the polypeptide, displayed all peaks in the same positions as those observed in uncrowded solution: amide chemical shift perturbation (CSP) values were close to zero for all residues (Fig. 1A), in agreement with the findings of Cino *et al.* [38] and Abriata *et al.* [37]. This observation indicates that Ficoll did not perturb the overall structure of Ub. ^{15}N-spin relaxation rate measurements were then carried out to explore possible changes in Ub main chain dynamics. Data reported in Fig. 1B,C show a significant decrease in longitudinal relaxation rate (^{15}N-R_1) values and increase in transverse relaxation rates (^{15}N-R_2), due to a slower molecular rotational diffusion resulting from increased viscosity (we calculated a global rotational correlation time, τ_R, of 4.0ns and of 9.2ns for Ub in uncrowded solution and in 200g/L Ficoll, respectively). It can be noted that ^{15}N-R_1 values exhibit lesser variation along the polypeptide sequence in the presence of crowder, and in particular, residues in the β1- β2 loop and in the C-terminus display ^{15}N-R_1 values similar to or even larger than those observed in the rigid elements. Although it is possible that the observed trend reflects some reduced local dynamics in crowded solution, an increase of the longitudinal relaxation rate is predicted to take place in flexible regions upon increase in τ_R if the correlation time for local motion is in the range of hundreds of picoseconds or slower [39].

In order to obtain mechanistic information on UBA2 binding to Ub under macromolecular crowding conditions, titration experiments were conducted by acquiring a series of ^{1}H,^{15}N-HSQC spectra in the presence and absence of 200g/L Ficoll. Unlabelled UBA2 was added stepwise to [^{15}N]Ub, and individual peaks were monitored throughout the titration (Fig. 2). The directions of peak movements along the titration in crowding conditions were unchanged compared to those in dilute solution (Fig. 2A,B), indicating that Ficoll did not perturb the mode of binding of Ub to UBA2.

Despite the increase in solution viscosity when the polymer was present, we noted that binding still occurred in the fast exchange regime on the chemical shift time scale, indicating that the dissociation rate was not slowed down considerably in crowded solution. The CSP patterns along the protein sequence in the presence of Ficoll closely resembled those measured in the absence of crowder (Fig. 2C), suggesting that the average structure of the complex was not affected by the presence of crowder molecules. The backbone regions most affected by UBA2-binding in both solution conditions contain residues surrounding Leu8, Ile44 and Val70, the key residues forming the well-characterized Ile44 hydrophobic patch [12,14,26].

Titration experiments in 200g/L Ficoll solution pointed out that binding site saturation was reached at a lower nominal concentration of UBA2 than in uncrowded solution. Indeed, by fitting the binding isotherms of selected residues (Fig. 2D,E) we determined K_d = 412 ± 52μM for the Ub/UBA2 interaction in buffer, while in the presence of Ficoll the corresponding value was 226 ± 48μM. Accounting for the actual protein concentrations (corrected for the volume occupied by Ficoll particles, see Materials and Methods for details) would result in at most ~10-15% increase of the K_d. Therefore, our data indicate that the apparent affinity of the complex increased in the presence of crowder. After repeating the titration experiment in 200g/L sucrose (the monomeric counterpart of Ficoll) (Fig. S1), the K_d was found to be 381 ± 52μM, indicating that the macromolecular nature of Ficoll was responsible for the increased affinity.

Solvent paramagnetic relaxation enhancement analysis reveals secondary UBA2 footprints on the surface of Ub

After assessment of formation of the stereospecific Ub/UBA2 complex in crowded solution, we set to perform a more comprehensive evaluation of the protein-protein contacts, possibly including less represented binding sites. To this aim we measured solvent paramagnetic relaxation enhancements (PRE), a sensitive NMR approach based on the use of soluble paramagnetic relaxation agents to reveal changes in macromolecular surface accessibility occurring during dynamic events such as protein

complex formation. Due to their inherently large sensitivity (caused by the large magnetic moment of unpaired electrons), paramagnetic probes have been used successfully to characterize the structure of protein-protein complexes [40,41], formation of encounter complexes [42], and transient interactions [43]. In particular, Johansson et al. [34], were able to identify residues of human growth hormone that are involved in either unspecific or specific protein self-interactions based on the PRE profile of amide protons induced by gadodiamide (a gadolinium-based paramagnetic relaxation agent, also referred to as Gd-DTPA-BMA). Gadodiamide-induced PRE effects on Ub were previously interpreted according to a relaxation model where the relaxation agent forms an unspecific, yet rotationally correlated, complex with the protein [44].

In our work, we measured the amide proton transverse relaxation rates, ^1H_N-R_2, on Ub in isolation, and in the presence of UBA2 at 1:0.5 and 1:1 molar ratios. To determine solvent PRE for ^1H_N of Ub, we repeated the same set of measurements in the presence of gadodiamide. A relaxation agent concentration of 2mM was chosen to allow measurement of sufficiently large ^1H_N-R_2 without excessively compromising the quality of signals for samples containing 200g/L of Ficoll (where resonances were already broadened compared to dilute solution conditions). Transverse ^1H_N PRE rates, ^1H_N-R_2p, were calculated as the difference in ^1H_N-R_2 between samples with and without gadodiamide.

The values of ^1H_N-R_2p determined for Ub display a significant variability along the protein sequence (Fig. 3, main panel), in qualitative agreement with the solvent accessibility along the Ub backbone (Fig. 3, top panel). In particular, residues 8-14, located in or adjacent to the β1- β2 loop, residues 45-49 forming the β3- β4 loop, and the C-terminal tail showed the largest ^1H_N-R_{2p} values, while residues 20-30 displayed the smallest PRE effects. We then examined the PRE trends upon subsequent additions of UBA2. In Ficoll solution, at Ub:UBA2 1:0.5 molar ratio, ^1H_N-R_2p values increased for all of the residues (Fig. 4A,B, red plots). As explained by Johansson *et al*. for their system [34], the observed increase in PRE with increasing total protein concentration (that of UBA2 in the present case) can be attributed to a reduced rate of diffusion of the relaxation agent along the protein surface, which in turn is caused by increased molecular crowding associated with more frequent non-specific transient collisions.

However, at a higher Ub:UBA2 molar ratio (1:1), the $^{1}H_{N}$-R_{2}p values were found to consistently increase for a subset of residues (Fig. 4A, red plots), while another set of residues displayed no increase or an increase smaller than that observed between 1:0 and 1:0.5 Ub:UBA2 molar ratios (Fig. 4B, red plots). The non-linear trend of $^{1}H_{N}$-R_{2p} values (which we refer to as a 'roof' pattern) can be explained by the progressive formation of long-lived specific protein-protein associations that reduce access of gadodiamide to residues at or near the binding interface. To quantify this effect, we calculated the deviation of $^{1}H_{N}$-R_{2}p values from a linear concentration dependence (Δ), as described previously [34]. We excluded from this analysis all of the residues that in free Ub displayed a $^{1}H_{N}$-R_{2}p value smaller than $8s^{-1}$ (Fig. S3), indicative of a buried position. Δ values were considered significant if the parameter exceeded twice its uncertainty (i.e. $\Delta/r\Delta > 2$, [34]).

As a control, we repeated the experiment using protein GB1 (not a partner of Ub) in place of UBA2 (Fig. 4A,B, green plots). In this case, we observed modest variations of Ub's $^{1}H_{N}$-R_{2p} values, attributable to less frequent protein-protein collisions at the surface of Ub, when in the presence of a non-interacting protein.

Residues displaying a significant Δvalue were: 2, 4, 8, 10, 11, 14, 20, 32-36, 40, 41, 43-46, 51, 54, 55, 58-60, 63, 67-69. These residues were mapped onto the Ub structure with distinct colors representing different surface patches (Fig. 5A,B). Notably, residues displaying the largest Δvalues ($> 5s^{-1}$) were residues 8, 14, 35, 36, 43, 44, 46, 59, 63, 67, 69, which include those forming the Ile44 patch. Therefore, this analysis revealed additional specific contact surfaces containing residues Phe4, Ile36, and Asp58, which were identified in alternative binding interfaces used by Ub to recognize different binding partners [11,45,46]. Some of the identified contact residues in the secondary patches (e.g. 32-36) also show a small CSP (Fig. 2C, 5C). Others, occupying positions 58-60 in a region between strands β4 and β5, do not exhibit significant CSP (Fig. 2C, 5C), probably due to the small population of the corresponding complex.

It must be noted that the approach described above proved unsuccessful in identifying specific binding patches in the uncrowded solution. Indeed, $^{1}H_{N}$-R_{2}p values obtained upon addition of UBA2 in the same concentration range as explored with Ficoll, did not display the typical 'roof' pattern even in the case of residues belonging to

the Ile44 patch (Fig. S4). Given the small size of the chosen test proteins, it is possible that Ub/UBA2 collisions and reorientations in dilute conditions are too fast to be detectable by our PRE method.

DISCUSSION

Ubiquitin function is determined by covalent modification of protein substrates as well as by non-covalent recognition of partner protein surfaces. Owing to its small size, a large part of Ub's amino acid residues are exposed to the exterior and are potentially available for biomolecular recognition. Among these, a prominent role is played by residues forming the Ile44 patch, which centralizes most of the interactions with Ub binding domains [4,11]. Remarkably, this region is able to attract UBDs displaying highly divergent structural traits, such as single helices, multiple helices, and even β sheets [14]. The interactive capacity and binding versatility of the Ile44 patch is ensured by a finely tuned combination of rigidity and plasticity [13,47]. In contrast to the high level of detail obtained in the description of protein binding at the Ile44 region, the role of the remainder of Ub's surface in protein-protein association events remains less well characterized.

Monomeric Ub and individual Ub moieties in polyUb chains offer limited contact size for binding, and most Ub-UBD interactions are characterized by low affinity *in vitro*. Weak transient protein-protein interactions can be influenced by the heterogeneous and crowded intracellular environment to a larger extent than high-affinity associations [48], however investigations of Ub-UBD binding in cell-mimicking media are lacking. In our work, we aimed at assessing the effect of macromolecular crowding on the interaction between Ub and a representative UBD, UBA2, that was previously shown to target the canonical Ile44 patch on Ub [26,41]. Because of the multiple possible perturbations originating from macromolecular crowders, we focused our study on a hydrophilic sucrose polymer, Ficoll 70, eliciting viscosity and excluded-volume effects [35], but not establishing chemical interactions with the test proteins.

Ub experienced minimal perturbations in the presence of up to 200g/L Ficoll (a concentration close to the total macromolecular concentration in the average eukaryotic

cytoplasm), except for viscosity-dependent hydrodynamic properties. Similarly, the overlapping features of site-resolved NMR spectral fingerprints collected with saturating amounts of UBA2 in dilute and in crowded solutions indicated that the structure of the Ub/UBA2 complex was retained. Thus, the Ile44 region mediates specific Ub/UBA2 recognition in buffer as well as in crowded solution. The most prominent crowding-induced perturbation concerns the apparent binding affinity, which was found increased with respect to uncrowded solution. By comparing the results obtained in polymeric (Ficoll) and non-polymeric (sucrose) crowder solutions, we found that the macromolecular nature of the cosolute was responsible for the increased complex stability. Protein-protein associations are predicted to be favored under macromolecular crowding conditions according to excluded volume and depletion force theories [49–53]. Depletion forces stabilize equilibrium-state and transition-state complexes, enhancing binding affinity and association rates [48,51]. Macromolecular crowding may also decrease dissociation rates by increasing rebinding probability from the encounter complex [48]. On the other hand, diffusion and collision between associating molecules are slower under macromolecular crowding (the viscosity of a Ficoll solution is higher than that of a sucrose solution at equal mass concentration, see Materials and Methods, and Fig. 1, S2) resulting in decreased association rate if the reaction is diffusion-limited, or unperturbed on-rate if it is under transition-state control [48]. Although the precise influence of viscosity on formation of the Ub/UBA2 complex remains unknown, it appears that depletion attraction is a major contributor to the observed increase of binding affinity in Ficoll solution.

In addition to the identification of the primary recognition interface, we explored the possibility to detect UBA2 binding to alternative surface patches. Indeed, transient complexes play an important role in macromolecular associations, and repositioning of binding partners after preliminary collisions efficiently leads to formation of longer-lived assemblies [54]. Furthermore, it is expected that alternative encounters are particularly evident in low-affinity protein-protein complexes [55], such as Ub-UBD, due to the high concentration of UBD necessary to obtain saturation of the primary binding site. Paramagnetic relaxation enhancement approaches have established capability to detect ultra-weak interactions [42,43,56]. Here, we adopted the solvent PRE method in place

of covalent conjugation by paramagnetic tags in order to avoid potential artifacts due to chemical modification of a small-sized protein. From the systematic analysis of ^{1}H PREs of Ub at increasing concentrations of UBA2 in Ficoll solution, a differential behaviour of surface residues was observed. Specific interactions, revealed by a non-linear increase in PRE attributed to the exclusion of the paramagnetic molecular probe from the protein-protein interface, were detected for residues of the primary binding site at the Ile44 patch and for residues previously identified in non-canonical binding surfaces on Ub [11,14,45,46]. For example, a 'polar' surface centered on Asp58, involving Ub residues 51, 54, 55, and 57-60, was recognized as a novel interaction interface in Ub-Rabex5 complex in addition to the canonical Ile44 hydrophobic area [57,58]. Also, in Ubch5b-Ub-NEDD4L complex, Ub contacts the E3 partner via the Ile36 patch (Ile36-Leu71-Leu73) and other residues including Gly35, Gln40 and Leu69 [59]. Binding of linear ubiquitin chains by NEMO involves distinct patches on the Ub surface: while the distal Ub binds via its C-terminal tail and the Ile44 area, the proximal Ub employs residues Gln2, Phe4, Lys6, Gly10, Thr12, Ile13, Thr14, Glu16, Glu64, and Thr66 [60].

Thus, our data suggest that, in equilibrium conditions, low-populated specific UBA2/Ub associations coexist with the main stereospecific complex. It remains to be established whether these interactions are unintended (noise) or productive (part of the signal, i.e. functional encounter complexes), and whether they are specific of the investigated UBA2 or paradigmatic for all UBDs. By investigation of further (larger) protein pairs it should become possible to determine if macromolecular crowding affects the relative free energies of primary and secondary interactions, a notion that could be exploited to select among alternative bound conformations.

CONCLUSIONS

Weak protein-protein interactions, such as those between Ub and UBD, are potentially prone to be influenced by macromolecular crowding, a distinctive feature of the intracellular environment. In our work we found that high concentrations of a crowding agent did not influence the preferential binding of UBA2 to the canonical Ile44 patch of Ub. On the other hand, from a more comprehensive exploration of

Ub/UBA2 contacts based on a solvent PRE approach, secondary contact surfaces were detected. The regions were classified as specific based on the non-linear PRE trend observed at varying UBA2 concentration, although the absence of strong concomitant CSP hinted at ultra-weak affinity of the corresponding interactions and/or a significant heterogeneity of the ensemble of complex conformations. Thus, Ub/UBA2 complexes populate high energy local minima of the free energy landscape which may be in equilibrium with the low energy minimum corresponding to the stereospecific complex. Alternatively, the identified patches may mediate formation of futile complexes that are in competition with the main conformation. It is probably not accidental that other UBD target these regions, which can thus be considered pivotal for biomolecular recognition. In the broader context, identification of weak secondary interaction surfaces in cell-mimicking crowded solutions by use of PRE methods could improve our understanding of dynamic protein-protein interaction networks and, ultimately, of the molecular-level structural organization of the intracellular milieu.

ACKNOWLEDGMENTS

This work was supported by the Italian Ministry for Education and Research via departmental funding. We thank "Centro Piattaforme Tecnologiche" of the University of Verona for access to NMR instrumentation. SZ and FM received a fellowship grant (Assegno di Ricerca) from the Department of Biotechnology. DF was supported by NIH grant GM065334.

AUTHOR CONTRIBUTION

MA conceived and supervised the study; AB prepared protein samples; AB and FM performed experiments and analyzed data; MD and SZ analyzed data; DF critically revised the manuscript; MA and FM wrote the manuscript.

REFERENCES

1 Pickart CM & Eddins MJ (2004). *Biochim. Biophys. Acta* **1695**, 55–72.

2 Komander D & Rape M (2012). *Annu. Rev. Biochem.* **81**, 203–229.

3 Hicke L (2001) . *Nat. Rev. Mol. Cell Biol.* **2**, 195–201.

4 Pickart CM & Fushman D (2004). *Curr. Opin. Chem. Biol.* **8**, 610–616.

5 Alfano C, Faggiano S & Pastore A (2016). *Trends Biochem. Sci.* **41**, 371–385.

6 Hicke L, Schubert HL & Hill CP (2005). *Nat. Rev. Mol. Cell Biol.* **6**, 610–621.

7 Dikic I, Wakatsuki S & Walters KJ (2009). *Nat. Rev. Mol. Cell Biol.* **10**, 659–671.

8 Varadan R, Assfalg M, Raasi S, Pickart C & Fushman D (2005). *Mol. Cell* **18**, 687–698.

9 Sims JJ, Haririnia A, Dickinson BC, Fushman D & Cohen RE (2009). *Nat. Struct. Mol. Biol.* **16**, 883–889.

10 Fushman D & Wilkinson K (2011). *F1000 Biol. Rep.* **3**.

11 Husnjak K & Dikic I (2012). *Annu. Rev. Biochem.* **81**, 291–322.

12 Beal R, Deveraux Q, Xia G, Rechsteiner M & Pickart C (1996). *Proc. Natl. Acad. Sci. U. S. A.* **93**, 861–866.

13 Lange OF, Lakomek N-A, Fares C, Schroder GF, Walter KFA, Becker S, Meiler J, Grubmuller H, Griesinger C & de Groot BL (2008). *Science* **320**, 1471–1475.

14 Winget JM & Mayor T (2010). *Mol. Cell* **38**, 627–635.

15 Ellis RJ & Minton AP (2003). *Nature* **425**, 27–28.

16 Gierasch LM & Gershenson A (2009). *Nat. Chem. Biol.* **5**, 774–777.

17 Zhou H-X (2013). *FEBS Lett.* **587**, 1053–1061.

18 Theillet F-X, Binolfi A, Frembgen-Kesner T, Hingorani K, Sarkar M, Kyne C, Li C, Crowley PB, Gierasch L, Pielak GJ, Elcock AH, Gershenson A & Selenko P (2014). *Chem. Rev.* **114**, 6661–6714.

19 Minton AP (2013). *Biopolymers* **99**, 239–244.

20 Miklos AC, Li C, Sharaf NG & Pielak GJ (2010). *Biochemistry* **49**, 6984–6991.

21 Kuznetsova I, Turoverov K & Uversky V (2014). *Int. J. Mol. Sci.* **15**, 23090–23140.

22 Crowley PB, Brett K & Muldoon J (2008). *ChemBioChem* **9**, 685–688.

23 Ceccon A, Busato M, Pérez Santero S, D'Onofrio M, Musiani F, Giorgetti A & Assfalg M (2015). *ChemBioChem* **16**, 2633–2645.

24 Pérez Santero S, Favretto F, Zanzoni S, Chignola R, Assfalg M & D'Onofrio M (2016). *Arch. Biochem. Biophys.* **606**, 99–110.

25 Ryu K-S, Lee K-J, Bae S-H, Kim B-K, Kim K-A & Choi B-S (2003). *J. Biol. Chem.* **278**, 36621–36627.

26 Varadan R, Assfalg M, Haririnia A, Raasi S, Pickart C & Fushman D (2004). *J. Biol. Chem.* **279**, 7055–7063.

27 Mueller TD, Kamionka M & Feigon J (2004). *J. Biol. Chem.* **279**, 11926–11936.

28 Clore GM & Iwahara J (2009). *Chem. Rev.* **109**, 4108–4139.

29 Göbl C, Madl T, Simon B & Sattler M (2014). *Prog. Nucl. Magn. Reson. Spectrosc.* **80**, 26–63.

30 Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995). *J. Biomol. NMR* **6**, 277–293.

31 Iwahara J, Tang C & Marius Clore G (2007). *J. Magn. Reson.* **184**, 185–195.

32 Groen J, Foschepoth D, te Brinke E, Boersma AJ, Imamura H, Rivas G, Heus HA & Huck WTS (2015). *J. Am. Chem. Soc.* **137**, 13041–13048.

33 Walker O, Varadan R & Fushman D (2004). *J. Magn. Reson.* **168**, 336–345.

34 Johansson H, Jensen MR, Gesmar H, Meier S, Vinther JM, Keeler C, Hodsdon ME & Led JJ (2014). *J. Am. Chem. Soc.* **136**, 10277–10286.

35 Zhou H-X, Rivas G & Minton AP (2008). *Annu. Rev. Biophys.* **37**, 375–397.

36 Benton LA, Smith AE, Young GB & Pielak GJ (2012). *Biochemistry* **51**, 9773–9775.

37 Abriata LA, Spiga E & Peraro MD (2016). *Biophys. J.* **111**, 743–755.

38 Cino EA, Karttunen M & Choy W-Y (2012). *PLoS ONE* **7**, e49876.

39 Fushman D, Cahill S & Cowburn D (1997). *J. Mol. Biol.* **266**, 173–194.

40 Camacho-Zarco AR, Munari F, Wegstroth M, Liu W-M, Ubbink M, Becker S & Zweckstetter M (2015).*Angew. Chem. Int. Ed Engl.* **54**, 336–339.

41 Varadan R, Assfalg M & Fushman D (2005). *Methods Enzymol.* **399**, 177–192.

42 Volkov AN, Ubbink M & van Nuland NAJ (2010). *J. Biomol. NMR* **48**, 225–236.

43 Clore GM, Tang C & Iwahara J (2007). *Curr. Opin. Struct. Biol.* **17**, 603–616.

44 Pintacuda G & Otting G (2002). *J. Am. Chem. Soc.* **124**, 372–373.

45 Searle MS, Garner TP, Strachan J, Long J, Adlington J, Cavey JR, Shaw B & Layfield R (2012). *Biochem. Soc. Trans.* **40**, 404–408.

46 Kulathu Y & Komander D (2012). *Nat. Rev. Mol. Cell Biol.* **13**, 508–523.

47 Fenwick RB, Esteban-Martín S, Richter B, Lee D, Walter KFA, Milovanovic D, Becker S, Lakomek NA, Griesinger C & Salvatella X (2011). *J. Am. Chem. Soc.* **133**, 10336–10339.

48 Phillip Y, Sherman E, Haran G & Schreiber G (2009). *Biophys. J.* **97**, 875–885.

49 Asakura S & Oosawa F (1958). *J. Polym. Sci.* **33**, 183–192.

50 Minton AP (1998). *Methods Enzymol.* **295**, 127–149.

51 Minton AP (1983). *Mol. Cell. Biochem.* **55**, 119–140.

52 Phillip Y & Schreiber G (2013). *FEBS Lett.* **587**, 1046–1052.

53 Marenduzzo D, Finan K & Cook PR (2006). *J. Cell Biol.* **175**, 681–686.

54 Tang C, Iwahara J & Clore GM (2006). *Nature* **444**, 383–386.

55 Blundell TL & Fernández-Recio J (2006). *Nature* **444**, 279–280.

56 Liu Z, Zhang W-P, Xing Q, Liu M & Tang C (2012). *Angew. Chem. Int. Ed.* **51**, 469–472.

57 Penengo L, Mapelli M, Murachelli AG, Confalonieri S, Magri L, Musacchio A, Di Fiore PP, Polo S & Schneider TR (2006). *Cell* **124**, 1183–1195.

58 Lee S, Tsai YC, Mattera R, Smith WJ, Kostelansky MS, Weissman AM, Bonifacino JS & Hurley JH (2006). *Nat. Struct. Mol. Biol.* **13**, 264–271.

59 Kamadurai HB, Souphron J, Scott DC, Duda DM, Miller DJ, Stringer D, Piper RC & Schulman BA (2009). *Mol. Cell* **36**, 1095–1102.

60 Rahighi S, Ikeda F, Kawasaki M, Akutsu M, Suzuki N, Kato R, Kensche T, Uejima T, Bloor S, Komander D, Randow F, Wakatsuki S & Dikic I (2009). *Cell* **136**, 1098–1109.

61 Fraczkiewicz R & Braun W (1998). *J. Comput. Chem.* **19**, 319–333.

62 Vijay-Kumar S, Bugg CE & Cook WJ (1987). *J. Mol. Biol.* **194**, 531–544.

FIGURE

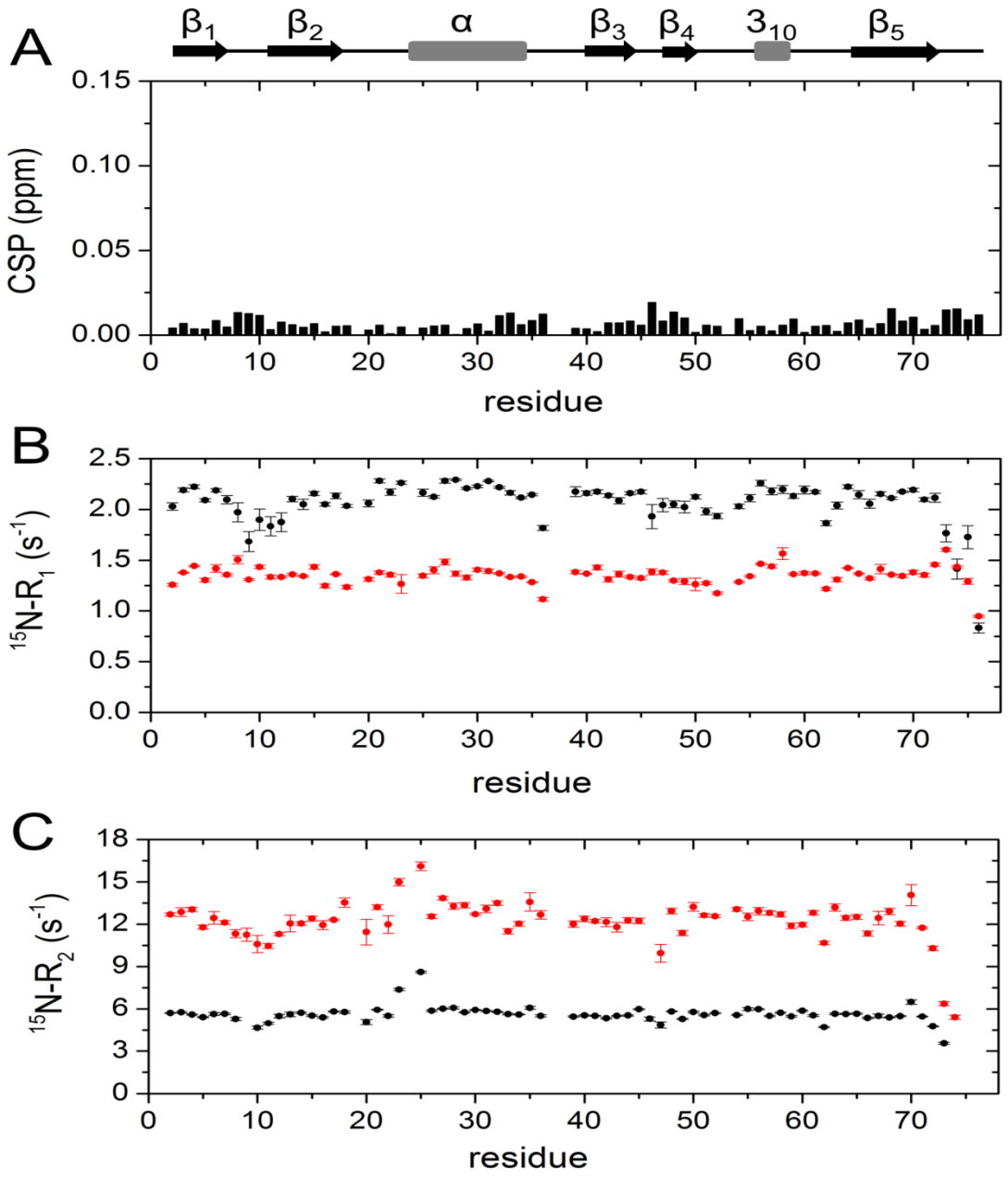


Figure 1. *NMR analysis of ^{15}N-Ub in crowded solution.* A) Amide chemical shift perturbations (CSP) obtained from ^{1}H,^{15}N-HSQC spectra of [^{15}N]Ub in the presence of 200 g/L of Ficoll with respect to [^{15}N]Ub in buffer. The CSP data are plotted as a function of residue number. B,C) ^{15}N–spin relaxation rates of [^{15}N]Ub with/without Ficoll. Shown are ^{15}N-R_1 (B) and ^{15}N-R_2 (C) values as a function of residue number, obtained with (red) and without (black) 200 g/L Ficoll at 25 °C.

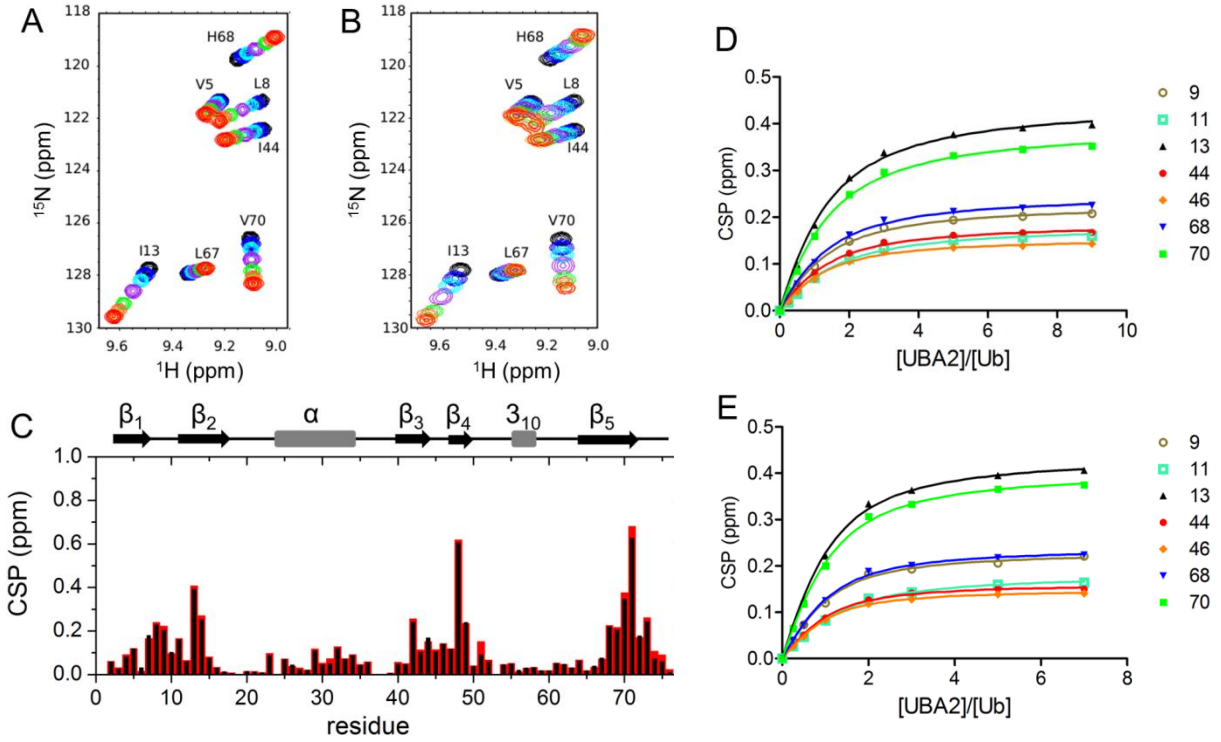


Figure 2. *Ub-UBA2 titration monitored by $^{1}H,^{15}N$-HSQC.* A, B) Overlay of portions of ^{15}N-Ub spectra collected in buffer without (A) and with (B) 200 g/L Ficoll, upon successive additions of UBA2. The displayed spectra correspond to UBA2/Ub molar ratios of 0 (black), 0.25 (blue), 0.5 (cyan), 1 (purple), 2 (green), 3 (tomato), 5 (orange), 7 (red). C) Plot of the chemical shift perturbation (CSP) data from ^{1}H,^{15}N-HSQC spectra of [^{15}N]Ub in the presence of 7-fold molar excess UBA2 with respect to free [^{15}N]Ub, in the absence (black) and presence (red) of 200 g/L Ficoll. D,E) Ub/UBA2 binding isotherms for selected Ub residues based on CSP data collected in buffer without (D) and with (E) 200 g/L Ficoll. Residue numbers are indicated in the legend.

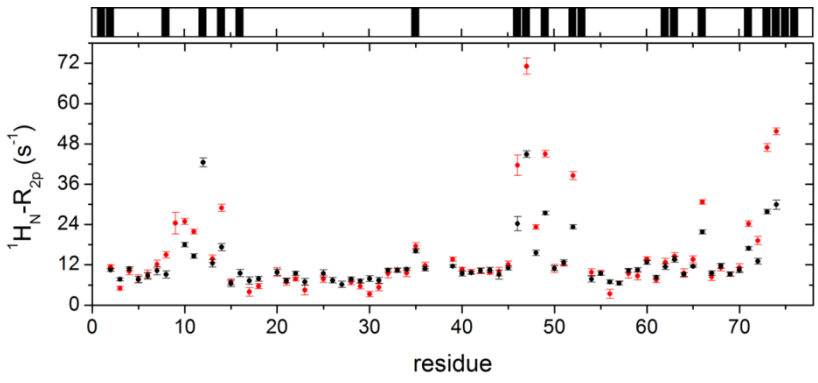


Figure 3. *Solvent PRE on Ub.* Main panel: $^{1}H_N$-R_{2p} rates of Ub in phosphate buffer solution without (black) and with 200 g/L Ficoll (red). Top panel: black bars indicate Ub residues whose backbone nitrogen atoms display significant (> 0.3 Å^2) solvent accessible surface area

(calculations performed using the software GetArea [121] with the protein structure, PDB: 1ubq [122]).

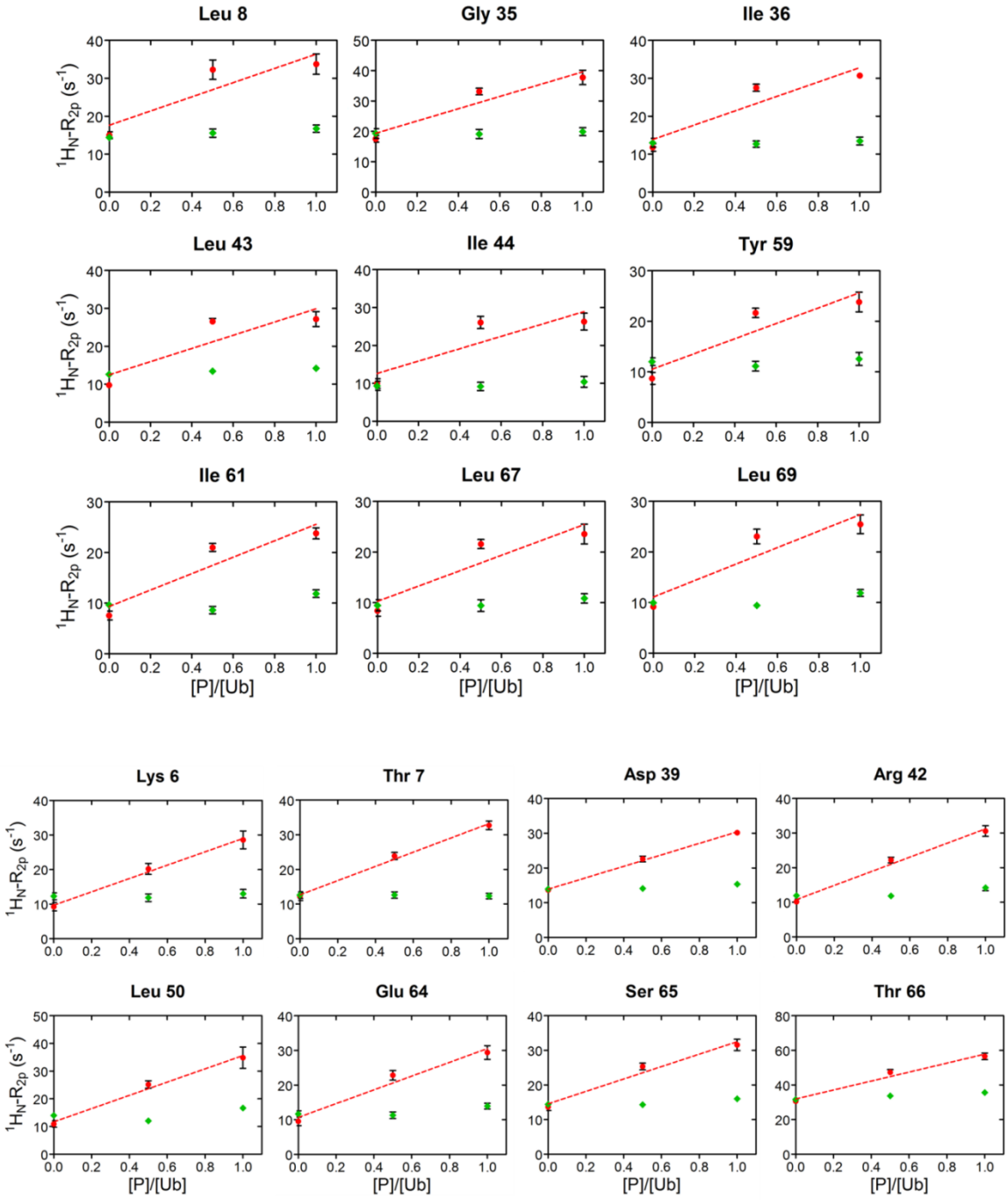


Figure 4. *Solvent PRE on Ub upon UBA2 addition*. A) In red, ^1H_N-R_{2p} rates of a representative group of Ub residues, that increase linearly by addition of UBA2 to Ub. The dashed line corresponds to the least-squares linear fit over three experimental points. As control, ^1H_N-R_{2p}

rates measured upon addition of GB1 to Ub, are reported in green. Rates are plotted versus the protein/Ub molar ratio. B) $^{1}H_{N}$-R_{2p} rates of representative Ub residues, whose value at 1:0.5 was above the linear trace connecting points 1:0 and 1:1, are shown in red. The dashed line corresponds to the least-squares linear fit over three experimental points. $^{1}H_{N}$-R_{2p} rates measured upon addition of GB1 to Ub are shown in green. Rates are plotted versus the protein/Ub molar ratio.

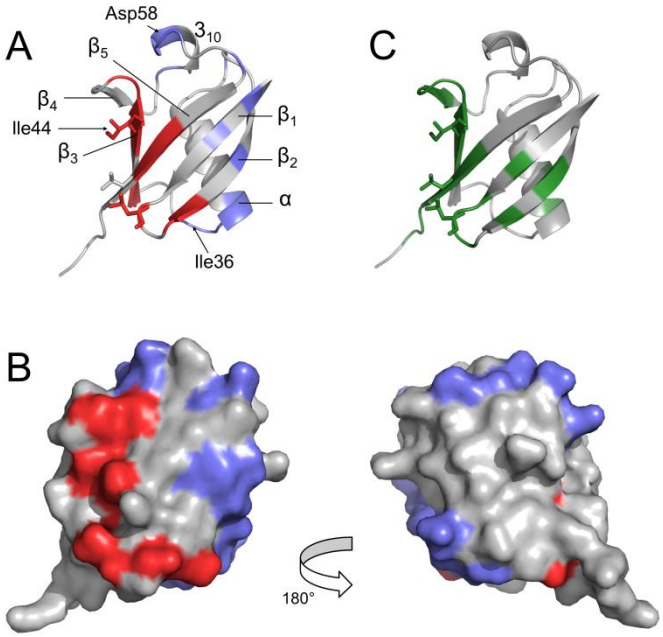


Figure 5. *Mapping of interaction surfaces on Ub*. Surface residues ($^{1}H_{N}$-$R_{2p}(0)$ > 8 s^{-1}) displaying Δ/rΔ > 2 are mapped onto the Ub structure (PDB: 1ubq [122]), represented with a ribbon (A) or surface (B) models; two groups are identified: the red area surrounding the Ile44 patch and the blue surfaces comprising the additional contact sites. C) CSP mapping: Ub residues displaying CSP > 0.1 ppm upon addition of a seven-fold molar excess of UBA2 are colored in green. All data refer to experiments performed in 200g/L Ficoll solution. Residues Leu8, Ile44, and Val70, belonging to the canonical Ile44 patch, are represented in sticks in panels A and C.

Supplementary info of Publication 1

SUPPORTING INFORMATION

Identification of primary and secondary UBA footprints on the surface of ubiquitin in cell-mimicking crowded solution

Francesca Munari[1], Andrea Bortot[1], Serena Zanzoni[1], Mariapina D'Onofrio[1], David Fushman[2] and Michael Assfalg[1]

Materials

Deuterium oxide (99.9%) and $^{15}NH_4Cl$ were purchased from Spectra2000 s.r.l. (Roma, IT). Ficoll® (Poly(sucrose-co-epichlorhydrin)) PM 70 (#F2878) from Sigma–Aldrich. Gadodiamide (gadolinium(III) 5,8-bis(carboxylatomethyl)-2-[2-(methylamino)-2-oxoethyl]-10-oxo-2,5,8,11-tetraazadodecane-1-carboxylate hydrate) was purchased from Selleck Chemicals.

Protein expression and purification

Recombinant human Ub and the UBA2 domain of HHR23A were expressed overnight at 20 °C in *Escherichia coli* Rosetta cells induced with 0.4mM IPTG.

Ub was purified by precipitation of bacterial proteins through addition of perchloric acid to cell lysate. After centrifugation and dialysis of the supernatant in ammonium acetate pH 4.5, Ub was further purified by cation exchange chromatography on SP-Sepharose (GE Healthcare). Fractions of Ub eluted with a 0-0.5 M NaCl gradient were collected and dialyzed into the final NMR buffer.

UBA2 was produced as GST-fusion protein as described by Raasi *et al.*(JMB, 2004, 341, 1367) and purified using Glutathione Sepharose (GE Healthcare) according to standard protocols. The GST tag was then removed by cleavage with thrombin followed by gel-filtration on a Superdex-75 (GE Healthcare). UBA2 was then dialyzed into final

NMR buffer. ^{15}N-labelled samples were produced in M9 minimal medium supplemented with 1g/L $^{15}NH_4Cl$ as sole source of nitrogen.

GB1 (streptococcal protein G B1 domain) protein was produced in *BL21(DE3) Escherichia coli* cells as His-tagged recombinant protein in fusion with C-terminal Green Fluorescent Protein (GFP). Protein purification was performed by IMAC chromatography using a Ni^{2+} charged-chelating sepharose (GE Healthcare). Next, GFP was cleaved from His-GB1 by TEV protease (Sigma Aldrich) and removed by size exclusion chromatography (Sephacryl S-100 HR, GE Healthcare).

Protein samples were concentrated using centrifugal filter units (Millipore). All samples for NMR measurements were prepared in 10mM potassium phosphate buffer at pH 6.8, also containing 0.02% NaN_3 and 8% D_2O.

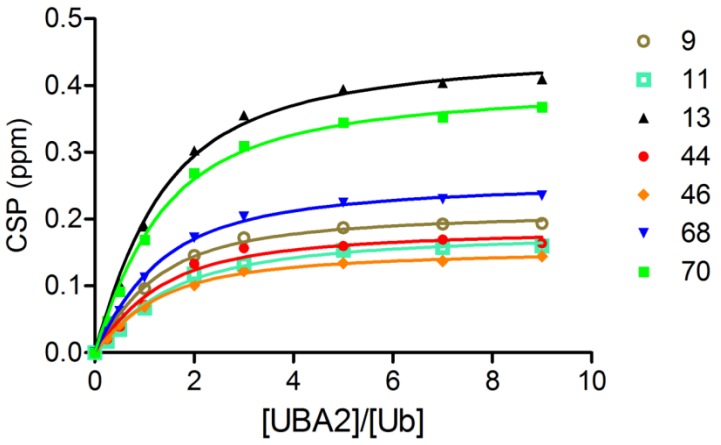


Figure S1. *NMR binding isotherms of Ub/UBA2 in sucrose*. ^{15}N-Ub/UBA2 binding isotherms for selected Ub residues based on CSP data collected in 200g/L sucrose. Residue numbers are indicated in the legend.

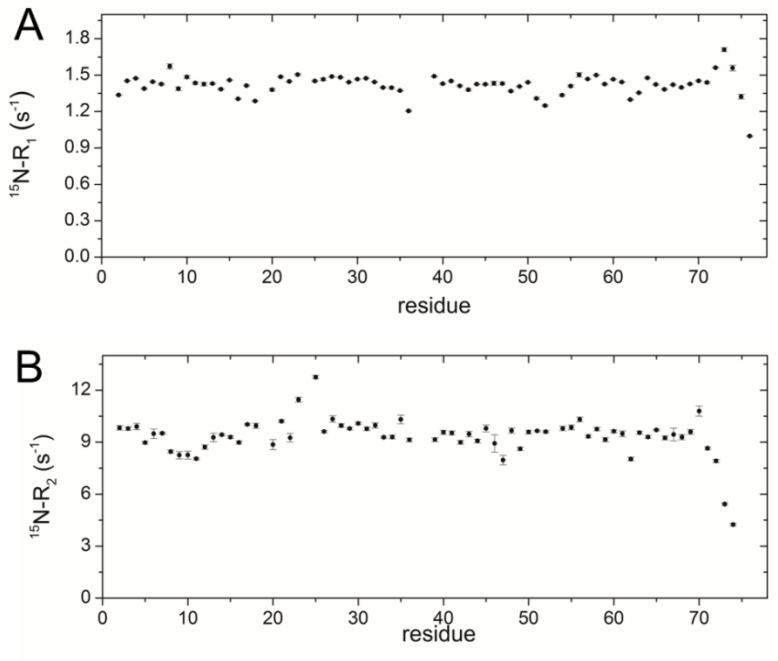


Figure S2. *NMR relaxation data for Ub in sucrose*. A) Main-chain ^{15}N–R_1 values of [^{15}N]Ub in 200g/L sucrose solution. B) Main-chain ^{15}N–R_2 values of [^{15}N]Ub in 200g/L sucrose solution.

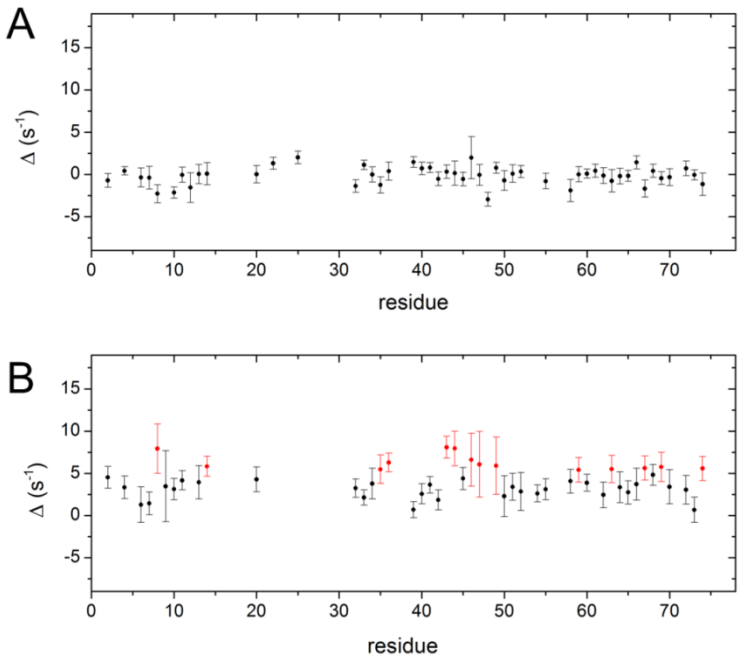


Figure S3. *Analysis of Δ values*. Plot of Δ values against residue number. Error bars are σ_Δ values. Δ is the deviation from a linear concentration dependence of ^1H_N-R_{2p} values. Measurements were performed in A) uncrowded solution, B) 200g/L Ficoll. Only values for exposed residues (^1H_N-$R_{2p}(0) > 8$ s^{-1}) are shown. Δ values exceeding 5 s^{-1} are highlighted in red.

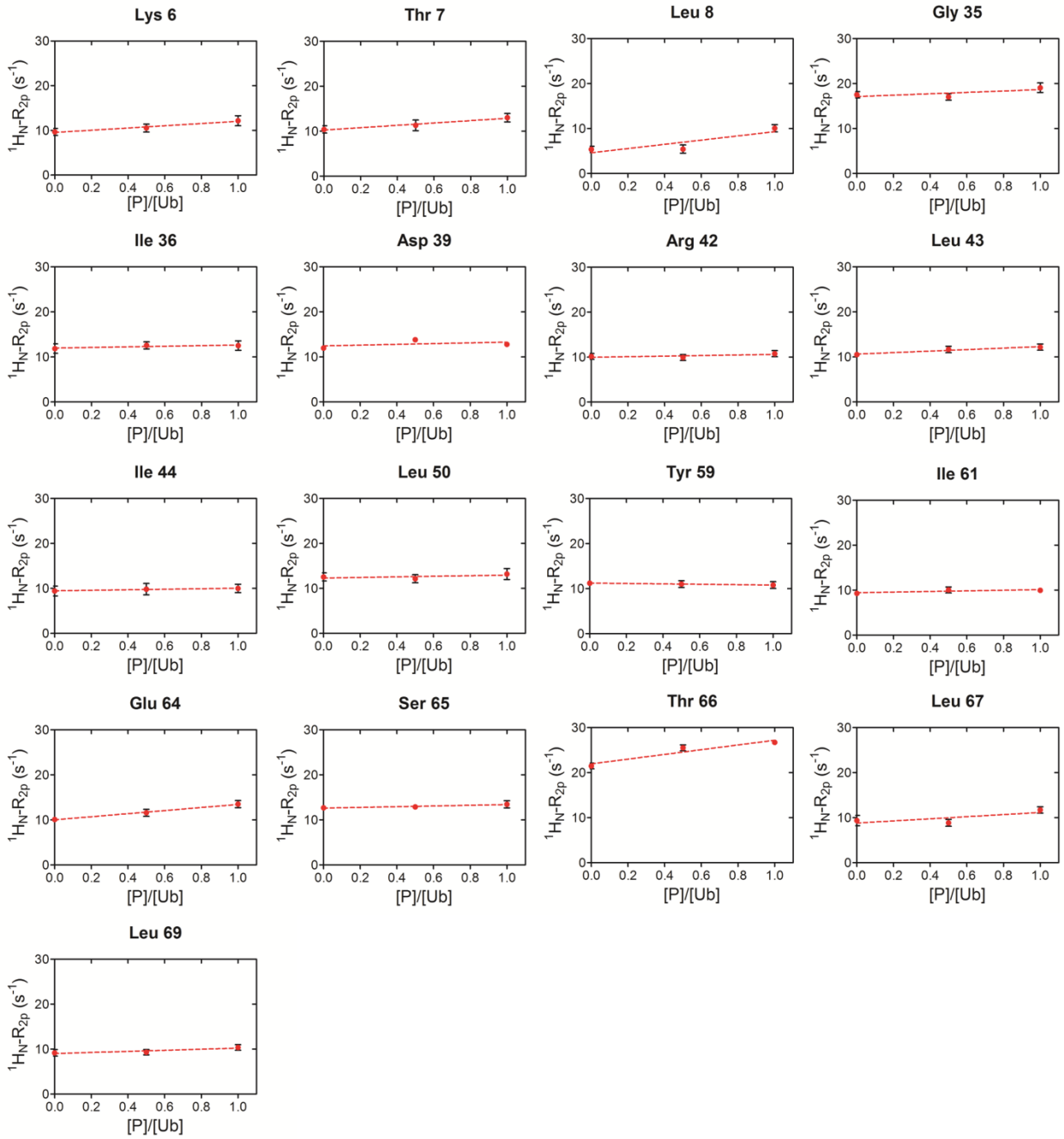


Figure S4. *^1H_N-R_{2p} rate analysis for Ub/UBA2 in uncrowded solution*. Data are shown for the same residues as in Figure 4A,B.

3.2 Publication 2:

"Specific interaction sites determine differential adsorption of protein structural isomers to nanoparticle surfaces."

Chem - A Eur J. 2018. doi:10.1002/chem.201705994

Specific interaction sites determine differential adsorption of protein structural isomers to nanoparticle surfaces

A. Bortot,[a] S. Zanzoni,[a] M. D'Onofrio[a] and M. Assfalg*[a]

Department of Biotechnology, University of Verona, 37134 Verona (Italy)

Abstract

In biological milieus, nanoparticles (NPs) elicit bioactivity upon interaction with proteins. As a result of post-translational modification, proteins occur in a variety of alternative covalent forms, including structural isomers, which present unique molecular surfaces. We aimed at a detailed description of the recognition of protein isomeric species by NP surfaces. The transient adsorption of isomeric Ub dimers to NPs was investigated by solution NMR spectroscopy. Lys63-linked Ub_2 and Lys48-linked Ub_2 adsorbed to large anionic NPs with different affinity, while the binding strength was similar in the case of smaller particles. After incorporation of paramagnetic tags into NPs, the observed site-resolved paramagnetic footprints provided a high-resolution map of the different protein surfaces binding to NPs. The described approach could be extended to further protein isoforms and more specialized NP systems, allowing better control of interactions between NPs and protein targets.

Introduction

Continuous progress in nanotechnological research has aroused tremendous interest in the interactions between biomacromolecules and nanoscale objects.[1–4] Noncovalent protein-nanoparticle (P-NP) interactions play fundamental roles in numerous recent technological developments, particularly in the fields of biomedicine and sensing.[2,3,5–8] Indeed, selective binding of biomacromolecule surfaces by NPs provides a means to elicit specific responses and to control biomolecular communication.[9] Clearly, NPs present several advantages for protein recognition, including a large surface available for interaction, however achieving the specificity found in natural protein-protein (P-P) interactions represents a major challenge.

NPs devoid of complex decorations expose relatively featureless surfaces which appear unsuitable for complementary surface interactions with biomacromolecules. Yet, several examples have been presented in which proteins adsorb to non-functionalized NP surfaces *via* preferential binding sites, thereby providing evidence that even simple NPs are capable of establishing specific interactions with biomacromolecules.[10–16] This notion has important implications for the design of nano-medical tools as well as non-medical devices,[1,6] but many aspects of specific P-NP interactions remain to be addressed and a more fundamental understanding of the principles underlying nano-bio complexations is required.

In order to be able to control the orientation of protein molecules relative to NP substrates, the overall complexity of protein surface recognition needs to be taken into account. Natural proteins present an extremely diversified array of surface features, leading to unique P-P interfaces for each molecular pair or assembly.[17–19] The chemical diversity of protein surfaces results from unique combinations of amino acids forming polypeptide chains as well as from a huge number of possible covalent modifications of amino acid side chains. Post-translational modifications such as phosphorylation, acetylation, methylation, and ubiquitination introduce chemical groups at distinct sites on a protein structure, generating multiple protein products and structural isomers.[20] The specificity of P-P (and predictably of P-NP) interactions crucially depends on the differential recognition of isomeric species by their receptors. Additionally, dynamic

aspects of molecular recognition must be considered, as these may significantly contribute to shape the interaction energy landscape.[21,22]

In our work, we strived to add new insight into protein-NP recognition mechanisms. While recent studies described the adsorption of monomeric globular proteins to NP surfaces,[10–16] here we addressed the specificity of protein-NP interactions in relation to protein structural isomerism and conformational dynamics. The small protein ubiquitin (Ub) represents an ideal model system for such investigations for several reasons. First, Ub was shown to interact specifically with a number of NPs.[10–12] Second, the covalent conjugation of Ub to protein substrates is a major post-translational modification mechanism (ubiquitination).[23] Third, ubiquitination can modify Ub itself, generating polymeric Ub named polyUb chains. Ub units in polyUb are linked by isopeptide bonds and generate polyUb isomers that differ in the site of covalent conjugation.[23] Fourth, polyUb isomers exhibit different conformational dynamics and elicit profoundly different functional responses.[24]

For our investigation, two isomeric forms of dimeric Ub (Ub_2) were presented with a small collection of bare NPs. Among these, lipid nanovesicles (LNV) constitute a highly versatile NP system, they can be easily assembled in a wide range of sizes and obtained in a variety of chemical formulations. LNV are largely employed for biomedical applications and represent models for lipid-coated NPs. We also used silica NP (SNP), and micelles obtained from an amphiphilic synthetic derivative. In order to gain atomic level information on adsorption mechanisms, we adopted NMR spectroscopy. In particular, 1D ^{1}H-NMR was used to estimate the relative strength of the interactions, and site-resolved paramagnetic-NMR was used to identify binding sites. Ub_2 constitutional isomers were found to adsorb differently to anionic LNV of 100 nm size, while differences were less pronounced with the other formulations. The paramagnetic footprints provided insight into the adsorption mechanisms.

Results

Preparation and description of NPs and protein products

We selected lipid nanovesicles, LNV, as our main NP model system due to the straightforward preparation procedure, the possibility to control particle size and physicochemical properties, as well as to incorporate spectroscopic molecular probes. Large unilamellar vesicles (LUV) and small unilamellar vesicles (SUV) were prepared by mixing phospholipids of defined surface charge with cholesterol, thereby obtaining LNV in the size range of 30-120 nm with strongly or moderately negative surface charge, or positive charge (Figure 1, Supplementary Figure S1, and Table 1). In order to produce LNV of smaller size, an amphiphilic compound made from a C_{17} aliphatic chain conjugated to an hydrophilic cage (C_{17}-AAZTA) was used to form micelles of ~7 nm diameter (Scheme S1, Figure S1, Table 1). The AAZTA cage can host metal ions which can act as spectroscopic probes. Finally, we purchased silica nanoparticles (SNP) of similar size as SUV, bearing negative surface charge (Supp. Figure S1, Table 1).

Monomeric human Ub was produced by recombinant expression. Ub is a small 76-amino acid polypeptide which folds into a highly stable β-grasp conformation. A solvent-exposed hydrophobic area, centered around residues Leu8, Ile44, and Val70 (Ile44 patch, yellow in Figure 2),[24,25] stands out from the predominantly polar protein surface. The side chains of seven lysine residues are on the surface and each one can potentially participate in the formation of an isopeptide bond with the carboxyl group of another Ub molecule (Figure 2B), forming polyUb chains. We selected Lys48- and Lys63-linked polyUb since these linkages are the most characterized and relevant of all types of conjugation in natural chains. Ub dimers, Ub_2, were produced by enzyme-catalyzed synthesis combining mutants Ub(K48R or K63R) and Ub(D77) in order to prevent the formation of longer polyUb chains. $^{K63}Ub_2$ adopts an extended conformation, in which the only intersubunit contact is the covalent linkage (Figure 2C).[26] $^{K48}Ub_2$ occurs mainly as a compact structure, with extensive contacts between the Ub units, and buried Ile44 patches. This dimer undergoes rapid conformational equilibrium between the closed state and a more open state that does not present a significant intersubunit interface (Figure 2D).[27]

Strength of protein-NP interactions

Protein NMR spectra are expected to be perturbed when proteins adsorb to NPs.[28] The most dramatic effect is the strong broadening or cancellation of protein signals upon association with NPs, which originates from the reduced tumbling rate of the protein when forming a larger assembly. In case of dynamic association equilibria, the type of perturbation to the observable signal is determined by the chemical exchange regime, but signal intensities are always reduced. If the interaction takes place in the slow exchange regime, it becomes possible to use signal intensities to quantify the adsorbed protein molecules.[29] In the faster regimes, signal intensities are not linearly dependent on the population fraction of the observed species, but their reduction upon successive addition of NPs can be semi-quantitatively related to binding strength, as shown by lineshape simulation analysis (Supp. Figure S2). The analysis of one-dimensional ^{1}H-NMR spectra is a straightforward and rapid approach to rank protein-NP pairs based on their relative affinity. ^{1}H-NMR experiments were performed on the same samples used for heteronuclear NMR experiments, enabling a direct link between binding curves and perturbation analysis. Herein, binding curves that showed hyperbolic behaviour were fitted to determine pseudo-association constants, *K*, whose values were used to estimate the apparent relative binding strength $f_{63/48} = {}^{K63}K/{}^{K48}K$ of protein isomers to the same substrate.

Titration experiments were performed by recording ^{1}H-NMR spectra on protein samples at successive additions of NPs. The downfield portion of the ^{1}H-NMR spectra displayed signals exclusively from protein molecules, while no signals of NPs appeared in this region. The spectral envelopes were integrated and binding curves were obtained for each combination of protein and NP (Figure 3). After mixing (POPG:POPC:Chl)LUV with Ub, the protein signal intensity decrease appeared rather modest (Figure 3A), indicative of weak binding. The presence of an additional negatively charged residue at the C-terminus in the mutant Ub(D77) further reduced the interaction. Since the two investigated Ub_2 isomers bear the same mutation at the C-terminus, it was interesting to observe that $^{K63}Ub_2$ retained affinity for LUVs, while almost no perturbation was observed in $^{K48}Ub_2$ spectra. Using more negatively charged LUVs, (POPG:Chl)LUV, all proteins were found to be attracted more strongly, except

for Ub(D77) (Figure 3B), which was experiencing electrostatic repulsion due to the negative charge of the terminal aspartate. Again, the $^{K63}Ub_2$ species displayed higher affinity for the NP surface than its isomer $^{K48}Ub_2$ ($f_{63/48}$ = 2.77). The same trend was also found with smaller particles, (POPG:Chl)SUV, although in this case the binding curves for the two dimeric Ub forms appeared less separated (Figure 3C) and the relative binding strength was $f_{63/48}$ = 1.61. Fluorescence titration data for the binding to (POPG:Chl)LUV confirmed the result of NMR titrations (Supp. Figure S3).

Similar to the case of (POPG:Chl)SUV, both monomeric and dimeric Ub were attracted by SNP displaying negative ζ-potential and almost identical size as SUVs (Figure 3D). The relative affinity of $^{K48}Ub_2$ and $^{K63}Ub_2$ was similar ($f_{63/48}$ = 1.54) to that observed with SUV. When C_{17}-AAZTA micelles, displaying the same ζ-potential as SUV, were presented to dimeric Ub, we observed no difference in the behaviour of the two isomers (Figure 3E), as indicated by the value $f_{63/48}$ = 0.96. Finally, we investigated the interaction of positively charged (DOTAP:DOPE)LUV with both monomeric and dimeric Ub. While there was almost no adsorption of wild-type Ub, the other protein molecules displayed significant interaction with the NPs (Figure 3F). However, the binding curves departed significantly from a hyperbolic trend after the first few titration steps, and intensities reached a plateau with only partial signal cancellation. This behaviour can be ascribed to the loss of colloidal stability and particle precipitation. Thus, it remained unclear whether Ub_2 species interacted similarly or differently with these NPs.

Identification of NP binding sites on protein surfaces

Given the observation that Ub species could interact differently with the diverse NPs, we set to explain the specificity of binding by determining the protein surface regions contacting the NPs. Site-resolved perturbations can be observed in heteronuclear correlation NMR spectra such as $^{1}H,^{15}N$-HSQC. Although the signals of NP-bound proteins are generally invisible due to excessive line broadening, bound states can be accessible to NMR experiments in case of transient association equilibria due to exchange-mediated signal averaging.[30,31] Under favorable conditions, chemical shift (peak position) changes are detectable, providing a straightforward means to identify

residues whose chemical environment is perturbed upon interaction of the protein with the NP.[28,31,32] While powerful, the chemical shift approach has the drawback that it does not allow to distinguish primary from secondary interactions, nor P-P interfaces from P-NP binding sites. Thus, we adopted the recently described paramagnetic perturbation mapping method.[31] In this approach, paramagnetic relaxation agents are incorporated into the NPs, and paramagnetic-induced spin relaxation enhancements (PRE) are detected. PRE effects are distance dependent and thus contain information about the protein residues that get closer to the NP surface.

We focused our binding site investigation on two LNVs that elicited different binding behaviour of $^{K48}Ub_2$ and $^{K63}Ub_2$. The two protein isomers had very distinct affinity for (POPG:Chl)LUV, while they produced identical binding curves when mixed with C_{17}-AAZTA micelles. Importantly, paramagnetic probes could be introduced easily into both NPs: in the case of LUV, part of the cholesterol molecules were replaced by doxyl cholestane, while in the case of micelles, diamagnetic Y^{3+}-bound C_{17}-AAZTA compounds were in part replaced by Gd^{3+}-bound analogues.[33] Both the aminoxyl radical and the trivalent Gd ion have isotropic or nearly isotropic magnetic susceptibility tensors, thereby causing negligible paramagnetic shifts.[34–37] The unpaired electrons, however, relax slowly and determine strong PRE on nearby nuclear spins.[34,35] Residue-specific PRE effects observed in protein spectra after addition of paramagnetic-labelled NPs can be ascribed exclusively to direct protein-NP interactions.

PRE effects were monitored by evaluating HSQC peak intensity attenuations, with respect to the corresponding diamagnetic references, at varying protein:lipid molar ratios. For each titration series, the protein:lipid molar ratios that showed clear attenuation patterns and that were obtained with good signal/noise were selected for detailed analysis (Figure 4). The addition of paramagnetic LUV to [^{15}N]Ub caused region-specific HSQC peak intensity attenuations (Figure 4A). The smallest intensity ratios (I_{para}/I_{dia} <0.8, where I_{para} is measured in the presence of paramagnetic LUV, and I_{dia} in the presence of diamagnetic LUV) were found for residues 8, 9, 44, 46, 47, 49, 69, 71, 72, 73, forming three sub-regions (displayed in orange in Figure 5A) that overlapped with the Ile44 patch. These experiments were repeated with dimeric Ub. In order to resolve the spectroscopic equivalence of the two Ub units, only the proximal

unit (carrying a free C-terminus) was ^{15}N-labelled. By inspecting the intensity attenuation profiles of $^{K63}Ub_2$, it appeared that the strongest attenuations were localized in all of the sub-regions I-III, including residues 8-11, 47-49, 71, 73-76 (Figure 4B), again partly overlapping with the Ile44 patch (Figure 5B). In the case of $^{K48}Ub_2$, the strongest attenuations were observed for residues 44, 46, 47, 49, 71-73, while no effect was observed in sub-region I at any protein:lipid ratio (Figure 4C, 5C).

Finally, both dimeric Ub were presented with paramagnetic C_{17}-AAZTA micelles. The attenuation profiles appeared rather similar in this case. Significant intensity attenuations were found for residues in the polypeptide stretches corresponding to sub-regions I-III (Figure 4D).

Discussion

Recent studies demonstrated the possibility to obtain high-resolution details about nano-bio interfaces.[10,30,31,38–40] It emerged that simple protein molecules could adsorb to NPs *via* specific binding sites, in analogy with the mode of recognition of natural protein partners. In the present work, we considered additional layers of complexity in protein-NP recognition by interrogating the adsorption mechanisms of dynamic isomeric protein species to diverse NP surfaces. Both structural isomerism and conformational dynamics potentially modulate protein-NP interactions and their elucidation should enable better control of binding specificity and selectivity.

The protein Ub undergoes transient association to negatively charged NPs without experiencing significant conformational changes.[38] Adsorption is mediated by a preferential binding surface of Ub that coincides approximately with the canonical hydrophobic patch involved in the recognition of many Ub-binding partners.[38] The presence of several cationic amino acid side chains in this area suggests that the interaction is supported by electrostatic attraction. Since Ub can form polymers, it represents an ideal system to study the effects of intersubunit linkage and conformational dynamics on its adsorption to NP surfaces.

We mixed Ub dimers with a diverse set of NPs and evaluated the binding strength based on protein signal intensities as a function of NP concentration (Figure 3). With

the smallest particles of this study, C_{17}-AAZTA micelles, we observed identical binding curves for the two isomers $^{K48}Ub_2$ and $^{K63}Ub_2$, indicating that they bound with the same affinity. The $^{K63}Ub_2$ species displayed stronger affinity to negatively charged SNP than $^{K48}Ub_2$. With particles of different chemical composition and electrokinetic surface potential but similar size, (POPG:Chl)SUV, the relative binding affinity was almost unchanged. However, the difference in binding was even more evident when using LUVs, with $f_{63/48}$ = 2.77 for (POPG:Chl)LUV and $f_{63/48}$ = 7.14 for (POPG:POPC:Chl)LUV. These data indicate that the tendency of Ub dimers to bind to NPs is related to the topology of the Ub chain, i.e. the position of the isopeptide bond. In fact, the two Ub_2 species are constitutional isomers, therefore the differences must originate from structural diversity. It appears that small, featureless NPs are not capable to differentially recognize the isomers, while larger particles characterized by low surface curvature are less ideally suited to bind $^{K48}Ub_2$ in comparison to $^{K63}Ub_2$. Interestingly, both Ub_2 species also associated with positively charged NPs, although we could not establish if there was any preferential binding due to the loss of colloidal stability at increasing NP concentration.

In order to understand the molecular determinants of the differential recognition of Ub dimers by negatively charged NP surfaces, residue-resolved NMR experiments were performed. We introduced paramagnetic tags into LUVs and micelles, and analyzed the PRE effects at individual amino acid positions in terms of intensity attenuations. Based on the observed intensity profiles, we found that the interactions of monomeric and dimeric Ub to LNVs involved defined surface regions (Figure 4, 5), consistent with specific P-NP binding. Residues experiencing the strongest effects invariably localized to a surface area almost coincident with the Ile44 patch. In fact, the perturbations mapped to two close regions and residues between them experienced minor or no PRE effect. Based on the limited size of the individual surface areas and on previous studies on Ub/NP interactions,[31,38] it can be concluded that the two patches do not constitute separate binding sites but they belong to the same interface. The reduced PRE effects on residues Arg42 or Val70, located between patches, is attributable to the positioning of the corresponding amide atoms deeper inside the protein framework resulting in a larger distance from the NP surface in the bound state. It must also be noted that, on

average, we expect a NP-bound protein molecule to experience the PRE effect predominantly from a single paramagnetic center.

The intensity attenuation profiles determined for Ub_2 species in the presence of micelles appeared almost identical, explaining the observed comparable binding affinity. On the contrary, differences could be observed for Ub_2 isomers binding to LUVs. First, perturbations were absent for $^{K48}Ub_2$ in sub-region I, while they were present in the case of $^{K63}Ub_2$. Second, PRE effects could be detected in the C-terminal amino acids of $^{K63}Ub_2$ but not of $^{K48}Ub_2$. It should be emphasized that $^{K48}Ub_2$ adopts preferentially a compact conformation in which the Ile44 patch is not accessible. The observation of perturbations in this area indicates that the binding competent conformation is an extended one in which the interdomain interface is lost. Opening of the compact structure must occur through conformational isomerization. Yet, the absence of observable intensity attenuations in the proximal domain's sub-region I indicates that the linkage topology determines a reduced accessibility of that area, resulting in decreased affinity of this dimer compared to its structural isomer. The modified orientation of $^{K48}Ub_2$ compared to $^{K63}Ub_2$ might also explain the different attenuation pattern at the C-terminus. The dissimilar patterns observed in the case of micelles and of LUVs suggest that NP surface curvature had a major influence on the involved contact surfaces of the Lys48-linked dimer. Therefore, in agreement with several previous studies, our results confirm that NP size and surface charge have a profound impact on protein adsorption.[16,41,42]

Selectivity of P-NP interactions can be attained by optimal tuning of the interaction forces. The adsorption of Ub to LNVs bearing different surface charge suggests that Coulombic forces contribute significantly to such interactions, and indeed the identified recognition sites display positive electrostatic potential.[18] Electrostatic interactions have been shown to play an important role in the binding of several proteins to surface-charged NPs or NPs modified with charged ligands.[3,43–45] Importantly, such interactions have been found to provide a significant level of selectivity, whose molecular basis is protein charge anisotropy.[46–48] Region-specific short-range (hydrogen bonding, dipole-dipole, hydrophobic) forces may further enhance the selectivity of the interactions.[49] Herein, we have shown that NMR perturbation

mapping can serve as a fundamental tool to describe surface chemistry anisotropy and to identify recognition patches on the surfaces of closely related protein molecules, such as isomeric products of post-translational modification, also in the presence of conformational dynamics.

It is important to remark that it remains as yet unknown whether Ub dimers form compact monolayers around the NP surface, in analogy with the monomeric protein,[11,29] or rather more complex adsorbed layers are formed. Indeed, in addition to steric repulsion between protein molecules, it is possible that intermolecular interactions are established at the NP surface. In the latter case, the linkage topology could have significant influence on the supramolecular arrangements because it would constrain the relative orientation of interaction sites.

Finally, we note that the possibility to exploit NMR perturbation mapping is dependent on the exchange regime of the P-NP interaction. Associations of weak and moderate strength (down to the low micromolar range[50]) are readily accessible, while strong interactions would remain excluded because they normally entail slow exchange kinetics. Because general paradigms concerning protein-targeted NPs are yet to be established, the range of application of the approach remains undetermined. In its current implementation, the described methodology appears suitable for a number of applications such as sensing, diagnostics, and bioanalytics, while its potential as a complementary tool for the improvement of NP-based drugs is linked to the specific features of the target biological pathways.

Conclusions

It is now widely accepted that most nanomaterials display bioactivity as a consequence of their interaction with biomolecules,[1,51,52] however the specificity and selectivity of such interactions remain poorly explored. The capability of NPs to distinguish closely related covalent forms of proteins, such as structural isomers, has important implications for the design of nanomaterials and could be exploited for the development of NP devices eliciting specific biological responses.

Distinct covalent forms of proteins, produced by post-translational modification, display different recognition properties and perform different functions.[53] This notion

is well exemplified by the Ub system, in which differently conjugated polymeric chains have unique binding features and play distinct roles in cellular biology.[24,54] In our work, we found that structural isomers of polyUb interacted with negatively charged NP surfaces through preferential binding sites. Interestingly, the adsorption of structural isomers of Ub_2 to large NPs involved different surface patches, resulting in reduced binding affinity of $^{K48}Ub_2$ compared to $^{K63}Ub_2$. We further showed that recognition of $^{K48}Ub_2$ required the dynamic interconversion between the prevalent 'closed' conformation and the binding competent 'open' state.

In conclusion, by use of NMR experiments, we were able to provide site-resolved binding information on complex protein-NP adsorption equilibria. We provided evidence of the role of structural isomerism and conformational isomerization in the regulation of adsorption, contributing to a better understanding of the principles governing protein recognition by NP surfaces.

In our work, we used a simplified model system with charged nanoparticles to illustrate the differential binding of protein isomers, however the approach could be applied to more specialized NP systems, exploiting the possibility to tailor functional groups at the NP surface to increase biomolecular binding specificity and selectivity. In the particular case of Ub, appropriately designed NPs could act as artificial receptors for certain isomeric forms, with possible applications in biomedicine and bioanalytics.

Experimental Section

Materials

1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) sodium salt (POPG), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP) chloride salt, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), and cholesterol (Chl) were purchased from Avanti Polar. Ludox® TMA colloidal silica nanoparticles and 3β-doxyl-5α-cholestane (DCh) were purchased from Sigma-Aldrich. 6-amino-6-methylperhydro-1,4-diazepinetetraacetic acid (AAZTA) was kindly provided by Prof. Silvio Aime

(University of Torino). All other reagents necessary for this study were from Sigma-Aldrich.

Polyubiquitin synthesis and SDS-PAGE analysis

Unlabeled and uniformly ^{15}N-labeled recombinant Ub and protein mutants, Ub(K48R), Ub(K63R) and Ub(D77), were expressed and purified as described.[26] Segmentally isotope-labeled Ub_2 chains were synthesized as described in Varadan et al.[27] using enzymes E1 and E2 in overnight reactions at 37 °C added with ATP, TCEP, and an ATP reconstituting cocktail. Recombinant human His-tagged E1 and GST-tagged E2-25K enzymes were used to obtain Lys48-linked Ub_2 while recombinant human His-tagged E1, yeast His-tagged Mms2 and yeast GST-tagged Ubc13 enzymes were used to produce Lys63-linked Ub_2 molecules. Products of polyUb synthesis were inspected by SDS-PAGE. Proximal ^{15}N-labeled Ub_2 (here referred to as Ub_2-P) was synthesized by using ^{15}N Ub(D77) and unlabeled Ub (K48R or K63R). Ub_2 was separated from unreacted Ub monomers at the end of the reaction by cation-exchange chromatography and the Ub_2 concentration was determined from UV absorbance.

Preparation of nanoparticle suspensions

Lipid vesicles were prepared by dissolving powdered pure phospholipids in chloroform/methanol (2:1, v/v) to form a homogeneous solution. The lipid mixture was dried under a nitrogen gas flux in a round-bottom glass flask and the resulting lipid film was then dispersed in 10 mM phosphate buffer (K_2HPO_4/KH_2PO_4) at pH 6.8, to a final stock concentration of ~120 mM in lipid. Large unilamellar vesicles, LUV, were formed by pressure extrusion using a hand-held miniextruder (Avanti, Alabaster, AL), by repeated passage through polycarbonate filters of 100 nm pore size after prefiltering through membranes of 1 μm and 400 nm pore sizes. The following LUV formulations were prepared: POPG:POPC:Chl (40:40:20, w/w), POPG:Chl (80:20, w/w), DOTAP:DOPE (50:50, w/w). Paramagnetic LUV were prepared with the following composition: POPG:Chl:DCh (80:10:10, w/w), simply replacing cholesterol molecules with the aminoxyl derivative in the initial lipid mixture. Small unilamellar vesicles, SUV, were prepared using a probe tip sonicator (Misonix, NY) without cooling at 40 W until the solution became clear. After sonication, SUV solutions were centrifuged to

sediment small metal pieces released by the sonication tip. The vesicle dispersions were stored in the dark at room temperature until analysis.

Ludox® TMA silica nanoparticles were obtained from the commercial source as 36 wt% suspension in H_2O. The nanoparticle suspensions were diluted with ultrapure water (Milli-Q) to the desired concentration (30–40 mg/mL), dialyzed for 24 h at 4 °C against 10 mM potassium phosphate buffer, pH 6.8, with a 10-kDa regenerated cellulose membrane, and finally filtered with 0.22 μm Durapore membrane. Nanoparticle concentration in the purified sample was determined by weighing a dried aliquot of the solution.

The gadolinium- and yttrium-bound amphiphilic AAZTA derivatives, prepared as described previously,[55] were kindly supplied by Prof. Silvio Aime. The C_{17}-AAZTA micelles were obtained by dissolving the powder in 10 mM potassium phosphate buffered solution, pH 6.8, at a concentration higher than the critical micellar concentration (cmc = 0.108 mM). Appropriate mixtures of Y^{3+}-bound and Gd^{3+}-bound molecules were used.

Dynamic light scattering

The size and polydispersity of NPs were determined using dynamic light scattering (DLS). DLS measurements were performed with a Zetasizer Nano ZS instrument (Malvern Instruments, USA) operating at λ = 633 nm with backscatter detection at 173°. The colloidal solutions were allowed to equilibrate for 10 min at the measurement temperature of 25 °C before starting acquisition.

All samples displayed monomodal and monodisperse size distributions. NP size was obtained after cumulant fit of autocorrelation functions and is given as intensity-weighted harmonic mean particle diameter (Z-Average). Each value corresponds to the average of three separate measurements and the experimental uncertainty is indicated as standard deviation.

For ζ-potential measurements, NP dispersions in 10 mM potassium phosphate buffer were placed in a zeta dip cell. Measurements were performed with automatically determined number of runs and applied voltage. All measurements were of good

technical quality based on inspection of phase and frequency plots. ζ-potential was evaluated from electrophoretic mobility applying the Henry equation.

NMR spectroscopy

NMR experiments were run at 600.13 MHz on an Avance III 600 spectrometer (Bruker) equipped with a triple resonance TCI cryoprobe and z-axis gradient.

The protein concentration was 0.2 mM in experiments with lipid vesicles and 0.1 mM in experiments with TMA NPs. The measurement temperature was 25 °C.

One-dimensional ^{1}H-NMR experiments were acquired with a standard pulse sequence incorporating the excitation sculpting water suppression scheme. A total of 128 transients were acquired over a spectral width of 12019 Hz and 32768 complex points, using a recycle delay of 2 s. The spectra were processed applying an exponential window function prior to the Fourier transformation.

Typical ^{1}H,^{15}N-HSQC experiments were recorded with a data matrix consisting of 2048 (F2, ^{1}H) × 128 (F1, ^{15}N) complex points. The number of scans was 8 and the relaxation delay 2 s. Spectral windows of 7211 (F2) and 2189 (F1) Hz were used.

Titration data analysis

NMR signal intensities, *I*, were measured on protein samples containing increasing amounts of NPs. *I* was obtained as the total integral of the ^{1}H-NMR spectral envelope in the amide proton region and plotted as a function of molar ratios. Titration data were found to display hyperbolic-like behaviour and were analyzed by least-squares fitting of the following function, corresponding to a Langmuir binding model:

$$y = \left(C_P + C_L + K^{-1} - \sqrt{\left(C_P + C_L + K^{-1}\right)^2 - 4C_P C_L} \right) / (2C_P)$$

where $y = 1\text{-}I/I_0$, with I and I_0 corresponding to the intensity determined in the presence or absence of NPs, respectively. y would correspond to the fraction of bound protein in the absence of broadening of the free protein signal due to exchange. K is a fitting parameter correlated with the binding strength and corresponding to the association constant if y is equal to the bound protein fraction. C_P and C_L are the total concentrations of protein and lipid (or NP), respectively. The concentration of the initial

protein solutions was 200 or 100 μM. Dilution-corrected values for C_P and C_L were used at successive titration steps.

The stoichiometry was not introduced as fitting parameter to simplify the analysis. Ratios of *K* values are reported throughout the text in place of the single values of *K* because this parameter does not represent the actual association constant. *K* ratios were used as a semi-quantitative measure of relative binding strength. Fitting of experimental data was performed with the software GraphPad Prism 5.00 (GraphPad Software Inc., La Jolla, CA, USA).

The first titration points (where the exchange rates are lowest) for each protein:NP pair were used to obtain a crude estimate of the maximum adsorption capacity (the maximum number of protein molecules bound per NP, N_{max}).[29] We found that experimentally-derived values were generally in agreement with those calculated from simple geometric considerations, suggesting that compact monolayers of proteins formed around NP surfaces.

Acknowledgements

This work was supported by the Joint Projects 2017 funding program of the University of Verona. We thank "Centro Piattaforme Tecnologiche" of the University of Verona for giving access to NMR and DLS instrumentation. SZ received a fellowship grant (Assegno di Ricerca) from the Department of Biotechnology.

Conflict of interest

The authors declare no conflict of interest.

Keywords: biomolecular recognition • nano-bio interface • NMR spectroscopy • paramagnetic relaxation enhancement • protein-nanoparticle interactions

REFERENCES

[1] M. Mahmoudi, I. Lynch, M. R. Ejtehadi, M. P. Monopoli, F. B. Bombelli, S. Laurent, *Chem. Rev.* **2011**, *111*, 5610–5637.

[2] I. Lynch, K. A. Dawson, *Nano Today* **2008**, *3*, 40–47.

[3] D. F. Moyano, V. M. Rotello, *Langmuir* **2011**, *27*, 10376–10385.

[4] I. Nasir, W. Fatih, A. Svensson, D. Radu, S. Linse, C. Cabaleiro Lago, M. Lundqvist, *PLOS ONE* **2015**, *10*, e0136687.

[5] C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2001**, *40*, 4128–4158.

[6] J. E. Gagner, S. Shrivastava, X. Qian, J. S. Dordick, R. W. Siegel, *J. Phys. Chem. Lett.* **2012**, *3*, 3149–3158.

[7] C. Cabaleiro-Lago, F. Quinlan-Pluck, I. Lynch, S. Lindman, A. M. Minogue, E. Thulin, D. M. Walsh, K. A. Dawson, S. Linse, *J. Am. Chem. Soc.* **2008**, *130*, 15437–15443.

[8] B. Pelaz, C. Alexiou, R. A. Alvarez-Puebla, F. Alves, A. M. Andrews, S. Ashraf, L. P. Balogh, L. Ballerini, A. Bestetti, C. Brendel, et al., *ACS Nano* **2017**, *11*, 2313–2381.

[9] C.-C. You, A. Verma, V. M. Rotello, *Soft Matter* **2006**, *2*, 190.

[10] S. Zanzoni, A. Ceccon, M. Assfalg, R. K. Singh, D. Fushman, M. D'Onofrio, *Nanoscale* **2015**, *7*, 7197–7205.

[11] L. Calzolai, F. Franchini, D. Gilliland, F. Rossi, *Nano Lett.* **2010**, *10*, 3101–3105.

[12] V. Mangini, M. Dell'Aglio, A. D. Stradis, A. D. Giacomo, O. D. Pascale, G. Natile, F. Arnesano, *Chem. - Eur. J.* **2014**, *20*, 10745–10751.

[13] G. Brancolini, D. B. Kokh, L. Calzolai, R. C. Wade, S. Corni, *ACS Nano* **2012**, *6*, 9863–9878.

[14] G. Brancolini, A. Corazza, M. Vuano, F. Fogolari, M. C. Mimmi, V. Bellotti, M. Stoppini, S. Corni, G. Esposito, *ACS Nano* **2015**, *9*, 2600–2613.

[15] M. Calvaresi, F. Arnesano, S. Bonacchi, A. Bottoni, V. Calò, S. Conte, G. Falini, S. Fermani, M. Losacco, M. Montalti, et al., *ACS Nano* **2014**, *8*, 1871–1877.

[16] S. Shrivastava, J. H. Nuffer, R. W. Siegel, J. S. Dordick, *Nano Lett.* **2012**, *12*, 1583–1587.

[17] S. Jones, J. M. Thornton, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13–20.

[18] M. Mallon, S. Dutt, T. Schrader, P. B. Crowley, *ChemBioChem* **2016**, *17*, 774–783.

[19] J. Janin, R. P. Bahadur, P. Chakrabarti, *Q. Rev. Biophys.* **2008**, *41*, 133–180.

[20] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, *Angew. Chem. Int. Ed.* **2005**, *44*, 7342–

7372.

[21] D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5*, 789–796.

[22] C.-S. Goh, D. Milburn, M. Gerstein, *Curr. Opin. Struct. Biol.* **2004**, *14*, 104–109.

[23] A. Hershko, A. Ciechanover, *Annu. Rev. Biochem.* **1998**, *67*, 425–479.

[24] C. M. Pickart, D. Fushman, *Curr. Opin. Chem. Biol.* **2004**, *8*, 610–616.

[25] R. Beal, Q. Deveraux, G. Xia, M. Rechsteiner, C. Pickart, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 861–866.

[26] R. Varadan, M. Assfalg, A. Haririnia, S. Raasi, C. Pickart, D. Fushman, *J. Biol. Chem.* **2004**, *279*, 7055–7063.

[27] R. Varadan, O. Walker, C. Pickart, D. Fushman, *J. Mol. Biol.* **2002**, *324*, 637–647.

[28] M. Assfalg, L. Ragona, K. Pagano, M. D'Onofrio, S. Zanzoni, S. Tomaselli, H. Molinari, *Biochim. Biophys. Acta - Proteins Proteomics* **2015**, *1864*,102-14.

[29] A. Wang, K. Vangala, T. Vo, D. Zhang, N. C. Fitzkee, *J. Phys. Chem. C* **2014**, *118*, 8134–8142.

[30] A. Ceccon, M. Lelli, M. D'Onofrio, H. Molinari, M. Assfalg, *J. Am. Chem. Soc.* **2014**, *136*, 13158–13161.

[31] S. Zanzoni, M. Pedroni, M. D'Onofrio, A. Speghini, M. Assfalg, *J. Am. Chem. Soc.* **2016**, *138*, 72–75.

[32] S. Shrivastava, S. A. McCallum, J. H. Nuffer, X. Qian, R. W. Siegel, J. S. Dordick, *Langmuir* **2013**, *29*, 10841–10849.

[33] M. D'Onofrio, E. Gianolio, A. Ceccon, F. Arena, S. Zanzoni, D. Fushman, S. Aime, H. Molinari, M. Assfalg, *Chem. - Eur. J.* **2012**, *18*, 9919–9928.

[34] I. Bertini, C. Luchinat, *Coordination Chemistry Reviews: NMR of Paramagnetic Substances*, Lever A. B. P., **1996**.

[35] G. Otting, *Annu. Rev. Biophys.* **2010**, *39*, 387–405.

[36] A. R. Camacho-Zarco, F. Munari, M. Wegstroth, W.-M. Liu, M. Ubbink, S. Becker, M. Zweckstetter, *Angew. Chem. Int. Ed.* **2015**, *54*, 336–339.

[37] R. Varadan, M. Assfalg, D. Fushman, *Methods Enzymol.* **2005**, *399*, 177–192.

[38] A. Ceccon, V. Tugarinov, A. Bax, G. M. Clore, *J. Am. Chem. Soc.* **2016**, *138*, 5789–5792.

[39] A. Ceccon, G. Marius Clore, V. Tugarinov, *J. Biomol. NMR* **2016**, *66*, 1–7.

[40] A. Ceccon, V. Tugarinov, A. J. Boughton, D. Fushman, G. M. Clore, *J. Phys. Chem. Lett.* **2017**, *8*, 2535–2540.

[41] M. Lundqvist, J. Stigler, G. Elia, I. Lynch, T. Cedervall, K. A. Dawson, *Proc. Natl.*

Acad. Sci. U. S. A. **2008**, *105*, 14265–14270.

[42] S. Tenzer, D. Docter, S. Rosfa, A. Wlodarski, J. Kuharev, A. Rekik, S. K. Knauer, C. Bantz, T. Nawroth, C. Bier, et al., *ACS Nano* **2011**, *5*, 7155–7167.

[43] N. O. Fischer, C. M. McIntosh, J. M. Simard, V. M. Rotello, *Proc. Natl. Acad. Sci.* **2002**, *99*, 5018–5023.

[44] A. Verma, V. M. Rotello, *Chem. Commun.* **2005**, 303–312.

[45] L. Shang, G. U. Nienhaus, *Acc. Chem. Res.* **2017**, *50*, 387–395.

[46] K. Chen, Y. Xu, S. Rana, O. R. Miranda, P. L. Dubin, V. M. Rotello, L. Sun, X. Guo, *Biomacromolecules* **2011**, *12*, 2552–2561.

[47] P. Maffre, K. Nienhaus, F. Amin, W. J. Parak, G. U. Nienhaus, *Beilstein J. Nanotechnol.* **2011**, *2*, 374–383.

[48] L. Treuel, S. Brandholt, P. Maffre, S. Wiegele, L. Shang, G. U. Nienhaus, *ACS Nano* **2014**, *8*, 503–513.

[49] K. Chen, S. Rana, D. F. Moyano, Y. Xu, X. Guo, V. M. Rotello, *Nanoscale* **2014**, *6*, 6492.

[50] K. Srinivasan, S. Parimal, M. M. Lopez, S. A. McCallum, S. M. Cramer, *Langmuir* **2014**, *30*, 13205–13216.

[51] A. A. Shemetov, I. Nabiev, A. Sukhanova, *ACS Nano* **2012**, *6*, 4585–4602.

[52] S. R. Saptarshi, A. Duschl, A. L. Lopata, *J. Nanobiotechnology* **2013**, *11*, 26.

[53] Y. L. Deribe, T. Pawson, I. Dikic, *Nat. Struct. Mol. Biol.* **2010**, *17*, 666–672.

[54] D. Komander, M. Rape, *Annu. Rev. Biochem.* **2012**, *81*, 203–229.

[55] E. Gianolio, G. B. Giovenzana, D. Longo, I. Longo, I. Menegotto, S. Aime, *Chem. - Eur. J.* **2007**, *13*, 5785–5797.

FIGURES

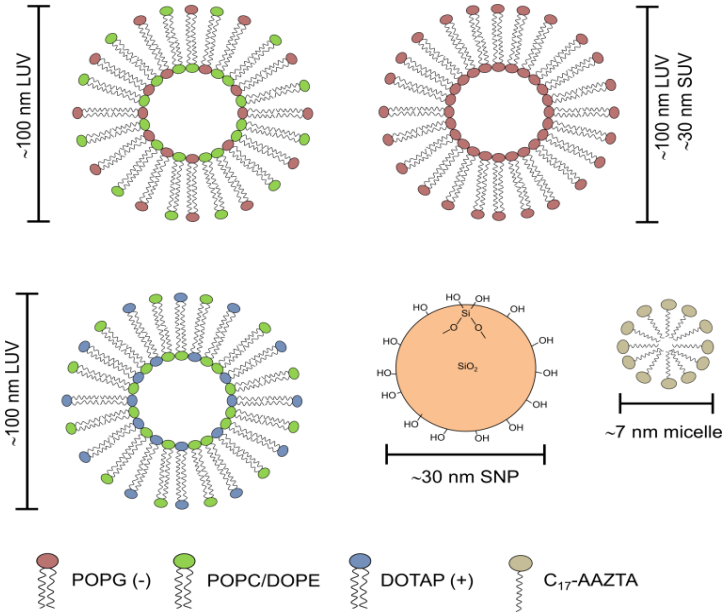


Figure 1. *Illustration of nanoparticles used in this study*. LUV, large unilamellar vesicles; SUV, small unilamellar vesicles; SNP, silica nanoparticles. POPG, POPC, DOPE, and DOTAP are diacyl phospholipids. AAZTA is a multidentate ligand able to chelate inorganic cations. Cholesterol, incorporated in negatively charged lipid vesicles, is not drawn for simplicity.

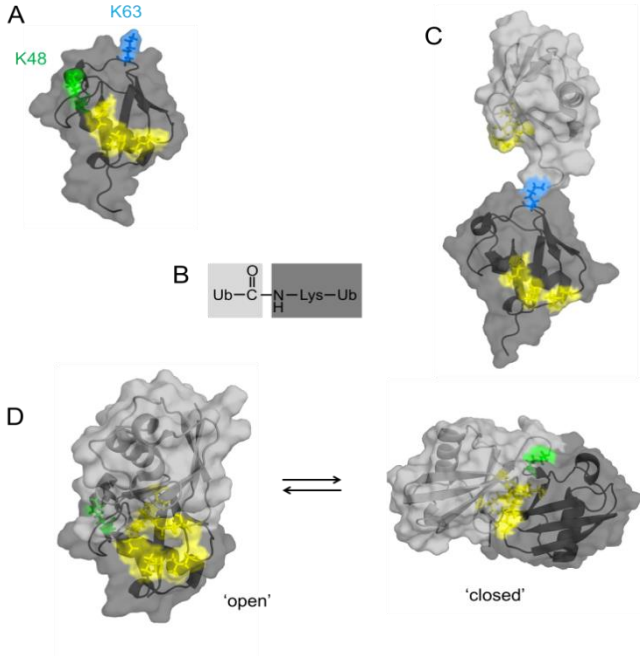


Figure 2. *Illustration of the structures of monomeric and dimeric ubiquitin*. Molecular surfaces and secondary structure cartoon representations of: A) ubiquitin (Ub), C) lysine-63-linked di-

ubiquitin ($^{K63}Ub_2$); D) lysine-48-linked di-ubiquitin ($^{K48}Ub_2$). The chemical formula of the isopeptide linkage between Ub units is displayed in B). The I44 hydrophobic patches are highlighted in yellow. The distal Ub unit in Ub_2 is displayed in light gray, the proximal unit (free C-terminus) is in dark gray. Residues forming the Ile44 patch, Lys63, and Lys48 are represented in sticks. Atomic coordinates were obtained from the Protein Data Bank (PDB id: 1d3z, solution structure of Ub; 2jf5 crystal structure of $^{K63}Ub_2$; 2pe9, solution structure of open state of $^{K48}Ub_2$; 1akk, crystal structure of closed state of $^{K48}Ub_2$).

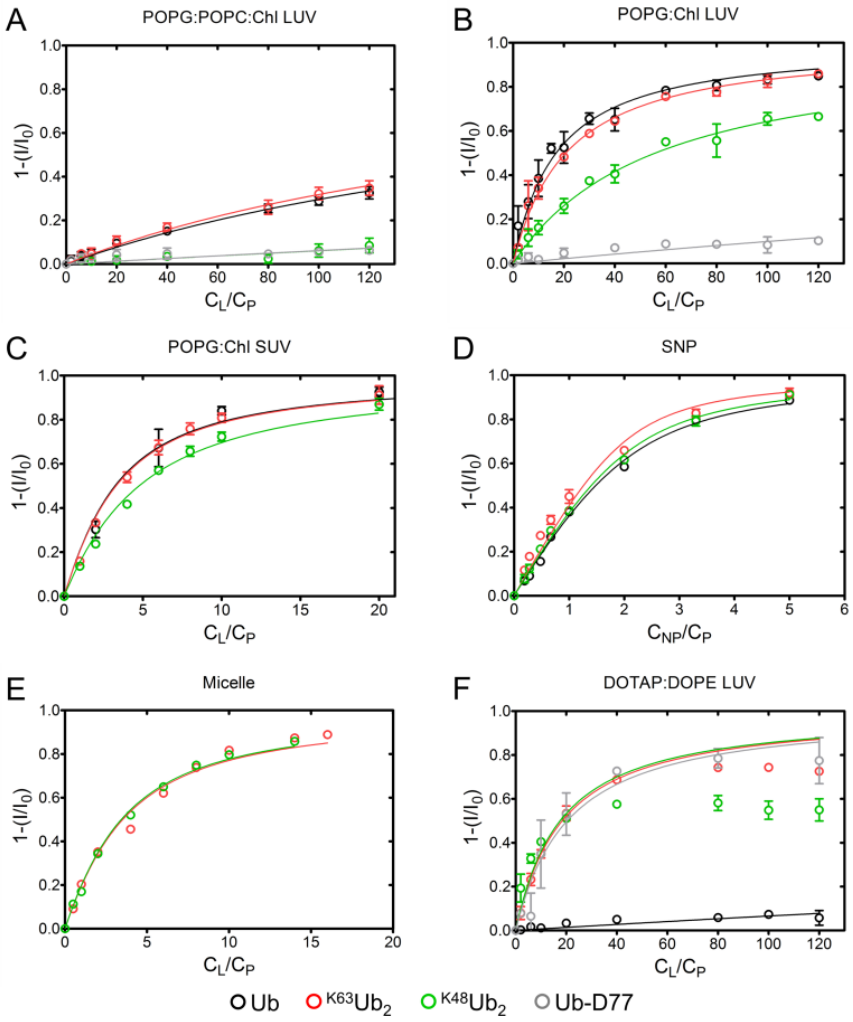


Figure 3. *NMR binding isotherms*. Protein NMR signal intensities were monitored upon stepwise addition of NPs and reported as 1-(I/I_0), where I is the total integral of the ^{1}H-NMR spectral envelope in the amide region, determined in the presence of NPs, and I_0 is the corresponding value measured in the absence of NPs. Data are reported as a function of total NP concentration, C_{NP}, or total lipid concentration, C_L, over total protein concentration, C_P. Protein products

(indicated in the legend) were presented with distinct NP formulations: A) POPG:POPC:Chl(40:40:20) LUV; B) POPG:Chl(80:20) LUV; C) POPG:Chl(80:20) SUV; D) silica NP; E) C_{17}-AAZTA:Y micelles; F) DOTAP:DOPE(50:50) LUV. Each data set (A-E) was obtained in duplicate and experimental points are presented as the mean ± σ. Continuous curves are the best-fit lines calculated according to a Langmuir association model for each data set. In panel F, the fitted curves were determined using only data corresponding to $C_L/C_P < 40$ due to the departure of the second part of the titration from a hyperbolic trend.

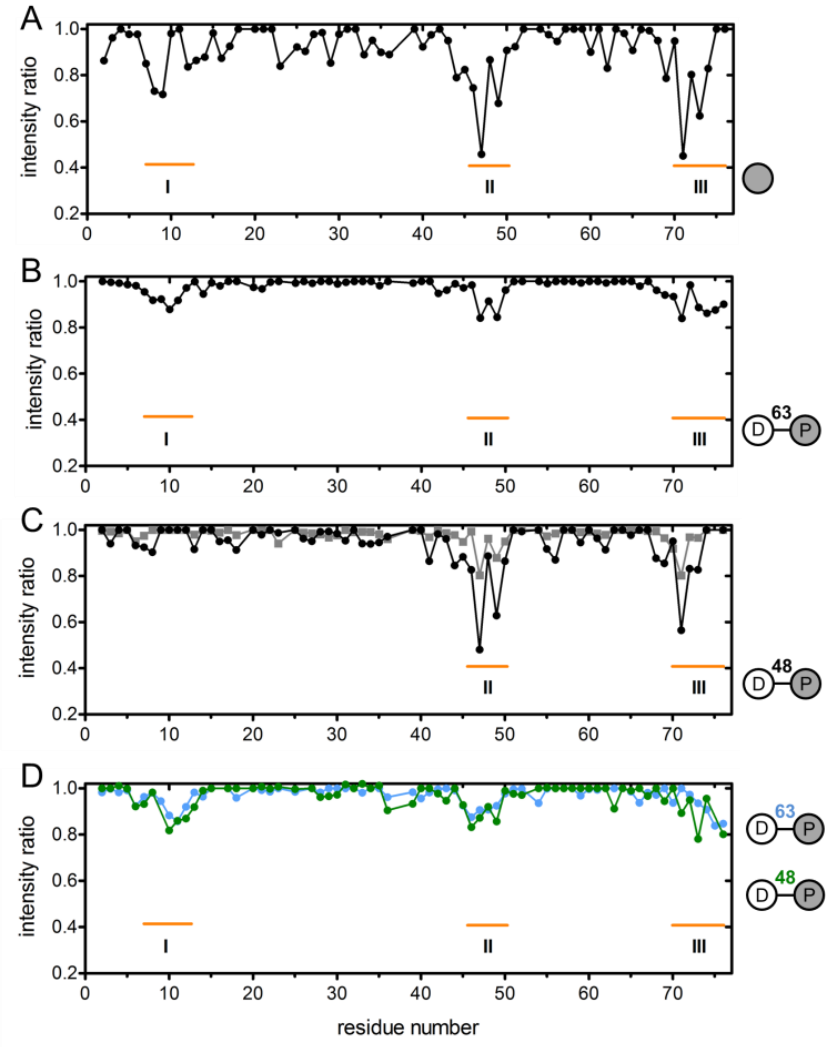


Figure 4. *Paramagnetic perturbations*. Backbone amide $^{1}H,^{15}N$-HSQC peak intensity perturbations measured on 100 μM protein in the presence of (POPG:Chl)LUVs (A-C) or C_{17}-AAZTA micelles (D). Residue-specific intensity ratios, I_{para}/I_{dia}, were calculated from peak heights measured in the presence of paramagnetic NPs (I_{para}) and in the presence of identical

amounts of diamagnetic NPs (I_{dia}). Protein samples were: A) Ub; B) $^{K63}Ub_2$-P; C) $^{K48}Ub_2$-P; D) $^{K63}Ub_2$-P or $^{K48}Ub_2$-P. An illustration of protein samples is displayed next to each panel (D: distal domain, P: proximal domain, grey shading: ^{15}N-labelled unit). Protein:lipid molar ratios were 1:20 (A), 1:1 (B), 1:20 and 1:80 (C), 1:4 (D). Sub-regions of the NP contact surface are labelled I, II, III.

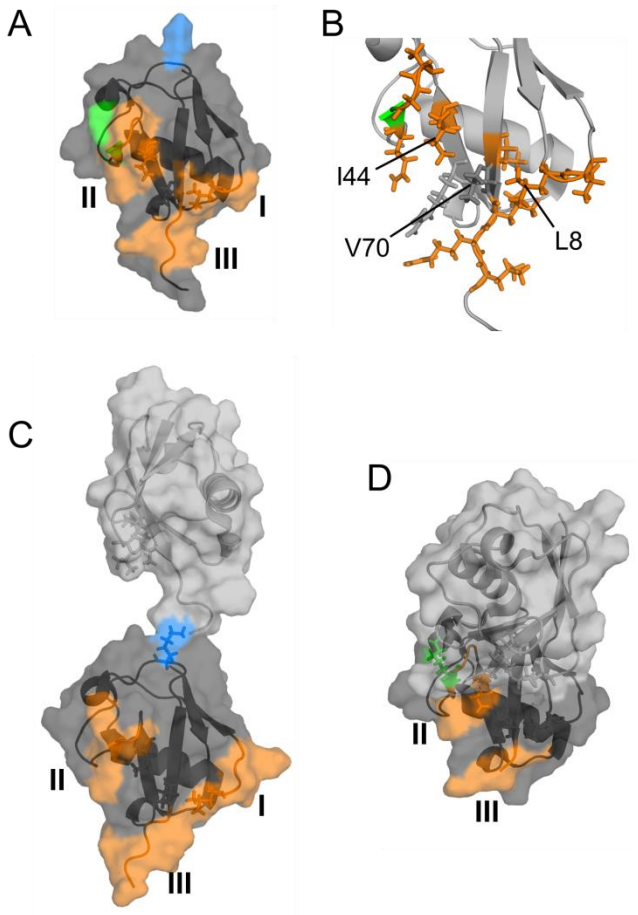


Figure 5. *Paramagnetic perturbation mapping*. Molecular surfaces and secondary structure cartoon representations of: A) ubiquitin (Ub), C) $^{K63}Ub_2$; D) $^{K48}Ub_2$. The patches formed by residues that experienced strong PRE effects are colored in orange. Sub-regions of the NP contact surface are labelled I, II, III. The distal Ub unit in Ub_2 is displayed in light gray, the proximal unit is in dark gray. Lys63 is colored in blue, Lys48 in green. Proximal units have approximately the same orientation in all panels. B) Enlarged view of the Ub region containing residues that experienced PRE effects (displayed as sticks); the residues belonging to the Ile44 patch are in sticks and labelled; Arg42 is in grey sticks. The source of atomic coordinates is indicated in the caption to Figure 2.

Table 1. Hydrodynamic diameter and electrokinetic surface potential (ζ) of nanoparticles.

Nanoparticles	diameter (nm)	PdI[A]	ζ-potential (mV)
POPG:Chl LUV	118.6 ± 1.3	0.067	-85.5 ± 2.8
POPG:POPC:Chl LUV	115.7 ± 1.6	0.068	-65.9 ± 0.9
DOTAP:DOPE LUV	105.7 ± 1.0	0.287	+75.8 ± 5.0
POPG:Chl SUV	34.5 ± 0.3	0.192	-56.2 ± 2.3
SNP (TMA LUDOX)	31.8 ± 1.2	0.175	-23.4 ± 0.8
C_{17}-AAZTA micelle	6.8 ± 1.4	0.040	-57.2 ± 2.5

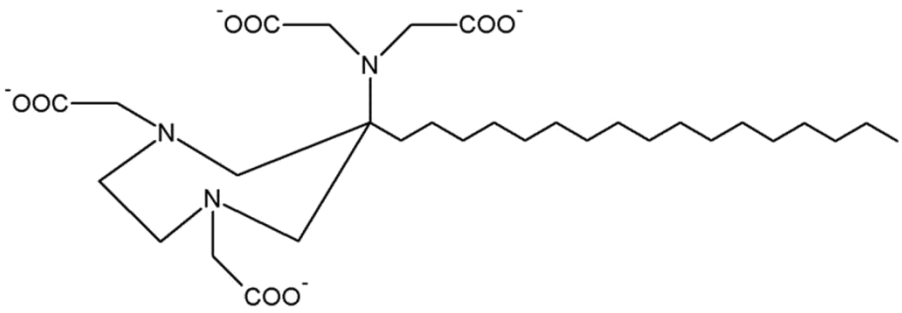
[a] polydispersity index

Supplementary info of Publication 2

SUPPORTING INFORMATION

Specific Interaction Sites Determine Differential Adsorption of Protein Structural Isomers on Nanoparticle Surfaces

CONTENT



Scheme S1. Molecular structure of the amphiphilic compound C_{17}-AAZTA.

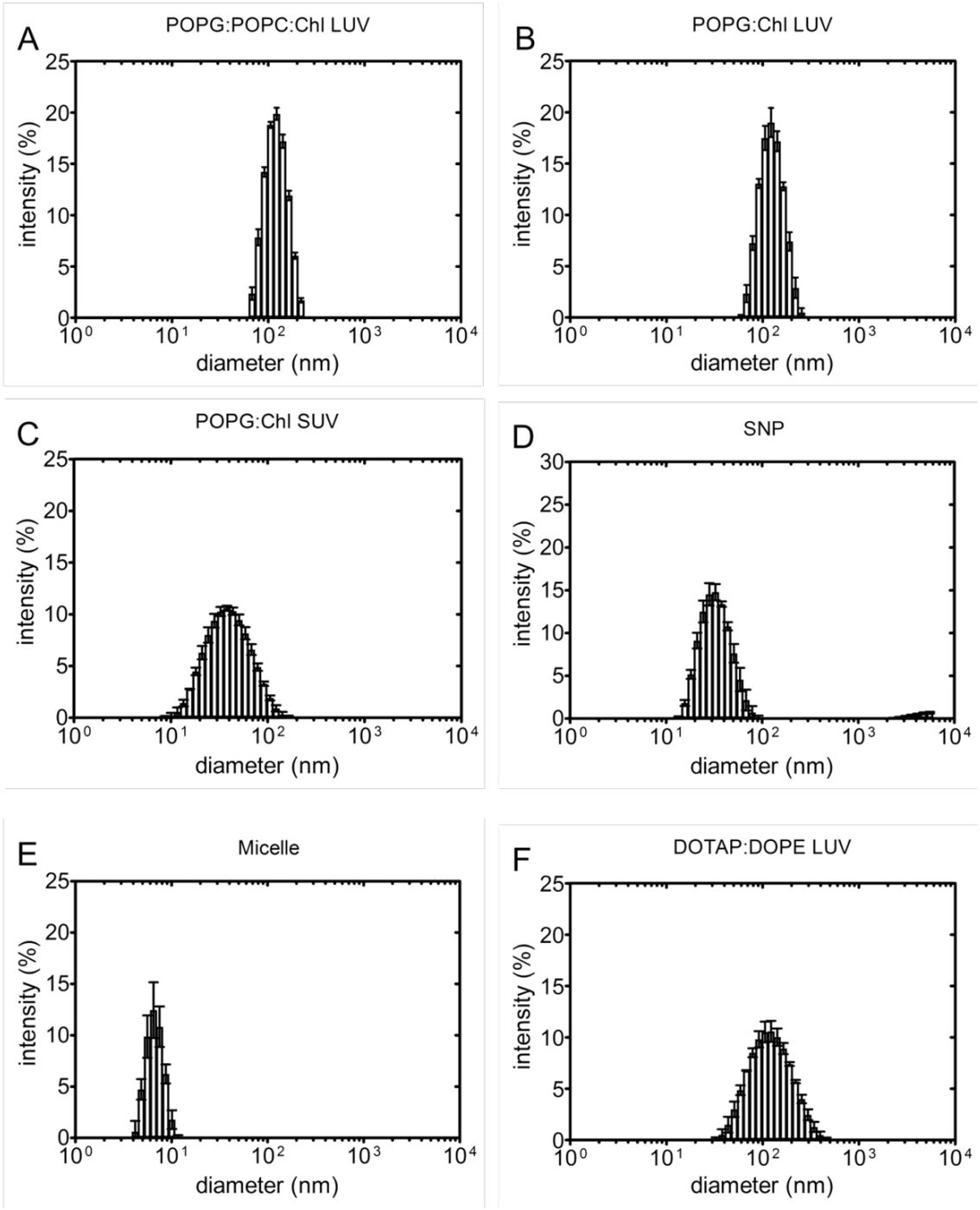


Figure S1. *Nanoparticle size distribution*. Histograms are hydrodynamic diameter distribution plots determined by dynamic light scattering. Error bars are standard deviations obtained from three independent measurements.

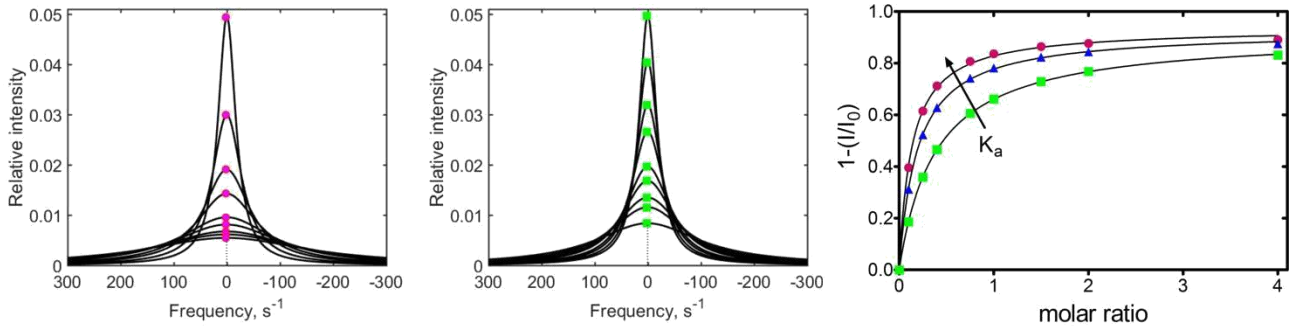


Figure S2. *Lineshape simulation*. The NMR signal of a protein binding to a larger particle was simulated according to a two-state association model. The following kinetic and equilibrium constant values were used: k_{off} = 5.1x10^4 s^{-1}, k_{on} = 4x10^9 M^{-1} s^{-1}, K_a = 8x10^4 M^{-1} (left panel); k_{off} = 4.6x10^5 s^{-1}, k_{on} = 4x10^9 M^{-1} s^{-1}, K_a = 9x10^3 M^{-1} (middle panel). Relaxation rate values were set to: R_2^{free} = 20 s^{-1}, R_2^{bound} = 200 s^{-1}. The chemical shift difference was set to zero to show the behaviour in the absence of chemical exchange line broadening. The right panel displays the normalized simulated peak heights against particle/protein molar ratio and illustrates how different intensity curves relate to binding affinity. The data points in magenta and green correspond to the normalized peak maxima of the first two panels, the blue triangles correspond to a binding equilibrium with K_a value intermediate between those of the magenta and green datasets. Solid lines are best-fit curves of a hyperbolic function (Langmuir association model) to the simulated data points. Simulations were performed with the program LineShapeKin Simulation 4.1. The pseudo-association constant values, *K*, obtained by fitting the simulated data are ranked in the same order as the K_a values used in the simulation, demonstrating that *K* can be used for semi-quantitative comparative analysis of the binding strength.

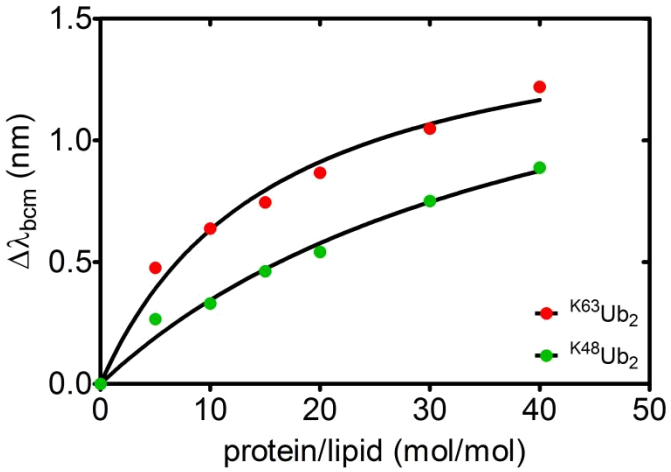


Figure S3. *Fluorimetric titration*. Fluorimetric titration plots display the variation of the barycentric mean fluorescence ($\Delta\lambda_{bcm}$) of (POPG:Chl)LUVs (10 μm lipid, incorporating 10% fluorescent dansyl-DHPE) upon addition of ubiquitin dimers. Solid lines are best-fit hyperbolic curves. Fluorescent LUVs were excited at 336 nm, and fluorescence was monitored between 450 and 600 nm. The barycentric mean fluorescence is given by the expression: $\lambda_{bcm}=\Sigma_\lambda\lambda I(\lambda)/\Sigma_\lambda I(\lambda)$. $\Delta\lambda_{bcm}=\lambda_{bcm}(0)-\lambda_{bcm}$ (step i) and positive values correspond to a blue-shift. Measurements were performed on a FP-8200 spectrophotometer (JASCO).

3.3 Publication 3:

"Alzheimer's disease-associated ubiquitin mutant Ubb^{+1}: properties of the carboxy-terminal domain and its influence on biomolecular interactions"

Int J Biol Macromol. 2018;108:24-31.

Alzheimer's disease-associated ubiquitin mutant Ubb^{+1}: properties of the carboxy-terminal domain and its influence on biomolecular interactions

Francesca Munari, Andrea Bortot, Michael Assfalg, and Mariapina D'Onofrio*

Abstract

Ubb^{+1}, a ubiquitin (Ub) mutant protein originating from misreading of the Ub B gene, is found accumulated in brain tissues of Alzheimer's disease patients. The mutant attracts strong interest due to its possible participation in the molecular events leading to neurodegeneration. Ubb^{+1} is composed of the globular domain of Ub, linked to a 19-residue C-terminal peptide. Based on NMR relaxation and solvent accessibility measurements we obtained new insight into the molecular properties of Ubb^{+1}. We further determined the thermal stability of Ubb^{+1} in the monomeric form, and in Lys48- and Lys63-linked dimers. Finally, we explored the influence of the C-terminal fragment on the interactions of Ubb^{+1} with an isolated UBA2 domain and with membrane mimics.

Our data indicate that the C-terminal fragment of Ubb^{+1} is overall highly flexible, except for a short stretch which appears less solvent-exposed. While influencing the hydrodynamic properties of the globular domain, the fragment does not establish long-lived interactions with the globular domain. It results that the structure and stability of Ub are minimally perturbed by the peptide extension. However, binding to UBA2 and to membrane mimics are both affected, exemplifying possible changes in biomolecular recognition experienced by the disease-associated Ubb^{+1} compared to the wild-type protein.

1. Introduction

Ubiquitin (Ub) is a small and highly conserved protein that is covalently linked to protein targets to regulate numerous fundamental processes in eukaryotic cells, such as progression of cell cycle and division, DNA damage response, organelle biogenesis, protein trafficking, and protein turnover [1–3]. Ubiquitin conjugation, which consists in the formation of an isopeptide bond between the C-terminus of Ub and the ε-amino group of a specific lysine of the substrate protein, is accomplished by the sequential action of Ub-activating (E1), Ub-conjugating (E2), and Ub-ligating (E3) enzymes. Ub itself has seven lysine residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, Lys63) that can act as acceptor sites, leading to the formation of polyUb chains endowed with a variety of linkage types and signaling functions. As an example, Lys48-chains mark substrates for rapid proteasomal degradation, while Lys63-chains play a role in endocytosis, DNA-damage response, cell signaling [1–3], and autophagic clearance of protein aggregates [4]. Ubiquitination is a reversible modification: a variety of deubiquitinating enzymes (DUBs) hydrolyze Ub from the target protein, thereby remodeling or reversing the polyUb signal [5].

Dysfunction of Ub-related enzymes and pathways has been linked to the pathogenesis of severe human diseases, including Alzheimer's (AD), Parkinson's (PD), Huntington's (HD) diseases, Amyotrophic Lateral Sclerosis (ALS), cancer, metabolic syndromes, and genetic disorder [6–8]. In particular, a mutated version of Ub, named Ubb^{+1}, was found specifically accumulated in neurofibrillary tangles, neuropil threads and dystrophic neurites in brain tissues of AD patients [9]. Ubb^{+1} originates from misreading of the Ub B gene: the deletion of a dinucleotide adjacent to the GAGAG motif in the first repeat of the Ub transcript leads to the loss of Ub stop codon and the generation of a new one. This produces Ubb^{+1}, a Ub mutant protein where the carboxy(C)-terminal Gly76 is replaced by a tyrosine and linked to a 19-residue peptide [9]. Molecular misreading is a rare event that originates from mistakes of RNA polymerase II transcription activity, leading to generation of frameshift mRNA and protein mutants in the absence of genetic mutation. In physiological conditions, the Ub-proteasome system (UPS) well compensates for transcription errors by removing aberrant proteins through degradation.

However, proteasomal degradation becomes progressively less efficient with aging, and accumulation of frameshift mutant proteins eventually occurs, such as in the case of AD [10].

Ubb^{+1} maintains the well-structured globular domain of Ub (Fig. 1) [11] and its lysine side chains can act as acceptors for polyUb chain linkage [11–14]. However, as a result of the lack of the C-terminal Gly76, Ubb^{+1} does not modify substrate proteins and instead terminates the elongation of polyUb chains. The resulting Ubb^{+1}-capped polyUb chains ($polyUbb^{+1}$) were shown to inhibit the proteasome [14,15], and were recalcitrant to disassembly mediated by DUBs [14]. Interestingly, it was shown that at low expression levels, Ubb^{+1} can be degraded by the 26S-proteasome, however after exceeding a threshold level of expression, Ubb^{+1} accumulates and inhibits the proteasome in a dose-dependent manner [16]. In neuroblastoma cells, overexpression of Ubb^{+1} leads to neuronal cell death [13]. Due to the essential function of the proteasome in protein turnover and clearance of misfolded proteins, the inhibition of proteasome by Ubb^{+1} species was proposed to be one of the key mechanisms of neuronal toxicity in AD. In particular, interaction of Ubb^{+1} with the enzyme E2-25K was suggested to modulate Aβ neurotoxicity via proteasomal inhibition [17]. Additionally, it was recently suggested that accumulation of extended Ub variants was due to their potent ability to inhibit specific DUBs [18].

Due to a possible participation of Ubb^{+1} in the molecular events leading to neurotoxicity and neurodegeneration in AD, there is large interest in elucidating the structural details of this frameshift mutant Ub protein and the consequent functional differences with respect to the wild-type protein. In our work, we investigated structural and dynamic features of Ubb^{+1} using NMR spectroscopy methods that are particularly suited to explore protein molecules containing flexible domains such as the C-terminal extension of Ubb^{+1}. Calorimetry measurements were carried out to evaluate the contribution of the tail to the protein's thermal stability in both its monomeric and dimeric forms. Finally, the ability of Ubb^{+1} to interact with different biomolecules, in particular the UBA2 domain of the human homologue of the yeast DNA repair protein RAD23 (HHR23A) and membrane mimics, was here explored and characterized.

2. Materials and Methods

2.1 Materials

Deuterium oxide (99.9%), ^{13}C-glucose and $^{15}NH_4Cl$ were purchased from Spectra2000 s.r.l. (Roma, IT). Gadodiamide (gadolinium(III) 5,8-bis(carboxylatomethyl)-2-[2-(methylamino)-2-oxoethyl]-10-oxo-2,5,8,11-tetraazadodecane-1-carboxylate hydrate) was purchased from Selleck Chemicals. Powder cholesterol, phosphatidylglycerol and phosphatidylcholine were purchased from Sigma. 4,4-dimethyl-4-silapentane-1-sulfonic acid-d6 was purchased from Sigma.

2.2 Protein expression and purification

Recombinant human Ub and the UBA2 domain of HHR23A were expressed and purified as described previously [19]. Ub mutants K48R, K63R, and D77 were produced with the same protocol used for wild-type Ub. Human Ubb^{+1}, cloned in pET3 vector, was obtained from PCR extension of the human Ub sequence. The purification of Ubb^{+1} was performed with the same procedure as for Ub, with an additional size exclusion chromatography step.

Homogeneous di-ubiquitin chains (Ub_2) were obtained from Ub mutants, following the strategy described in [20], in overnight enzymatic reactions at 37 °C complemented with ATP, TCEP, and an ATP reconstituting cocktail. Recombinant human His-tagged E1 and GST-tagged E2-25K enzymes were used to obtain Lys48-linked Ub_2 products: Ub(K48R)-^{48}Ub(D77) and Ub(K48R)-$^{48}Ubb^{+1}$. Recombinant human His-tagged E1, yeast His-tagged Mms2 and yeast GST-tagged Ubc13 enzymes were used for to produce Lys63-linked Ub_2 molecules: Ub(K63R)-^{63}Ub(D77) and Ub(K63R)-$^{63}Ubb^{+1}$. E2-25K, Mms2, and Ubc13 enzymes were produced in *Escherichia coli* BL21(DE3), while E1 was expressed in Rosetta cells. Purification of enzymes by affinity chromatography followed standard procedures. Ub_2 molecules were separated from unreacted Ub monomers by SP cation-exchange chromatography and further purified by size exclusion chromatography for calorimetry measurements. Protein samples were concentrated using centrifugal filter units (Millipore).

2.3 Liposome preparation

Liposomes were obtained by dissolving cholesterol (Chl) and phosphatidylglycerol (POPG) or phosphatidylglycerol/phosphatidylcholine (POPG/POPC) phospholipids in chloroform/methanol (2:1, v/v) to form a homogeneous solution. A molar ratio of POPG/Chl 80:20 or POPG/POPC/Chl 40:40:20 was used to produce liposomes with different surface charge. The lipid mixture was then dried under nitrogen flux and the resulting lipid film dispersed in 10 mM potassium phosphate buffer at pH 6.8, to a final lipid concentration of ~120 mM. Vesicles were obtained by pressure extrusion with a hand-held miniextruder (Avanti, Alabaster, AL), by repeated passage through polycarbonate filters of 100 nm pore size, after pre-filtering through membranes of 1 μm and 400 nm pore sizes. The obtained liposomes had hydrodynamic diameter of 120 nm and polydispersity index < 0.1 as determined by dynamic light scattering (DLS) measurements performed with a Zetasizer Nano ZS instrument (Malvern Instruments, USA) at 25 °C.

2.4 NMR Spectroscopy

NMR experiments were acquired at 25 °C on a Bruker Avance III spectrometer, operating at ^{1}H Larmor frequency of 600.13 MHz, equipped with a triple resonance TCI cryogenic probe. NMR data were processed with Topspin 3.2 (Bruker) or NMRpipe [21] and analyzed with the software Sparky (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

All samples for NMR measurements were prepared in 10 mM potassium phosphate aqueous buffer at pH 6.8, also containing protease inhibitors (Sigma) and 8% D_2O.

Sequence-specific backbone resonance assignment of Ubb^{+1} was obtained by analysis of CBCA(CO)NH, HNCACB, HNCO and HN(CA)CO spectra.

Secondary chemical shifts were calculated based on the random coil chemical shifts predicted by the Neighbor Corrected Structural Propensity Calculator [22] or the Camcoil software [23]. 4,4-dimethyl-4-silapentane-1-sulfonic acid-d_6 (DSS-d_6) was used for chemical shift referencing (0.0 ppm). Chemical shift perturbations were

calculated as: CSP = $[(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{0.5}$, where $\Delta\delta_H$ and $\Delta\delta_N$ were the chemical shift changes measured in the ^{1}H and ^{15}N frequency dimensions, respectively.

^{15}N relaxation experiments were performed on 1 mM [^{15}N]Ubb^{+1} protein samples. ^{15}N longitudinal relaxation rates (R_1) were measured using relaxation delays in the range 0.01-1.26 s and ^{15}N transverse relaxation rates (R_2) were measured with relaxation delays in the range 8-224 ms. Steady-state {^{1}H}^{15}N heteronuclear nuclear Overhauser effects (hetNOE) were measured with a 6 s recycle delay. hetNOE values were calculated taking the ratio of peak intensities in saturated and reference spectra.

Transverse ^1H_N paramagnetic relaxation rate enhancements (PRE), ^1H_N-R_{2p}, were obtained from the difference in ^1H_N-R_2 measured on samples containing or not containing 2 mM gadodiamide. The measurements were performed as described previously [19]. Seven relaxation delays between 11.7 and 79 (60) ms were used in experiments acquired on samples without (with) 2 mM gadodiamide and the signal intensity decays were fitted to a single exponential function to obtain the corresponding rates.

The solvent accessibility of Ubb^{+1} backbone was also evaluated through measurement of exchange rates between water and NH protons by performing CLEANEX-PM-FHSQC experiments [24] using mixing times of 10, 25, 50, 75, 100, and 150 ms. The peak intensity (V), measured as a function of mixing time (x), was fitted according to the following equation $V/V_0 = k/(R_{1A,app}+k-R_{1B,app})\cdot\{\exp(-R_{1B,app}\cdot x)-\exp[-(R_{1A,app}+k)\cdot x]\}$ to obtain k, the normalized rate constant related to the exchange rate constant between NH protons and water [24]. V_0 is the intensity of the protein signals in a reference FHSQC experiment, $R_{1A,app}$ and $R_{1B,app}$ are apparent relaxation rates for protein and water, respectively. For $R_{1B,app}$ we used the value of 0.6 s^{-1}, in analogy with previous work [24].

The experimental rotational correlation time constant, τ_c, was obtained from ^{15}N relaxation data R_1, R_2, and hetNOE with the program ROTDIF [25]. Prediction of the τ_c of Ubb^{+1} was done with the HYCUD software [26,27]. First, we generated an ensemble of 1000 random structures with the EOM program [28], using the crystal structure of human Ub [29] and the human Ubb^{+1} sequence as input files. Next, we calculated the

effective τ_c of the globular part of the full-length Ubb^{+1} with the HYCUD program [27]. By using AER of 2.9 Å and initial $\tau_{c,0}$ of 4.02 ns (our experimental value for Ub [19]), the algorithm predicted that the presence of a completely unrestricted C-terminal tail of Ubb^{+1} would raise the τ_c value to 5.7 ± 1.0 ns.

The dissociation constant value for the Ubb^{+1}/UBA2 interaction was obtained by fitting experimental binding isotherms obtained from ^{1}H-^{15}N HSQC-based titration experiments, assuming a one-site binding model and by use of the Matlab program Kdfit [30]. The initial concentration of ^{15}N-Ubb^{+1} was 0.367 mM and that of the titrant UBA2 solution was 10 mM. The reported K_d value is the average of values determined from binding isotherms of seven residues using dilution-corrected protein concentrations.

2.5 Differential scanning calorimetry (DSC)

Thermal denaturation data were acquired with a Nano DSC instrument (TA Instruments Inc.). Samples were heated from 20 to 100 or 110 °C at a scan rate of 1 K min^{-1}. Selected experiments were also performed at 0.5 K min^{-1} to evaluate the scan rate dependence of the thermograms. Reheating runs were carried out to establish the operational reversibility of the thermal unfolding processes. Samples contained 0.05-0.1 mM protein dissolved in 100 mM sodium phosphate, pH 7.4, 100 mM NaCl. Before measurements, sample and reference solutions were properly degassed in an evacuated chamber for 10 minutes at room temperature and carefully loaded into the cells to avoid bubble formation. Calorimetric cells (operating volume 300 μL) were kept under a pressure of 3 atmospheres. Exhaustive cleaning of the cells was undertaken before each experiment. A background scan collected with buffer in both cells was subtracted from each scan. Analysis was performed using NanoAnalyze (TA Instruments Inc.). The calorimetric enthalpy was determined from the total peak integral after baseline correction. The thermal transition midpoint was determined as the temperature corresponding to the peak top.

3. Results and discussion

3.1 The C-terminal amino acids of Ubb^{+1} do not exhibit canonical secondary structure propensities

In a recent work, the NMR structure of Ubb^{+1} (incorporating eight exogenous residues at the N-terminus) was determined [11]. The structure shows a compact globular domain in the region 1-75, corresponding to the Ub moiety, while the remainder of the polypeptide chain is undefined due to the lack of structural information (Fig. 1). However, a residual structure for residues 75-88 was proposed based on the larger values of heteronuclear NOE and smaller values of backbone RMSD in this region, compared to the rest of the terminal peptide [11].

To obtain further insight into the structural properties of the Ubb^{+1} 19-residue extension, we measured secondary chemical shifts, a very sensitive parameter used to probe local conformation [31,32] and structural propensities in unfolded proteins [33]. We thus assigned the backbone resonances of Ubb^{+1}, produced in the absence of affinity tags, by using a series of standard 3D heteronuclear NMR experiments. Then, we calculated secondary chemical shifts as the difference between Ubb^{+1} ^{13}Cα or ^{13}C' chemical shifts and the corresponding random coil values [22,23]. Data shown in Fig. 2 indicate that both ^{13}C' and ^{13}Cα secondary chemical shifts closely match the α and β structures of the globular region of Ubb^{+1}. By contrast, secondary chemical shift values are very close to zero for the 19-residue extension, indicating the absence of α and β secondary structure propensities. This finding was validated by the use of two independent datasets of random coil values [22,23].

3.2 The middle region of the C-terminal domain displays reduced solvent accessibility and mobility

^{15}N-spin relaxation rate constants, R_1 and R_2, and steady state heteronuclear NOE were measured to investigate the backbone dynamics of the 19-residue extension. Data reported in Fig. 3 reflect the modular architecture of the protein: the first region (residues 1-72) reports on the relatively rigid nature of the globular Ub domain, while the C-terminal tail is characterized by increased backbone mobility. From ^{15}N relaxation

data of Ubb^{+1}, we estimated a rotational correlation time constant (τ_c) for the globular domain of 5.77 ± 0.04 ns, much higher than the value of 4.02 ns estimated for Ub [19]. However, the determined τ_c value is in excellent agreement with the value of 5.7 ± 1.0 ns predicted using the HYCUD approach [27]. The prediction was made considering a completely unrestricted mobility of the C-terminal tail of Ubb^{+1} which exerts a dragging effect on the globular part *via* hydrodynamic coupling. The coincidence of experimental and predicted values strongly suggested that interactions between the two protein domains were absent.

While displaying overall larger backbone flexibility, the C-terminal tail did not exhibit uniform dynamics, in agreement with previous conclusions [11]. Of particular interest were residues Asp84 and Arg85, which showed $\{^1H\}^{15}N$-NOE (hetNOE) values that departed from the decreasing trend of the rest of residues in the C-terminal tail (Fig. 3A), suggesting the presence of reduced local backbone mobility. Interestingly, this feature was in qualitative agreement with the degree of solvent accessibility determined by two alternative approaches. In a first approach, we measured amide ^1H paramagnetic relaxation enhancements (PRE) in the presence of the soluble paramagnetic agent gadodiamide (Gd-DTPA-BMA) [34]. Fig. 4A shows the transverse PRE, ^1H_N-R_{2p}, as a function of the amino acid sequence. Data referred to the first part, 1-75, were in agreement with what found previously for Ub [19], with residues 8-14, 45-49, and 72-75 showing large ^1H_N-R_{2p} values, indicative of high accessibility to the paramagnetic cosolute. Interestingly, the 19-residue extension experienced large PRE, although non-uniformly distributed along the chain and with a marked dip in the region 78-88 (^1H_N-R_{2p} < 30 s^{-1}). Therefore, this region appeared less exposed to the solvent than the rest of the tail. The most protected residues of the tail were Asp82, Asp84 and Arg85, exhibiting PRE values of 13.9 ± 0.8 s^{-1}, 14.3 ± 0.3 s^{-1}, and 14.3 ± 0.5 s^{-1}.

To support these conclusions, we performed an independent measurement of amide proton exchange with bulk solvent based on the CLEANEX-PM-FHSQC experiment. In this NMR experiment, solvent accessibility was probed by monitoring water-protein proton exchange taking place during a mixing time which followed a selective water resonance excitation [24]. Fig. 4B shows the determined water-amide proton exchange

rates, *k*, against the Ubb^{+1} sequence. The majority of residues belonging to the globular part were protected from exchange with solvent and therefore did not show measurable signals after the given mixing times. However, resonances corresponding to residues 8-12, 39, 46, 47, 49, 63, and residues within the C-terminal tail were observable, indicative of a significant degree of solvent accessibility. For these residues, we determined the water-amide proton exchange rates. In the 19-residues extension, amino acids 78-87 exhibited the smallest *k* value ($k < 8\ s^{-1}$), confirming that this part of the C-terminal tail was less exposed to the solvent. In particular, the experiment confirmed that residues Asp84 and Arg85, showing *k* values of $0.9 \pm 0.1\ s^{-1}$ and $1.2 \pm 0.1\ s^{-1}$, were the least solvent-accessible residues of the C-terminal peptide.

3.2 Ubb^{+1} displays similar thermal stability as Ub in both monomeric and dimeric species

To determine if the 19-residue extension affected the thermal stability of Ub, differential scanning calorimetry (DSC) experiments were performed. Ub is a highly compact globular protein which undergoes thermal unfolding at elevated temperature [35,36]. According to previous investigations, the denaturation mechanism corresponds to a fully reversible two-state transition when Ub is dissolved in acidic solution [35]. The thermal transition was also found to be reversible in PBS solution, pH 7.4 [37], however some authors reported the onset of protein aggregation and significantly reduced reversibility in neutral pH solution [35,38]. The thermal unfolding of both Lys63- and Lys48-Ub_2 was found to be irreversible [37]. In our work, we established the calorimetric irreversibility of Ub_2 thermal unfolding and the partial reversibility (~30%) of Ub denaturation (Supplementary Fig. S1). The temperature of the peak maximum for Ub displayed an increase of less than 0.5 K when changing the scan rate from 0.5 K min^{-1} to 1 K min^{-1}, indicating little sensitivity of the endotherm to kinetic constraints. However, a slight concentration dependence of the thermograms was observed in the range 50-100 μM. Thus, comparative analyses between wild type and variant proteins were performed at equal scan rates on samples containing identical protein concentrations. To avoid the assumptions or simplifications inherent in model-based analyses, we limited our investigation to a phenomenological analysis of the

experimental data (measured calorimetric parameters are reported in Supplementary Table S1). Indeed, the midpoint of thermal unfolding curves provides an adequate indication of the relative thermal stability of related proteins.[39]

DSC data indicated that the apparent thermal transition midpoint of Ubb^{+1} was unchanged with respect to that of Ub (T_m = 369 K)(Fig. 5 and Table S1). The calorimetric enthalpy values of the two variants differed by 5% (Table S1), which is in the order of experimental uncertainty. Thus, the terminal polypeptide had essentially no influence on the thermal stability of the globular domain. We further inspected the thermal behaviour of Ub_2 molecules. In agreement with previous work [35], our experiments indicated a decreased thermal transition midpoint of both dimeric Ub-^{63}Ub (T_m = 353 K) and Ub-^{48}Ub (T_m = 358 K)(Fig. 5 and Table S1). Because of the reported formation of amyloid-like fibril assemblies [35], the stability of polyUb chains is of utmost relevance in the context of neurodegenerative diseases. Here, we found that substitution of Ub with Ubb^{+1} did not perturb the thermal transition of dimeric species and the change in calorimetric enthalpy was less than 5%, again indicating that the C-terminal tail of Ubb^{+1} does not interact with the globular units. In this respect, it should be noted that Ub-Ubb^{+1} dimers are structurally similar to their wild-type counterparts [12], therefore no changes are introduced by the extension on the intersubunit interfaces. Apparently, the conjugation of the terminal peptide to the flexible C-terminus of Ub has no influence on the intramolecular interactions that govern the folding/unfolding process.

3.3 The affinity of Ubb^{+1} for (HHR23A)UBA2 is enhanced compared to that of Ub

To investigate the influence of the C-terminal extension in Ubb^{+1} on biomolecular recognition, we carried out a binding experiment with a model Ub protein partner: the Ub-associated C-terminal domain (UBA2) of the human homologue of the yeast DNA repair protein Rad23 (HHR23A). Rad23, originally recognized as an important player in nucleotide excision repair, mediates targeting of ubiquitinated proteins to the proteasome for degradation [40]. Ubb^{+1}/UBA2 titration experiments monitored by ^{1}H-^{15}N HSQC spectra were carried out to describe the UBA2 interaction surface on Ubb^{+1} and to estimate the dissociation constant, K_d. Fig. 6A shows the chemical shift

perturbation (CSP) profile of Ubb^{+1} resonances upon binding to UBA2. The most affected regions are centered on residues Leu8, Ile44, and Val70, which form the well-known hydrophobic patch of the Ub surface, involved in several recognition events. The perturbation profile is highly similar to that determined for Ub upon binding to UBA2 [19] suggesting that the supramolecular arrangement was preserved, and for Ubb^{+1} binding to E2-25K enzyme which contains a UBA domain [11]. We found that the 19-residue extension did not experience perturbations, except for residues Glu81 and Asp82, for which a small effect was detected. The analysis of binding isotherms based on CSP data (Fig. 6B) revealed that the affinity of the Ubb^{+1}/UBA2 complex (K_d = 305 ± 18 μM) was slightly larger than that previously determined for the Ub/UBA2 system (K_d = 412 ± 52 μM) [19]. This result is consistent with previous findings on the interaction between E2-25K and Ubb^{+1}, which was found to be stronger than that of E2-25K and Ub. The observed increased affinity could be due to a slightly extended interaction surface area in the minimum-energy complex and/or to enhanced probability of formation of encounter complexes due to transient anchoring of the flexible tail.

3.4 C-terminal amino acids hinder adsorption of Ubb^{+1} to anionic lipid vesicles

Recently, it was shown that Ub binds transiently to liposomes through its hydrophobic patch and adjacent positively charged residues [41]. This finding is of interest considering that transient interactions between Ub and membranes might take place within cells due to the key role of Ub in regulating the sorting of membrane proteins [42] and autophagic processes [43]. Moreover, liposomes have been widely used as membrane mimics to investigate the interaction between proteins and membranes [31,44–46] We thus explored whether the 19-residue extension of Ubb^{+1} could affect the Ub/liposome interaction using an NMR-based binding assay where a series of ^{1}H-spectra of Ub or Ubb^{+1} were acquired in the presence of increasing amounts of negatively charged liposomes. Fig. 6C shows that the amide ^{1}H signal envelope, integrated in the region between 9.0 and 9.8 ppm, progressively decreases as the protein molecules bind to the negatively charged liposomes made of POPG and cholesterol at a molar ratio of 80:20. In comparison to Ub, the signal intensity loss displayed by Ubb^{+1} was much less pronounced, indicating a substantially reduced affinity. Similarly, Ub

experienced weak binding affinity to liposomes bearing reduced negative charge (composed of POPG, POPC and cholesterol at 40:40:20 molar ratio), while almost no interaction was observed between these liposomes and Ubb^{+1}. Clearly, electrostatic attraction plays a fundamental role in the formation of protein-liposome assemblies, therefore it seems likely that the reduced affinity of Ubb^{+1} was due to the acidic character of the C-terminal peptide (the calculated pI for residues Y76-Q95 is 4.7). Thus, translocation of Ubb^{+1} (and of polyUb chains terminated by Ubb^{+1}) within the cellular milieu and in proximity to lipid membranes may be significantly different compared to the wild-type species, possibly resulting in perturbed intracellular communication.

4. Conclusions

Our findings support the notion that the 19-residue extension of Ubb^{+1} lacks a well ordered structure and persistent canonical secondary structure elements. However, we observed that the flexible fragment displayed non-uniform structural and dynamic properties along its amino acid sequence. In particular, the region Asp78-Asp87 was characterized by diminished solvent exposure, with Asp84 and Arg85 being the most protected residues against proton exchange with bulk solvent and among the less accessible to a paramagnetic molecular probe. The relatively larger hetNOE values of Asp84 and Arg85, compared to adjacent residues, further indicated that the polypeptide backbone in those positions experienced reduced mobility on the ps-ns timescale. From analysis of global dynamics based on ^{15}N-spin relaxation data, it emerged that the mobility of the C-terminal tail was independent from that of the globular domain, and persistent interdomain interactions were absent. Hence, it is possible that a reduction in local flexibility might originate from structural constraints imposed by the sequence of amino acids, and in particular by the proximity of two proline residues (Pro83 and Pro90). Consistently with the above picture and with the absence of structural perturbations, inferred from chemical shift analysis, the peptide extension did not affect the protein thermal stability in either monomeric or dimeric species. However, biomolecular recognition was in part perturbed: the interaction of Ubb^{+1} with (HHR23A)UBA2 was characterized by enhanced affinity compared to that of Ub, and

the weak association of Ub with anionic lipid surfaces was significantly reduced in the case of the variant.

It can be concluded that the globular architecture of Ub is sufficiently robust not to be particularly influenced by the unintended covalent conjugation of a peptide to its flexible C-terminus. However, the presence of a fragment that extends by 25% beyond the regular polypeptide length is certainly not without consequences, and the selected interactions investigated in our work exemplify possible changes in biomolecular recognition and diffusive dynamics experienced by disease-associated mutant Ub species, compared to wild-type molecules, in physio(patho)logical conditions.

Acknowledgments

This work was supported by the University of Verona (Progetto Ricerca di Base 2015, to M.D.). F.M. thanks “Fondazione Umberto Veronesi” for granting a postdoctoral fellowship. We acknowledge "Centro Piattaforme Tecnologiche" of the University of Verona for giving access to NMR, DLS, and DSC instrumentation. We thank Prof. Dimitrios Fessas for useful discussions.

Conflicts of interest

The authors declare no conflict of interest.

References

[1]D. Komander, Biochem. Soc. Trans. 37 (2009) 937–953. doi:10.1042/BST0370937.
[2]C.M. Pickart, D. Fushman, Curr. Opin. Chem. Biol. 8 (2004) 610–616. doi:10.1016/j.cbpa.2004.09.009.
[3]J.-F. Trempe, Curr. Opin. Struct. Biol. 21 (2011) 792–801. doi:10.1016/j.sbi.2011.09.009.
[4]J.M.M. Tan, E.S.P. Wong, D.S. Kirkpatrick, O. Pletnikova, H.S. Ko, S.-P. Tay, M.W.L. Ho, J. Troncoso, S.P. Gygi, M.K. Lee, V.L. Dawson, T.M. Dawson, K.-L. Lim, Hum. Mol. Genet. 17 (2008) 431–439. doi:10.1093/hmg/ddm320.
[5]T.E.T. Mevissen, D. Komander, Annu. Rev. Biochem. (2017). doi:10.1146/annurev-biochem-061516-044916.
[6]G. Atkin, H. Paulson, Front. Mol. Neurosci. 7 (2014) 63. doi:10.3389/fnmol.2014.00063.
[7]Q. Huang, M.E. Figueiredo-Pereira, Apoptosis Int. J. Program. Cell Death. 15 (2010) 1292–1311. doi:10.1007/s10495-010-0466-z.
[8]D. Popovic, D. Vucic, I. Dikic, Nat. Med. 20 (2014) 1242–1253. doi:10.1038/nm.3739.
[9]F.W. van Leeuwen, D.P. de Kleijn, H.H. van den Hurk, A. Neubauer, M.A. Sonnemans, J.A. Sluijs, S. Köycü, R.D. Ramdjielal, A. Salehi, G.J. Martens, F.G. Grosveld, J. Peter, H. Burbach, E.M. Hol, Science. 279 (1998) 242–247.
[10]X. Chen, D. Petranovic, Wiley Interdiscip. Rev. Syst. Biol. Med. 8 (2016) 300–313. doi:10.1002/wsbm.1340.
[11]S. Ko, G.B. Kang, S.M. Song, J.-G. Lee, D.Y. Shin, J.-H. Yun, Y. Sheng, C. Cheong, Y.H. Jeon, Y.-K. Jung, C.H. Arrowsmith, G.V. Avvakumov, S. Dhe-Paganon, Y.J. Yoo, S.H. Eom, W. Lee, J. Biol. Chem. 285 (2010) 36070–36080. doi:10.1074/jbc.M110.145219.
[12]M. Chojnacki, D. Zhang, M. Talarowska, P. Gałecki, J. Szemraj, D. Fushman, M.A. Nakasone, FEBS Lett. 590 (2016) 4573–4585. doi:10.1002/1873-3468.12484.
[13]F.M. De Vrij, J.A. Sluijs, L. Gregori, D.F. Fischer, W.T. Hermens, D. Goldgaber, J. Verhaagen, F.W. Van Leeuwen, E.M. Hol, FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. 15 (2001) 2680–2688. doi:10.1096/fj.01-0438com.
[14]Y.A. Lam, C.M. Pickart, A. Alban, M. Landon, C. Jamieson, R. Ramage, R.J. Mayer, R. Layfield, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 9902–9906. doi:10.1073/pnas.170173897.
[15] K. Lindsten, F.M.S. de Vrij, L.G.G.C. Verhoef, D.F. Fischer, F.W. van Leeuwen, E.M. Hol, M.G. Masucci, N.P. Dantuma, J. Cell Biol. 157 (2002) 417–427. doi:10.1083/jcb.200111034.
[16]P. van Tijn, F.M.S. de Vrij, K.G. Schuurman, N.P. Dantuma, D.F. Fischer, F.W. van Leeuwen, E.M. Hol, J. Cell Sci. 120 (2007) 1615–1623. doi:10.1242/jcs.03438.
[17]S. Song, S.-Y. Kim, Y.-M. Hong, D.-G. Jo, J.-Y. Lee, S.M. Shim, C.-W. Chung, S.J. Seo, Y.J. Yoo, J.-Y. Koh, M.C. Lee, A.J. Yates, H. Ichijo, Y.-K. Jung, Mol. Cell. 12 (2003) 553–563.
[18]D. Krutauz, N. Reis, M.A. Nakasone, P. Siman, D. Zhang, D.S. Kirkpatrick, S.P. Gygi, A. Brik, D. Fushman, M.H. Glickman, Nat. Chem. Biol. 10 (2014) 664–670. doi:10.1038/nchembio.1574.
[19]F. Munari, A. Bortot, S. Zanzoni, M. D'Onofrio, D. Fushman, M. Assfalg, FEBS Lett. 591 (2017) 979–990. doi:10.1002/1873-3468.12615.
[20]C.M. Pickart, S. Raasi, Methods Enzymol. 399 (2005) 21–36. doi:10.1016/S0076-6879(05)99002-2.
[21]F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, J. Biomol. NMR. 6 (1995) 277–293.
[22]K. Tamiola, B. Acar, F.A.A. Mulder, J. Am. Chem. Soc. 132 (2010) 18000–18003. doi:10.1021/ja105656t.
[23]A. De Simone, A. Cavalli, S.-T.D. Hsu, W. Vranken, M. Vendruscolo, J. Am. Chem. Soc. 131 (2009) 16332–16333. doi:10.1021/ja904937a.

[24]T.L. Hwang, P.C. van Zijl, S. Mori, J. Biomol. NMR. 11 (1998) 221–226.
[25]O. Walker, R. Varadan, D. Fushman, J. Magn. Reson. San Diego Calif 1997. 168 (2004) 336–345. doi:10.1016/j.jmr.2004.03.019.
[26]N. Rezaei-Ghaleh, F. Klama, F. Munari, M. Zweckstetter, Angew. Chem. Int. Ed Engl. 52 (2013) 11410–11414. doi:10.1002/anie.201305094.
[27]N. Rezaei-Ghaleh, F. Klama, F. Munari, M. Zweckstetter, Bioinforma. Oxf. Engl. 31 (2015) 1319–1321. doi:10.1093/bioinformatics/btu824.
[28]P. Bernadó, E. Mylonas, M.V. Petoukhov, M. Blackledge, D.I. Svergun, J. Am. Chem. Soc. 129 (2007) 5656–5664. doi:10.1021/ja069124n.
[29]S. Vijay-Kumar, C.E. Bugg, W.J. Cook, J. Mol. Biol. 194 (1987) 531–544.
[30]R. Varadan, M. Assfalg, A. Haririnia, S. Raasi, C. Pickart, D. Fushman, J. Biol. Chem. 279 (2004) 7055–7063. doi:10.1074/jbc.M309184200.
[31]R. Bajaj, F. Munari, S. Becker, M. Zweckstetter, J. Biol. Chem. 289 (2014) 34620–34626. doi:10.1074/jbc.M114.595702.
[32]F. Munari, N. Rezaei-Ghaleh, S. Xiang, W. Fischle, M. Zweckstetter, PloS One. 8 (2013) e60887. doi:10.1371/journal.pone.0060887.
[33]M.D. Mukrasch, S. Bibow, J. Korukottu, S. Jeganathan, J. Biernat, C. Griesinger, E. Mandelkow, M. Zweckstetter, PLoS Biol. 7 (2009) e34. doi:10.1371/journal.pbio.1000034.
[34]H.G. Hocking, K. Zangger, T. Madl, Chemphyschem Eur. J. Chem. Phys. Phys. Chem. 14 (2013) 3082–3094. doi:10.1002/cphc.201300219.
[35]P.L. Wintrode, G.I. Makhatadze, P.L. Privalov, Proteins Struct. Funct. Genet. 18 (1994) 246–253. doi:10.1002/prot.340180305.
[36]G.I. Makhatadze, M.M. Lopez, J.M. Richardson, S.T. Thomas, Protein Sci. Publ. Protein Soc. 7 (1998) 689–697. doi:10.1002/pro.5560070318.
[37]D. Morimoto, E. Walinda, H. Fukada, Y.-S. Sou, S. Kageyama, M. Hoshino, T. Fujii, H. Tsuchiya, Y. Saeki, K. Arita, M. Ariyoshi, H. Tochio, K. Iwai, K. Namba, M. Komatsu, K. Tanaka, M. Shirakawa, Nat. Commun. 6 (2015) 6116. doi:10.1038/ncomms7116.
[38]G. Arena, R. Fattorusso, G. Grasso, G.I. Grasso, C. Isernia, G. Malgieri, D. Milardi, E. Rizzarelli, Chem. - Eur. J. 17 (2011) 11596–11603. doi:10.1002/chem.201101364.
[39]C.N. Pace, J.M. Scholtz, IRL Press, Oxford, 2002.
[40]N.P. Dantuma, C. Heinen, D. Hoogstraten, 8 (2009) 449–460. doi:10.1016/j.dnarep.2009.01.005.
[41]A. Ceccon, V. Tugarinov, A. Bax, G.M. Clore, J. Am. Chem. Soc. 138 (2016) 5789–5792. doi:10.1021/jacs.6b02654.
[42]L. Hicke, R. Dunn, Annu. Rev. Cell Dev. Biol. 19 (2003) 141–172. doi:10.1146/annurev.cellbio.19.110701.154617.
[43]A. Khaminets, C. Behl, I. Dikic, Trends Cell Biol. 26 (2016) 6–16. doi:10.1016/j.tcb.2015.08.010.
[44]A. Ceccon, M. D'Onofrio, S. Zanzoni, D.L. Longo, S. Aime, H. Molinari, M. Assfalg, Proteins. 81 (2013) 1776–1791. doi:10.1002/prot.24329.
[45]M. Pedò, F. Löhr, M. D'Onofrio, M. Assfalg, V. Dötsch, H. Molinari, J. Mol. Biol. 394 (2009) 852–863. doi:10.1016/j.jmb.2009.10.014.
[46]M. D'Onofrio, S. Zanzoni, F. Munari, H.L. Monaco, M. Assfalg, S. Capaldi, Biochim. Biophys. Acta BBA - Gen. Subj. 1861 (2017) 2315–2324. doi:10.1016/j.bbagen.2017.07.004.

FIGURES

Fig. 1. *Illustration of the structure of Ubb^{+1}.* Cartoon representation of the five lowest energy NMR structures of Ubb^{+1} (PDB 2KX0, [11]). The globular domain is colored in orange, the C-terminal 19-residue fragment is shown in green. The eight exogenous N-terminal residues are not displayed as they are absent in our Ubb^{+1} construct.

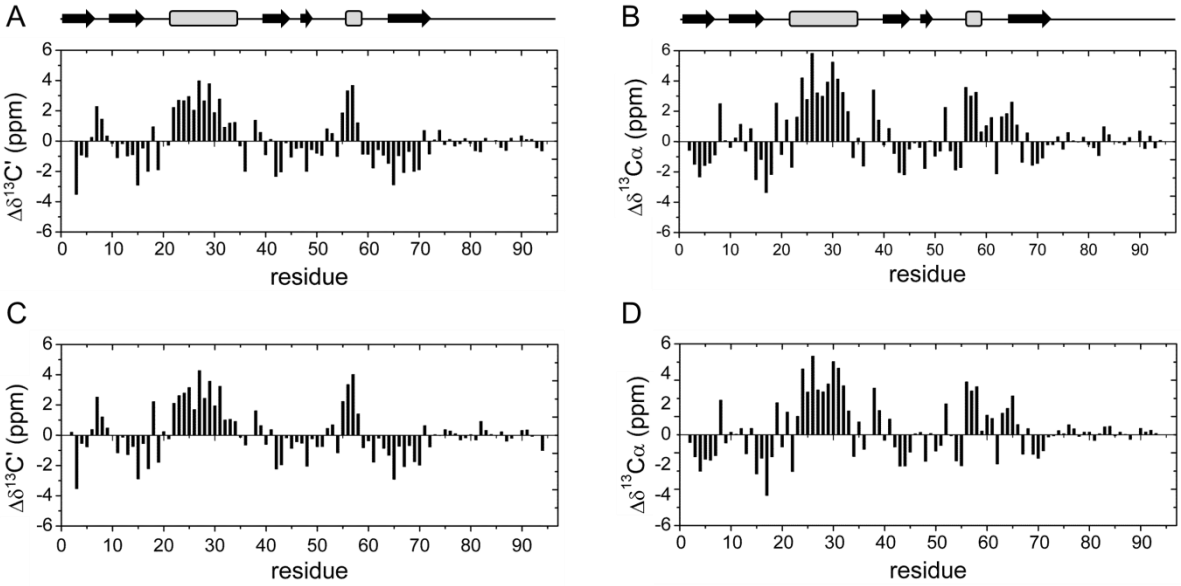


Fig. 2. *Analysis of Ubb^{+1} chemical shifts.* Secondary ^{13}C (A) and ^{13}Cα (B) chemical shifts ($\Delta\delta^{13}$C' and $\Delta\delta^{13}$Cα) of Ubb^{+1}, obtained with Camcoil random coil values, are plotted versus the protein sequence. Secondary ^{13}C' (C) and ^{13}Cα (D) chemical shifts ($\Delta\delta^{13}$C' and $\Delta\delta^{13}$Cα) of

Ubb^{+1}, obtained with the Neighbor Corrected Structural Propensity Calculator random coil values, are plotted versus the protein sequence.

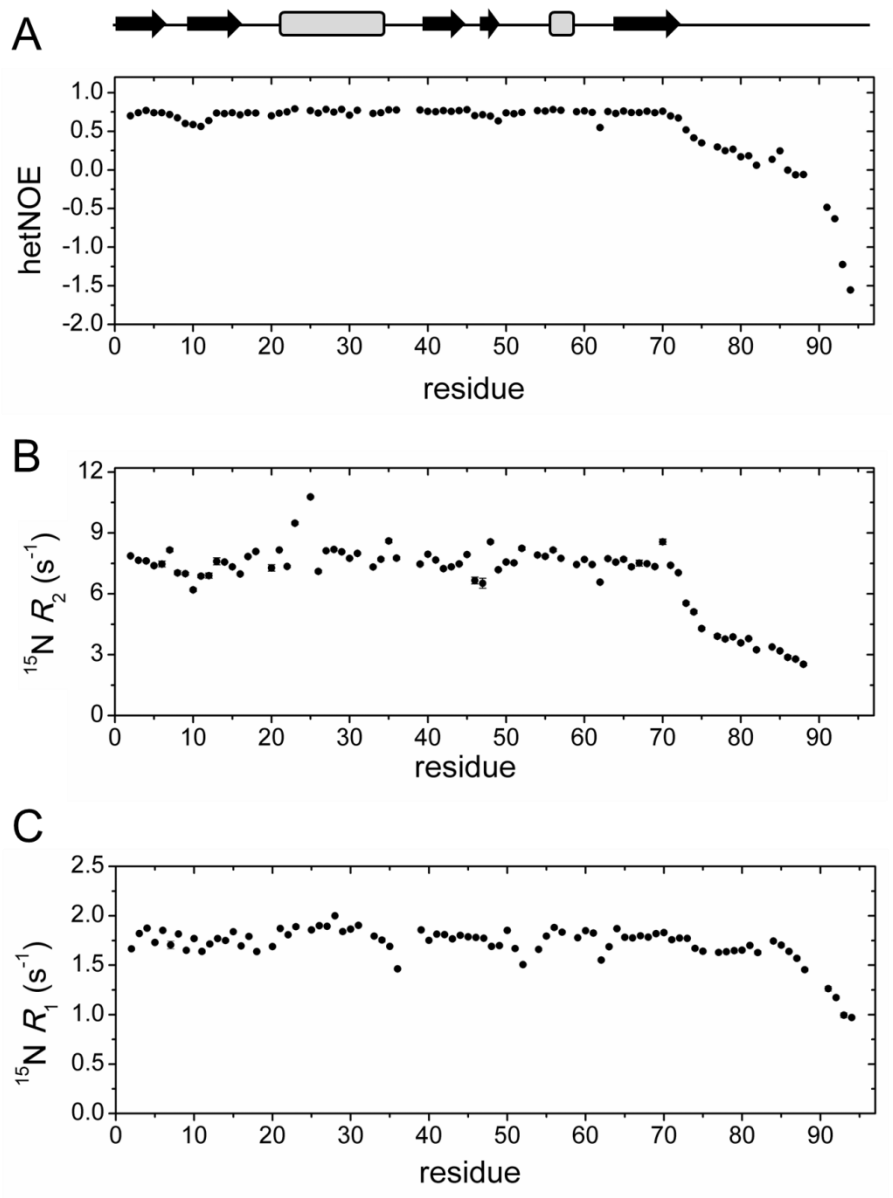


Fig. 3. *Analysis of Ubb^{+1} backbone dynamics*. ^{15}N–spin relaxation rates of [^{15}N]Ubb^{+1} are shown as a function of residue number: A) steady-state {^{1}H}^{15}N heteronuclear NOE (hetNOE); B) R_2 and C) R_1. Residues affected by signal overlap or with insufficient signal-to-noise ratio were excluded from the analysis.

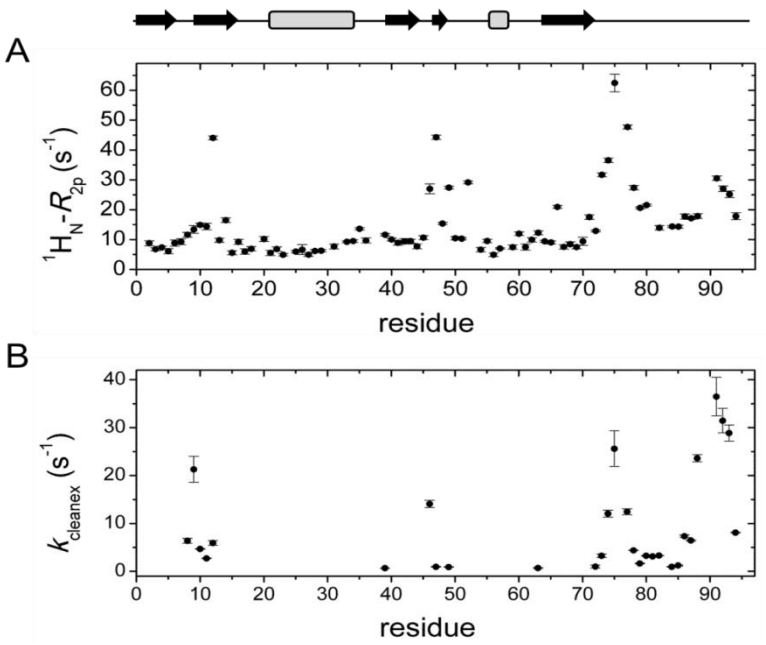


Fig. 4. *Solvent accessibility of Ubb^{+1}*. The solvent exposure of Ubb^{+1} has been probed by solvent PRE (A) or by measurement of water-amide proton exchange rates (B). In A) the $^{1}H_{N}$-R_{2p} rates of Ubb^{+1} obtained with 2 mM gadodiamide are plotted as a function of protein sequence. In B) the water-amide proton exchange rates (k) measured through a CLEANEX-PM-FHSQC experiment are reported versus the protein sequence. Error bars are standard errors obtained from data fitting. Residues affected by signal overlap or with insufficient signal-to-noise ratio were excluded from the analysis.

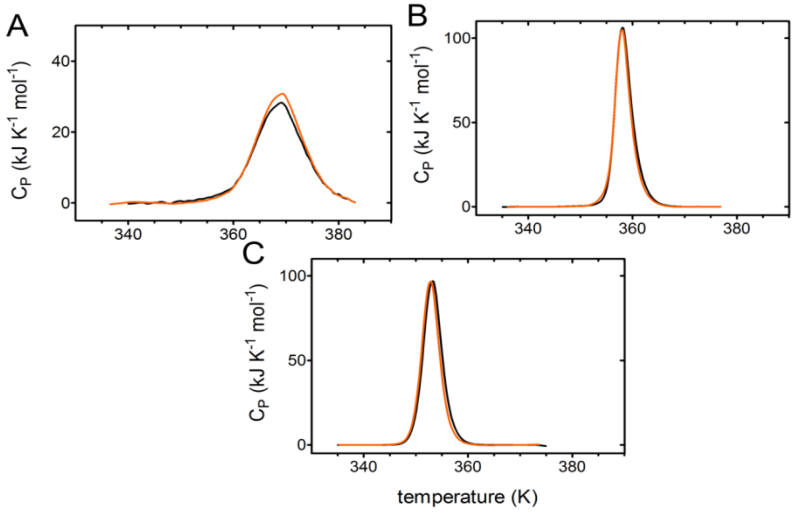


Fig. 5. *Thermal stability measured by calorimetry*. Differential scanning calorimetry thermograms recorded on samples of A) Ub, black, Ubb^{+1}, orange, B) Ub-^{48}Ub, black, Ub-^{48}Ubb^{+1}, orange, C) Ub-^{63}Ub, black, Ub-^{63}Ubb^{+1}, orange. DSC traces are baseline-corrected.

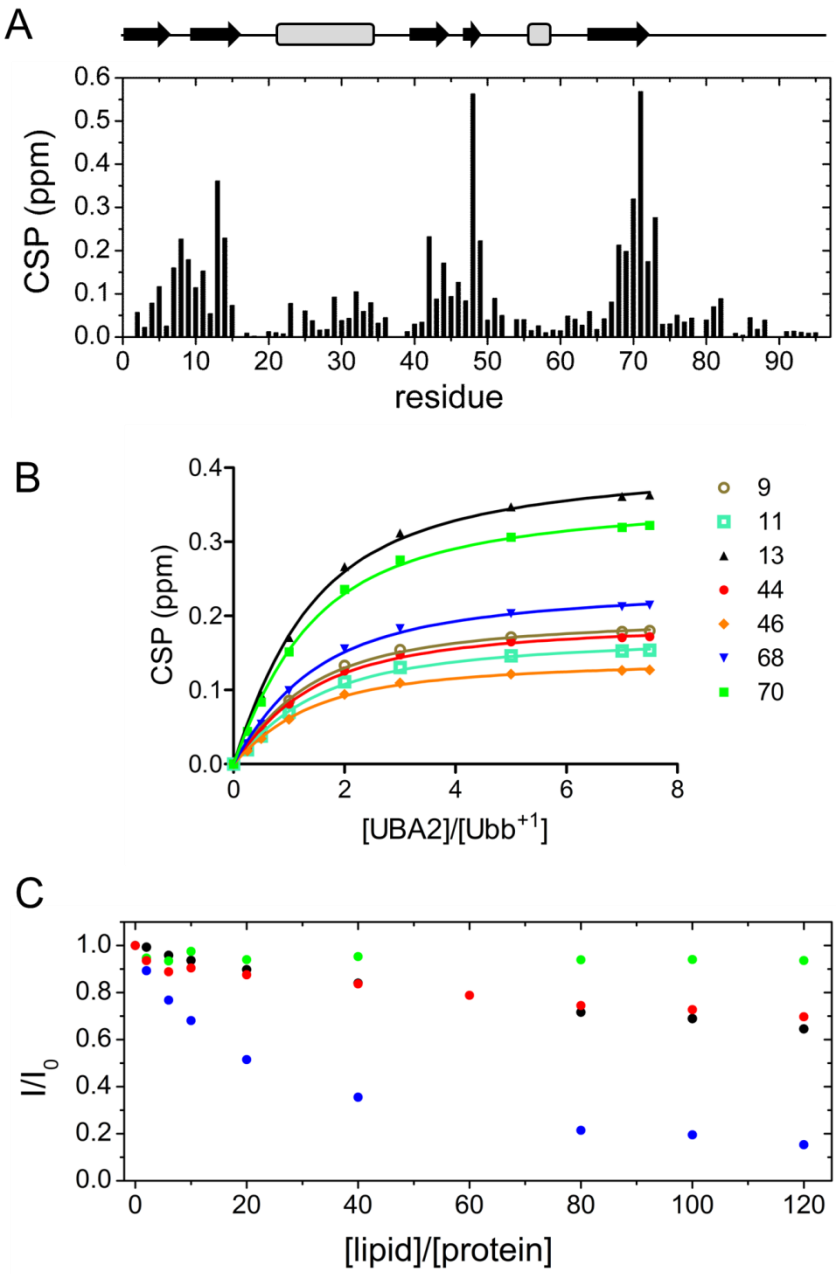


Fig. 6. *Biomolecular interactions of Ubb^{+1}.* A) Plot of the chemical shift perturbations (CSP) determined from ^{1}H,^{15}N-HSQC spectra recorded on [^{15}N]Ubb^{+1} in the presence of (HHR23A)UBA2 at 1:7 molar ratio and in the absence of UBA2. B) Binding isotherms for selected Ubb^{+1} residues based on CSP data from a UBA2/ Ubb^{+1} NMR titration experiment. C) Binding of Ubb^{+1} or Ub to liposomes monitored through analysis of protein ^{1}H signal intensity loss as a function of total lipids to protein molar ratio. Titration data for Ub/POPG-Chl 80:20 (blue), Ubb^{+1}/POPG-Chl 80:20 (red), Ub/POPG-POPC-Chl 40:40:20 (black) and Ubb^{+1}/POPG-POPC-Chl 40:40:20 (green) are shown.

Supplementary info of Publication 3

SUPPORTING INFORMATION

Alzheimer's disease-associated ubiquitin mutant Ubb^{+1}: properties of the carboxy-terminal domain and its influence on biomolecular interactions

Francesca Munari, Andrea Bortot, Michael Assfalg, and Mariapina D'Onofrio*

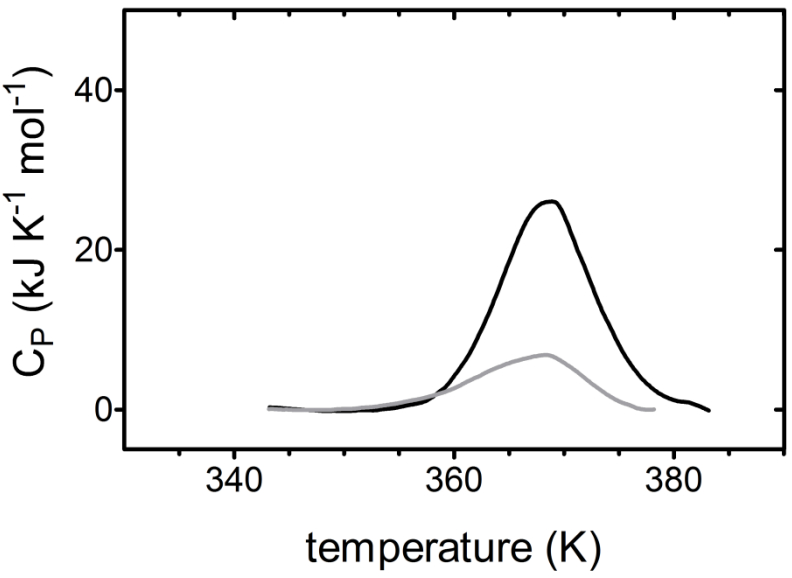


Figure S1A. Differential scanning calorimetry trace of monoubiquitin (Ub). The black trace shows initial heating and the grey line represents reheating. Data were acquired at a scan rate of 1 K min^{-1}. The sample contained 0.07 mM protein dissolved in 100 mM sodium phosphate, pH 7.4, 100 mM NaCl.

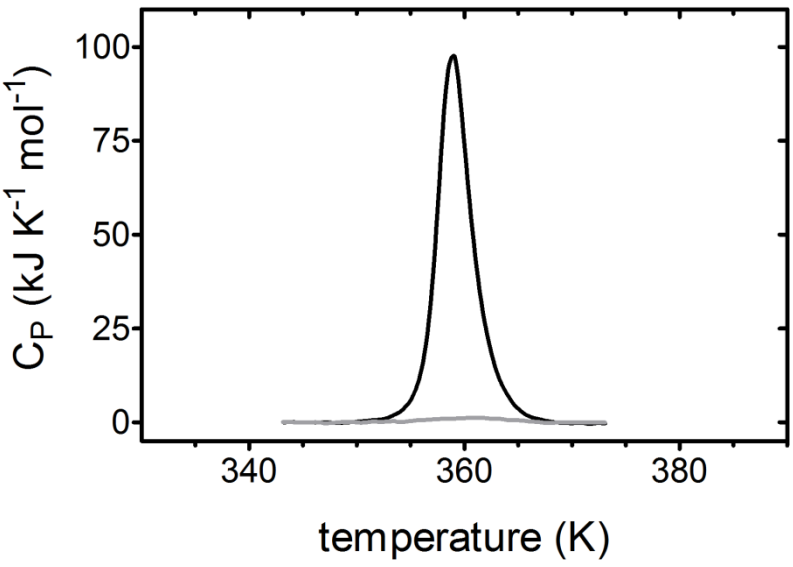


Figure S1B. Differential scanning calorimetry trace of Lys48-linked di-ubiquitin ($^{K48}Ub_2$). The black trace shows initial heating and the grey line represents reheating. Experimental conditions were the same as for Figure S1A.

Table S1. Calorimetric parameters determined from DSC thermograms.

	$\mathbf{T_m}^{(a)}$ / K	$\mathbf{\Delta H_{cal}}^{(b)}$ / kJ mol^{-1}
Ub	369.1	327
Ubb^{+1}	369.3	344
Ub-^{48}Ub	358.1	447
Ub-^{48}Ubb^{+1}	358.0	424
Ub-^{63}Ub	353.2	443
Ub-^{63}Ubb^{+1}	352.9	425

a. Temperature of the peak maximum

b. Determined as the total peak area

CHAPTER 4:
Conclusions

Conclusions

In this doctoral project we aimed at contributing to a better understanding of ubiquitin(Ub)-mediated signaling by investigating the influence of the intracellular environment on the formation of transient Ub-partner complexes. Weak protein-protein interactions, such as those between Ub and Ub binding domains, are potentially prone to be influenced by macromolecular crowding, a distinctive feature of the intracellular milieu. In our work we found that high concentrations of a crowding agent did not influence the preferential binding of the representative Ub-binding domain UBA2 to the canonical Ile44 patch of Ub. On the other hand, from a more comprehensive exploration of Ub/UBA2 contacts based on a solvent PRE approach, secondary contact surfaces were detected. The regions were classified as specific based on the non-linear PRE trend observed at varying UBA2 concentration, although the absence of strong concomitant chemical shift perturbation hinted at ultra-weak affinity of the corresponding interactions and/or a significant heterogeneity of the ensemble of complex conformations. Thus, Ub/UBA2 complexes populate high energy local minima of the free energy landscape which may be in equilibrium with the low energy minimum corresponding to the stereospecific complex. Alternatively, the identified patches may mediate formation of futile complexes that are in competition with the main conformation. It is probably not accidental that other Ub binding domains target these regions, which can thus be considered pivotal for biomolecular recognition. In the broader context, identification of weak secondary interaction surfaces in cell-mimicking crowded solutions by use of PRE methods could improve our understanding of dynamic protein-protein interaction networks and, ultimately, of the molecular-level structural organization of the cellular interior.

In a bid to lay the basis for the development of artificial Ub receptors to be used for biomedical applications, we studied the binding equilibria between Polyubiquitin (polyUb) chains and nanoparticle (NP) surfaces. It is now widely accepted that most nanomaterials display bioactivity as a consequence of their interaction with biomolecules, however the specificity and selectivity of such interactions remain poorly explored. The capability of NPs to distinguish closely related covalent forms of

proteins, such as structural isomers, has important implications for the design of nanomaterials and could be exploited for the development of NP devices eliciting specific biological responses. Distinct covalent forms of proteins, produced by post-translational modification, display different recognition properties and perform different functions. This notion is well exemplified by the Ub system, in which differently conjugated polymeric chains have unique binding features and play distinct roles in cellular biology. In our work, we found that structural isomers of polyUb interacted with negatively charged NP surfaces through preferential binding sites. Interestingly, the adsorption of structural isomers of diubiquitin (Ub2) to large NPs involved different surface patches, resulting in reduced binding affinity of K48Ub2 compared to K63Ub2. We further showed that recognition of K48Ub2 required the dynamic interconversion between the prevalent 'closed' conformation and the binding competent 'open' state. Thus, we have extended the scope of solution-state NMR spectroscopy approaches to the detailed description of complex protein-NP adsorption equilibria. We provided evidence of the role of structural isomerism and conformational isomerization in the regulation of adsorption, contributing to a better understanding of the principles governing protein recognition by NP surfaces. The here described approach could be applied to more specialized NP systems, exploiting the possibility to tailor functional groups at the NP surface to increase biomolecular binding specificity and selectivity. In the particular case of Ub, appropriately designed NPs could act as artificial receptors for certain isomeric forms, thereby providing new strategies for the treatment of human diseases such as cancer and neurodegeneration.

Neurodegeneration in particular has been found to involve dysfunction of the Ub system at various levels. For example, Ubb^{+1}, a ubiquitin mutant protein originating from misreading of the Ub B gene, is found accumulated in brain tissues of Alzheimer's disease patients. The mutant attracts strong interest due to its possible participation in the molecular events leading to neurodegeneration. Ubb^{+1} is composed of the globular domain of Ub, linked to a 19-residue C-terminal peptide. Based on NMR relaxation and solvent accessibility measurements we obtained new insight into the molecular properties of Ubb^{+1}. We further determined the thermal stability of Ubb^{+1} in the monomeric form, and in Lys48- and Lys63-linked dimers. Finally, we explored the

influence of the C-terminal fragment on the interactions of Ubb^{+1} with an isolated UBA2 domain and with membrane mimics. Our findings support the notion that the 19-residue extension of Ubb^{+1} lacks a well ordered structure and persistent canonical secondary structure elements. However, we observed that the flexible fragment displayed non-uniform structural and dynamic properties along its amino acid sequence. From analysis of global dynamics based on ^{15}N-spin relaxation data, it emerged that the mobility of the C-terminal tail was independent from that of the globular domain, and persistent interdomain interactions were absent. Hence, it is possible that a reduction in local flexibility might originate from structural constraints imposed by the sequence of amino acids, and in particular by the proximity of two proline residues (Pro83 and Pro90). Consistently with the above picture, the peptide extension did not affect the protein thermal stability in either monomeric or dimeric species. However, biomolecular recognition was in part perturbed: the interaction of Ubb^{+1} with (HHR23A) UBA2 was characterized by enhanced affinity compared to that of Ub, and the weak association of Ub with anionic lipid surfaces was significantly reduced in the case of the variant. It can be concluded that the globular architecture of Ub is sufficiently robust not to be particularly influenced by the unintended covalent conjugation of a peptide to its flexible C-terminus. However, the selected interactions investigated in our work exemplify possible changes in biomolecular recognition and diffusive dynamics experienced by disease-associated mutant Ub species, compared to wild-type molecules, in physio(patho)logical conditions.