

# Molecular recognition and dynamics of linear poly-ubiquitins: integrating coarse-grain simulations and experiments

Alexander Jussupow<sup>1</sup>, Ana C. Messias<sup>2,3</sup>, Ralf Stehle<sup>2,3</sup>, Arie Geerlof<sup>2,3</sup>, Sara M. Ø. Solbak<sup>4</sup>, Anders Bach<sup>4</sup>, Michael Sattler<sup>2,3,\*</sup>, and Carlo Camilloni<sup>1,5,\*</sup>

<sup>1</sup>*Department of Chemistry and Institute for Advanced Study, Technical University of Munich, Garching 85747, Germany*

<sup>2</sup>*Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg 85764, Germany.*

<sup>3</sup>*Center for Integrated Protein Science Munich at Department of Chemistry, Technical University of Munich, Garching 85747, Germany*

<sup>4</sup>*Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark.*

<sup>5</sup>*Dipartimento di Bioscienze, Università degli studi di Milano, 20133 Milano, Italy*

\*Corresponding authors: [sattler@helmholtz-muenchen.de](mailto:sattler@helmholtz-muenchen.de), [carlo.camilloni@unimi.it](mailto:carlo.camilloni@unimi.it)

**Keywords:** ubiquitin, nemo, molecular dynamics, small angle X-ray scattering, metainference, polymer theory

## Abstract

Poly-ubiquitin chains are flexible multidomain proteins, whose conformational dynamics enable their molecular recognition by a large number of partners in multiple biological pathways. By using alternative linkage, it is possible to obtain poly-ubiquitin molecules with different dynamical properties. This flexibility is further increased by the possibility to tune the length of poly-ubiquitin chains. Characterizing the dynamics of poly-ubiquitins as a function of their length is thus relevant to understand their biology. Structural characterization of poly-ubiquitin conformational dynamics is challenging both experimentally and computationally due to increasing system size and conformational variability. Here, by developing highly efficient and accurate small-angle X-ray scattering driven Martini coarse-grain simulations, we characterize the dynamics of linear M1-linked di-, tri- and tetra-ubiquitin chains. Our data show that the behavior of the di-ubiquitin subunits is independent of the presence of additional ubiquitin modules. We propose that the conformational space sampled by linear poly-ubiquitins, in general, may follow a simple self-avoiding polymer model. These results, combined with experimental data from small angle X-ray scattering, biophysical techniques and additional simulations show that binding of NEMO, a central regulator in the NF- $\kappa$ B pathway, to linear poly-ubiquitin obeys a 2:1 (NEMO:poly-ubiquitin) stoichiometry in solution, even in the context of four ubiquitin units. Eventually, we show how the conformational properties of long poly-ubiquitins may modulate the binding with their partners in a length-dependent manner.

## Significance

Protein conformational dynamics plays an essential role in molecular recognition mechanisms. The characterization of conformational dynamics is hampered by the conformational averaging of observable in experimental structural biology techniques and by the limitations in the accuracy of computational methods. By developing an efficient and accurate approach to combine small-angle X-ray scattering solution experiments and coarse-grain Martini simulations, we show that the conformational dynamics of linear poly-ubiquitins can be efficiently determined and to rationalize the role of poly-ubiquitin dynamic in the molecular recognition of the UBA1 domain upon binding to the signaling regulator NEMO. The analysis of the conformational ensembles allows us to propose a general model of the dynamics of linear poly-ubiquitin chains where they can be described as a self-avoiding polymer with a characteristic length associated with their specific linkage.

## Introduction

Ubiquitination is a reversible post-transcriptional modification system that regulates key physiological processes, such as protein degradation, cell cycle, apoptosis, DNA repair, and signal transduction<sup>1-3</sup>. Once a protein substrate is mono-ubiquitinated (e.g. a lysine of the substrate is conjugated through an iso-peptide bond to the C-terminus of a ubiquitin monomer), an additional ubiquitin may be conjugated to either one of the seven lysine residues of the first ubiquitin (K6, K11, K27, K29, K33, K48, and K63)<sup>4</sup> or its amino-terminal methionine residue (M1)<sup>5-7</sup>. This process can lead to the assembly of poly-ubiquitin chains of various lengths and topologies. The resulting polymeric chains are then associated with different cellular mechanisms<sup>8</sup>. Since all these polymers are made of the same single unit, the highly conserved 76-residues long ubiquitin domain, the ubiquitin code is an example of a conformation-based alphabet, where both the polymerization site<sup>8-9</sup> as well as the chain length<sup>10</sup> regulate the recognition by different partners, and thereby determine the cellular fate of the protein. The role of poly-ubiquitin length and dynamics in molecular recognition processes is poorly understood<sup>8, 10-11</sup>. An overall assessment of the typical length of different poly-ubiquitin chains in physiological conditions is missing, and only sporadic indications are available. For example, in the case of K48-linked poly-ubiquitin, a length of four is generally considered optimal for molecular recognition of the 26S proteasome<sup>12</sup>, while the nuclear protein localization protein 4 (Npl4) is selective for K48-linked chains longer than six<sup>13</sup>. Interestingly, it was reported that K48-linked tetra-ubiquitin (Ub<sub>4</sub>) slows down further ubiquitination<sup>14-16</sup>, while this is not the case for K63-linked Ub<sub>4</sub><sup>16</sup>.

Linear M1-linked poly-ubiquitin chains (**Figure 1**), whose assembly is catalyzed by LUBAC<sup>5</sup>, have been shown to play a role in inflammation, immune responses, and oncogenesis<sup>17-19</sup>. Their most studied function is the involvement in the activation of the canonical NF- $\kappa$ B pathway<sup>6-7, 17, 20-23</sup>. In this pathway, the IKK complex (or I $\kappa$ B kinase, formed by IKK $\alpha$ , IKK $\beta$ , and NEMO, also known as IKK $\gamma$ , the NF- $\kappa$ B essential modulator) is activated by LUBAC upon activation by various stimuli<sup>22</sup>. LUBAC preferentially recognizes and conjugates linear ubiquitin chains on NEMO. NEMO also possesses a specific linear di-ubiquitin-binding region referred to as the “ubiquitin binding in ABIN and NEMO” (UBAN) motif<sup>24</sup>, which forms a helical coiled-coil dimer in solution<sup>23</sup>. Recognition of a linear poly-ubiquitin conjugated to NEMO by the UBAN domain of another NEMO may trigger the clustering of the IKK complex as well as conformational changes that subsequently activate IKK<sup>25-26</sup>. Once active IKK can phosphorylate and inactivate the I $\kappa$ Bs (inhibitor of the NF- $\kappa$ B proteins) leading to the release of NF- $\kappa$ B<sup>27</sup>. Indeed, it was recently shown that it is possible to inhibit NF- $\kappa$ B activation upon UBAN-dependent TNF $\alpha$  and TCR/CD28 stimulation by small-molecules that inhibit the binding of linear poly-ubiquitins to the NEMO<sub>UBAN</sub> domain<sup>23</sup>. While the NEMO<sub>UBAN</sub> domain can bind linear di-ubiquitin, it has been observed that full-length NEMO can only bind Ub<sub>4</sub> or longer suggesting a length-dependent activation mechanism<sup>21</sup>. Furthermore,

another study suggested that the binding of NEMO to chains of 10 linear ubiquitins or longer induces a different conformation of NEMO compared to the binding of shorter chains<sup>20</sup>.

Characterizing the conformational space of poly-ubiquitin chains as a function of length is critical to understand their physiological behavior. Such structural characterization is nonetheless very challenging. Poly-ubiquitins, from di-ubiquitin to longer chains, exhibit a very dynamic behavior<sup>28</sup> that requires determining a statistical ensemble of all the relevant configurations populated in solution. The combination of molecular dynamics (MD) with experimental small-angle X-ray scattering (SAXS) data is very well suitable to study dynamic protein systems<sup>29</sup> including poly-ubiquitin of varying chain size. SAXS does not provide high-resolution structural information. Conversely, MD simulations may be used to determine the statistical ensemble of configuration populated by a system in equilibrium condition, but a full modeling base on MD simulations is hampered by the size of the system<sup>30-31</sup>. This problem can, in principle, be alleviated by coarse-grain force-fields<sup>32</sup>, eventually combined with enhanced sampling techniques<sup>33</sup>, that can massively speed up MD simulations although potentially at the expense of the accuracy<sup>32</sup>.

Here, we show that by integrating SAXS and MD simulations based on the Martini coarse-grain force-field<sup>34-35</sup> by means of Metainference<sup>36</sup> we can efficiently generate an ensemble of structures representing the dynamics of linear poly-ubiquitins (**Figure 1**). The ensembles allow the description of the dynamics of such complex systems at the single residue level. Our results show how poly-ubiquitins can populate multiple conformational states, but unexpectedly indicate that linear poly-ubiquitin chains and potentially poly-ubiquitins, in general, can be described by a simple self-avoiding polymer model. Various biophysical experiments are used to characterize the stoichiometry, kinetics and thermodynamic properties of the binding of poly-ubiquitin to the NEMO<sub>UBAN</sub> domain. Surprisingly, our data demonstrate that NEMO<sub>UBAN</sub> binds to di-, tri- and tetra-ubiquitin (Ub<sub>2</sub>, Ub<sub>3</sub>, and Ub<sub>4</sub>) in all cases forming a 2:1 NEMO<sub>UBAN</sub>:Ub<sub>N</sub> complex in solution. Notably, a conformational ensemble for the NEMO<sub>UBAN</sub>:Ub<sub>2</sub> complex rationalizes the 2:1 binding. Combined with our proposed poly-ubiquitin polymer model, this suggests how longer poly-ubiquitin chains may modulate NEMO recognition as well as bind more than one NEMO dimer.

## Results and Discussion

### *A simple Martini modification improves the simulation of linear di-ubiquitin*

We first evaluated the ability of the Martini coarse-grain force field to describe the dynamics of a linear Ub<sub>2</sub>. A metadynamics<sup>37</sup> simulation of Martini Ub<sub>2</sub> resulted in an extremely compact ensemble of structures (**Figure 2a**) which does not reproduce the measured SAXS intensities (**Figure 2d** and **Fig. S1** in the Supporting Information). In **Figure 2a** we report a free energy landscape (in kJ/mol) as a function of the distance between the centers of the two ubiquitin

domains and their relative orientation; the average distance between the two domains is very short, around  $2.41 \pm 0.02$  nm, with a preferential orientation of the two ubiquitin's domain (measured as the torsion angle between two axes defined using the first and second half of the sequence of each ubiquitin, cf. Methods). The average radius of gyration of  $1.73 \pm 0.01$  nm, strongly underestimates the value of  $2.23 \pm 0.02$  nm derived from SAXS (**Table S1**). The ensemble seems to be able only to capture compact Ub<sub>2</sub> configurations also when compared to the available crystal structures (PDB 2W9N<sup>38</sup> (open), 3AXC<sup>39</sup> (compact) and 4ZQS<sup>28</sup> (compact)). This result can indicate an imbalance between the protein-protein and the protein-solvent interaction in Martini, a result that is not unexpected<sup>40-41</sup> given the extremely simple description employed for the solvent (more complex descriptions, like the Martini polarizable water<sup>42</sup>, are available at the expense of performances). Interestingly, recent developments in atomistic force fields demonstrated the need for tuning solute-solvent interactions<sup>43-44</sup>. Following recent approaches that have successfully improved atomistic and coarse-grained force-fields, we repeated the same simulation after increasing by 5% the Martini water-protein Lennard-Jones interaction. This simple adjustment was sufficient to obtain a more expanded ensemble of structures as shown by the free energy landscape (**Figure 2b**) without any additional computational cost (**Table S2**). Importantly, the new ensemble resulted in an improved, even if not yet quantitative, agreement with the SAXS data (**Figure 2d**, blue curve; **Fig. S1**). The average distance between the domains increased to  $3.10 \pm 0.02$  nm, and the protein can explore a much wider conformational space that now includes open and closed structures. In terms of the radius of gyration, the ensemble average resulted in  $2.05 \pm 0.01$  nm to be compared with the  $2.23 \pm 0.02$  nm derived from SAXS. Of notice, an under-development version of the Martini force field (Martini 3, currently in beta phase) while showing promising behavior may still benefit from increased protein-water interaction (**Fig. S2**). Importantly, Martini simulations outperform full atomistic simulations for the same system. As reported in **Table S1** an explicit solvent atomistic Ub<sub>2</sub> simulation is around 500 times slower than a Martini simulation. Nonetheless, our aim here is to obtain ensembles in quantitative agreement with the SAXS data without a large-scale force field re-parameterization effort. This can be achieved at least, in principle, by integrating experimental information directly in the simulation by Metainference<sup>36</sup>.

### ***Metainference SAXS simulations of Martini di-ubiquitin quantitatively reproduce the experimental data***

Metadynamic metainference (M&M)<sup>45</sup> simulation (see Methods) for Martini linear Ub<sub>2</sub> (including our modified water) result in an ensemble of configurations characterized by a flatter and broader free energy landscape (**Figure 2c**) and in quantitative agreement with the experimental SAXS (**Figure 2d,e**; **Fig. S1**). With respect to the unrestrained simulation, the average distance between the two domains increased from  $3.10 \pm 0.02$  nm to  $3.32 \pm 0.02$  nm. The radius of gyration of the ensemble of  $2.23 \pm 0.01$  nm quantitatively agrees with that derived from SAXS of  $2.23 \pm 0.02$  nm. Qualitatively the topology of the free energy landscape is comparable to the unrestrained simulation but translated to larger relative

distances. Overall the free energy landscape is quite flat with relatively limited free energy differences indicating that the two ubiquitin-domain are relatively free to move with respect to each other. Therefore, Ub<sub>2</sub> shows highly dynamical behavior, which cannot be described by a few individual structures. Instead, a full ensemble is required in agreement with previous findings on linear as well as other di-ubiquitins.

From the performance point of view, the SAXS on-the-fly calculation used by Metainference is computationally demanding, but the use of a coarse-grained representation makes it far more affordable with respect to the same simulation performed at full atomistic resolution (**Table S2**). The loss of performance resulting from the use of SAXS is justified by the increased accuracy of the resulting simulations. Note, that it is not required to calculate the Metainference SAXS restraint at every step of the simulation. Indeed, by calculating it every 5 steps, we obtained a quantitatively equivalent ensemble (**Fig. S2**) at a fraction of the computational cost (**Table S2**). Notably, using Metainference allows us also to sample the scaling value, which is necessary to compare the experimental and computed SAXS curves. For Ub<sub>2</sub> we observed a 3% higher scaling value for the simulation with increased protein water interaction and a 9% higher scaling value just with the Martini force field compared to the Metainference solution (**Figure S1, Table S1**).

***Linear poly-ubiquitin chains are preferentially extended, do not show long-range correlations and can be described as self-avoiding polymers***

To investigate the dynamics of linear Ub<sub>3</sub> and Ub<sub>4</sub>, we performed SAXS experiments on both proteins at different concentrations (**Fig. S3**). The measured SAXS data were then employed to perform M&M simulations (cf. Methods, **Table S3**). Additionally, unrestrained simulations based only on Martini with our modified water were also performed. In **Figure 3** (see also **Fig. S1**), we show the comparison of the back-calculated SAXS with respect to the experimental measures of Ub<sub>3</sub> and Ub<sub>4</sub>. The effect of our improved water diminishes for the longer poly-ubiquitin chains. A comparison of the radius of gyration for the Ub<sub>2</sub>, Ub<sub>3</sub>, and Ub<sub>4</sub> ensembles show that while the unrestrained and restrained simulations sample a comparable range of compactness, the restrained simulations are shifted towards a more extended conformational space. The trend of the average radius of gyration (2.0, 2.7, and 3.3 nm for Ub<sub>2</sub>, Ub<sub>3</sub>, and Ub<sub>4</sub>, respectively, cf. Methods, **Table S2**) suggests an almost linear increase of the size of the protein with the number of ubiquitin monomers. The analysis of the free-energy landscape for the Ub<sub>2</sub> couples in Ub<sub>3</sub> and Ub<sub>4</sub> (**Fig. S4**) shows qualitatively the same behavior, suggesting that the interdomain interactions are essentially only those between neighbor domains (i.e. between Ub<sub>2</sub>). For Ub<sub>3</sub> and Ub<sub>4</sub> this is confirmed by analyzing the free energy landscape of non-neighbor ubiquitin domains. Overall the free energy landscape is flatter for larger poly-ubiquitin chains indicating that the interaction with neighboring ubiquitin becomes less and less specific. Also, the distance (centers of the two ubiquitin domains) distribution shifts from a bimodal distribution for Ub<sub>2</sub> to a flatter one for Ub<sub>3</sub> and Ub<sub>4</sub>. We also observe that Ub<sub>3</sub> samples more

extended conformations for large distances  $> 4.0$  nm, while both Ub<sub>4</sub> and Ub<sub>3</sub> are forming more compact conformation below 2.6 nm. In **Fig. S5a-c** the free energy profiles are shown for the first-and-third ubiquitin Ub<sub>3</sub>(1-3) in Ub<sub>3</sub> as well as for the first-and-third Ub<sub>4</sub>(1-3), second-and-fourth Ub<sub>4</sub>(2-4). These landscapes are all qualitatively similar showing that the interaction between two non-neighbor ubiquitins is quite rare. The average distance between a 1-3 or 2-4 ubiquitins pair is around 6 nm, with an average angle of around 140°. The first ubiquitin does not influence the relative orientation of the third ubiquitin. The first and fourth Ub<sub>4</sub>(1-4) ubiquitin couple, as shown in **Fig. S5d**, behaves similarly. Interaction between the first and fourth ubiquitin are also rare. In most cases, the distance between both ubiquitins is around 8.5 nm. There is also no strong preference for a specific torsion angle between all four ubiquitins.

To further assess the presence of short and long-range interactions between neighbor and non-neighbor ubiquitin couples we estimated the fraction of compact configurations by analyzing the minimum distance between neighbor and non-neighbor ubiquitin couples (**Figure 4a,b**). For neighbor and non-neighbor couples, there is a peak in the distribution around 0.5 nm. As already indicated by the free energy profiles, compact neighbor ubiquitin pairs represent around 40% to 50% of the ensemble, while contacts between non-neighbor couples are only present in around 8% for Ub<sub>3</sub> and around 2% for Ub<sub>4</sub>, indicating an overall lack of compact states in linear poly-ubiquitins. A contact analysis for the Ub<sub>2</sub> compact state indicates that this state is not structurally homogeneous. Even the most frequent contact is only present in 10 to 30% of all compact conformations, depending on the specific ubiquitin pair (**Figure 4c**). On the other hand, even the 10<sup>th</sup> most frequent contact has still a probability between 5 to 15% while the 100<sup>th</sup> most frequent one is still in the 1 to 5% range.

Nonetheless, all residues involved in the most frequent contacts belong to three distinct surfaces. These interactions define the preferred orientations between two adjacent ubiquitin pairs. All residues involved in the most frequent contacts from the first ubiquitin are on the same surface as the hydrophobic patch I44 (**Figure 4d, Fig. S6**), which is known to be also essential for interactions of ubiquitin with other proteins. The I44 surface interacts either with the surface around E18<sub>2</sub> or I13<sub>2</sub> (E18 or I13 would be the analog residues of the first ubiquitin). The E18<sub>2</sub> surface is located opposite to the I44 surface while the I13<sub>2</sub> surface is roughly 90° rotated to the I44 and E18<sub>2</sub> surface. In **Fig. S6**, the ten most frequent contacts between ubiquitin cores for all adjacent ubiquitin pairs are illustrated. Ub<sub>2</sub> is predominantly stabilized by salt bridges between the positive charged R42 and K48, and the negative charged E16<sub>2</sub> and E18<sub>2</sub>. However, going to Ub<sub>3</sub> and Ub<sub>4</sub>, electrostatic interactions become less important compared to Van der Waals (VdW) interactions (**Table S4**). For Ub<sub>2</sub>, the Coulomb interaction between charged amino acids is responsible for 27 % of the total interaction energy between two ubiquitin cores. This value goes as low as 9 % for the Ub(3-4) pair of Ub<sub>4</sub>. The lower electrostatic interactions may also explain the flatter free energy surfaces of Ub<sub>3</sub> and Ub<sub>4</sub> neighbor pairs (**Fig. S4**). On the other hand, the increased role of VdW interactions in Ub<sub>3</sub> and Ub<sub>4</sub> is compatible with the increase in the compact population of di-ubiquitin couples.

Finally, while in Ub<sub>2</sub> the I44 surface prefers to interact with the E18<sub>2</sub> surface, interactions between the I44 and the I13<sub>2</sub> surface are more important for the last pair of Ub<sub>4</sub>, causing a shift of the preferred orientation between both ubiquitins.

Overall, our linear Ub<sub>2</sub>, Ub<sub>3</sub>, and Ub<sub>4</sub> ensembles indicate that linear poly-ubiquitins are extended polymers, whose dynamics are mostly uncorrelated over a distance of more than one ubiquitin domain. In **Figure 4e** this behavior is further highlighted by plotting the end-to-end distance as a function of the number  $N$  of ubiquitin domain,  $e2e(N)$ . Remarkably, fitting the data with a power law (including the end-to-end distance for Ub<sub>1</sub>) resulted in  $e2e(N) = 3.9N^{0.60}$  in remarkable agreement with Flory theory for self-avoiding polymers<sup>46</sup>. This is remarkable, since generally, proteins do not behave as self-avoiding chain, showing also less entropy in the denatured state. It is tempting to speculate that all poly-ubiquitins may be described as self-avoiding polymers following the same relationship for the end-to-end distance but with a different pre-factor (i.e., characteristic length) associated with the distance between the C-terminus glycine and the specific linkage side chain. For K63-linked poly-ubiquitins, we could test this, even if only with the mono-ubiquitin and di-ubiquitin, making use of a SAXS based ensemble we recently published<sup>47</sup>. By fixing the exponent to 0.6 and setting the prefactor to the average distance between the C-terminus and K63, the expected  $e2e$  distance of Ub<sub>2</sub> fits the ensemble. This suggests that the fit can be used to predict the behavior for longer K63-linked Ub <sub>$n$</sub>  chains (**Fig. S7**). Extrapolating for other linkages we observe that K11- and K48-linked poly-ubiquitin, which are known to populate more compact states in solution<sup>48-49</sup>, have a shorter distance between the C-terminus and the lysine and would consequently show smaller pre-factors and populate systematically more compact states than K63- and M1-linked.

### ***The conformational entropy of long linear poly-ubiquitins modulates NEMO binding.***

In order to study the dynamics and changes of linear poly-ubiquitin dynamics upon binding to cognate proteins, we used simulations and experiments to characterize the interaction of the NEMO UBAN domain (NEMO<sub>258-350</sub>) to the linear poly-ubiquitins Ub<sub>2</sub>, Ub<sub>3</sub> and Ub<sub>4</sub>. The NEMO UBAN domain is a dimer in solution<sup>23</sup>. Previous studies have shown two different binding stoichiometries in solution and crystalline state for Ub<sub>2</sub>: with either two NEMO monomers bound to one Ub<sub>2</sub>(2:1) or two NEMO monomers bound to two Ub<sub>2</sub> (2:2)<sup>23-24</sup>. Poly-ubiquitin chains longer than Ub<sub>2</sub> harbor potentially more than one binding site and could thus bind more than one NEMO dimer with a theoretical stoichiometry for Ub <sub>$N$</sub>  of 2( $N$ -1):1 (NEMO:Ub <sub>$N$</sub> ). A crystal structure (PDB 5H07) shows the binding of two linear Ub<sub>3</sub> to four ABIN monomers (a homologous of NEMO which also form a dimer in solution)<sup>50</sup>. The observed binding mode requires Ub<sub>3</sub> to be in a relatively compact and univocally oriented configuration in order to avoid steric hindrances between the two ABIN dimers. This would dramatically decrease the entropy not only of each di-ubiquitin couple but also that of the overall chain and thus should be entropically disfavored in solution (**Fig. S5**). Indeed, isothermal titration calorimetry (ITC)

experiments show that in solution only one ABIN2 dimer binds to Ub<sub>3</sub><sup>50</sup>, arguing that higher stoichiometries are artifactual, induced by crystal packing and do not reflect the solution assembly.

To characterize the binding in solution of NEMO to Ub<sub>3</sub> and Ub<sub>4</sub> we performed SAXS (**Figure 5a,b, Fig. S3, Table S5**), ITC (**Figure 5c,d, Fig. S8, Table 1, Table S6**), Size Exclusion Chromatography (SEC) coupled with static light scattering (SLS) (**Fig. S9, Table 2, Table S7**), and surface plasmon resonance (SPR) (**Fig. S10, Table S8**). Extending Ub<sub>2</sub> to Ub<sub>3</sub> and Ub<sub>4</sub> does not affect the (2:1) binding stoichiometry, with either poly-ubiquitin protein binding two NEMO monomers (one NEMO dimer) (**Figure 5, Table 2**). In SEC experiments with SLS (**Table S7**) we detected NEMO:Ub<sub>3</sub> complexes with molecular weights ranging from (41-45 kDa), which is similar to the MW of the calculated 2:1 NEMO:Ub<sub>3</sub> complex (47.7 kDa), while for NEMO:Ub<sub>4</sub> complexes only a single peak is found with MW between 53-56 kDa (calculated MW 2:1 NEMO:Ub<sub>4</sub> 56.2 kDa) (**Table 2, Fig. S9, Table S5**). SAXS measurements (**Figure 5a-b, Table 2, Fig.S3**) confirmed the stoichiometry observed by SLS-SEC. ITC (**Figure 5c,d, Fig. S8, Table 1, Table S6**) indicates that NEMO binds to Ub<sub>3</sub> and Ub<sub>4</sub> with very similar enthalpy ( $\Delta H$  of -17.9 and -18.8 kJ/mol, respectively) suggesting that the molecular interactions and binding interfaces between NEMO and the different poly-ubiquitins are similar to the one described for NEMO:Ub<sub>2</sub><sup>23</sup> ( $\Delta H$  of -16.9 kJ/mol). Affinities for Ub<sub>3</sub> and Ub<sub>4</sub> are 1.6 and 4.1  $\mu$ M close to 1.8  $\mu$ M obtained for Ub<sub>2</sub>. ITC data comparing shorter and longer poly-ubiquitins may suggest that longer poly-ubiquitins can form long range, flanking interactions with NEMO, resulting in a gain of enthalpy and loss of entropy with respect to shorter ones. SPR confirmed the binding between NEMO and Ub<sub>3</sub> and Ub<sub>4</sub> with equilibrium dissociation constants ( $K_D$ ) of 9.6  $\mu$ M and 6.4  $\mu$ M for Ub<sub>3</sub> and Ub<sub>4</sub>, respectively (**Fig. S10, Table S8**).

To explain the contradictory observation of the stoichiometry in solution and in crystal and to better understand the molecular recognition between linear poly-ubiquitins and NEMO, we characterized the dynamics of a NEMO<sub>258-350</sub>:Ub<sub>2</sub> complex. An M&M Martini simulation was performed including SAXS data for the complex previously measured (**Figure 6a, Fig. S11**). The resulting ensemble of structures highlights how binding to NEMO strongly decreases the conformational freedom of linear Ub<sub>2</sub> (**Figure 6a, Figure 2**). Neither the ensemble nor the crystal structures of other bound Ub<sub>2</sub> are located close to the minima of the free Ub<sub>2</sub> ensemble, which has a different distance and orientation between both ubiquitins. Interestingly, Ub<sub>2</sub> residues building up the NEMO<sub>258-350</sub>:Ub<sub>2</sub> interface overlap with those involved in the interdomain interactions (**Fig. S11**), in particular residues around the hydrophobic patch I44 and the previously mentioned E18<sub>2</sub> surface. The observed interaction sites are in agreement with observed NMR chemical shifts perturbations reported in Vincendeau et al<sup>23</sup>. The ensemble also provides a possible explanation for the different binding stoichiometry observed in solution (NEMO:Ub<sub>2</sub> 2:1) and crystalline state (NEMO:Ub<sub>2</sub> 2:2). A detailed analysis of the NEMO binding sites indicates that almost all the residues of the NEMO unoccupied site are less exposed to the solvent

than those on the occupied one (**Fig. S12**). While the solvent accessible surface area (SASA) of the occupied site is 13 nm<sup>2</sup> the one of the unoccupied is 11.9 nm<sup>2</sup> showing that both binding sites are not equal after the binding of Ub<sub>2</sub>. Altogether these observations indicate that, in solution, binding of linear Ub<sub>2</sub> to NEMO<sub>UBAN</sub> induces allosteric effects that modulate the overall structure and dynamics of the NEMO dimer. These observations also suggest that the 2:2 highly symmetric binding mode observed in the dense and ordered crystalline state becomes entropically unfavorable to the more flexible and far less dense solution state of the complex.

# Discussion

Structural biology investigations on poly-ubiquitins have mostly focused on di-ubiquitins observing that different protein linkages correspond to different protein dynamics leading to different exposed regions for the binding with partners<sup>47-49, 51-56</sup>. Ubiquitin signaling has been found associated not only to the linkage type but also to the length of the ubiquitin chains<sup>8-10</sup>. Here we first develop an efficient and accurate integrative approach to characterize the conformational ensembles of linear poly-ubiquitin by combining the Martini coarse-grain force field with SAXS experiments in the framework of Metainference. We then employ our method to try to rationalize the length dependent behavior of linear poly-ubiquitins and the consequence for the interaction with their partner NEMO. **Figure 6** rationalizes the observed differences in binding by comparing our free Ub<sub>N</sub> ensembles with our NEMO-bound ensemble. The fraction of bound-like configurations in the Ub<sub>2</sub> ensemble is a small fraction of the total ensemble suggesting a large conformational entropy loss upon binding. This is likely compensated by a release of a large number of water molecules from the binding interfaces upon binding in order to result in a final entropy gain as indicated by ITC (**Table 1**). The probability of finding at least one di-ubiquitin pair in a bound-like configuration in the Ub<sub>3</sub> and Ub<sub>4</sub> ensemble increases slightly more than linearly (3.2%, 6.7%, and 12.0% for Ub<sub>2</sub>, Ub<sub>3</sub> and Ub<sub>4</sub>, respectively) suggesting that longer poly-ubiquitins are likely to bind NEMO more favorably with respect to shorter ones. In fact, the SEC-SLS experiments show that NEMO bound to Ub<sub>4</sub> eluted as a single bound peak in comparison with Ub<sub>3</sub> (**Table S7** and **Fig. S9**). Since both NEMO:Ub<sub>4</sub> and NEMO:Ub<sub>3</sub> have similar *K<sub>D</sub>*s this can indicate a difference in kinetic stability. To provide a structural interpretation for this hypothesis, we calculated the actual probability of finding the full poly-ubiquitin in a configuration compatible with the binding (and thus avoiding configurations that would lead to a steric clash with NEMO (**Figure 6b-d**)) from our free polyubiquitin ensembles. The probability decreases from Ub<sub>2</sub> to Ub<sub>4</sub> (3.2%, 1.5% and 1.4% for Ub<sub>2</sub>, Ub<sub>3</sub>, and Ub<sub>4</sub>, respectively) which can lead to entropy loss. At the same time, non-specific flanking interactions between poly-ubiquitin and NEMO far from the binding site can increase the enthalpy. This is also in agreement with previous measures where, using a longer NEMO construct (NEMO<sub>242-419</sub>) that could provide more surface for interactions, affinities of 3 and 0.3

$\mu\text{M}$  were reported for Ub<sub>2</sub> and Ub<sub>4</sub>, respectively<sup>57</sup>. This principle is also common for intrinsically disordered proteins possibly modulating the lifetime of complexes<sup>58</sup>. Such long-range effects would be less pronounced for a less entropic chain and can play a length dependent role in the overall interaction.

Interestingly, the probability of finding two di-ubiquitin pairs in a bound like configuration is essentially negligible for both Ub<sub>3</sub> and Ub<sub>4</sub> providing a rationale why a 2:1 NEMO<sub>258-350</sub>:Ub<sub>4</sub> interaction is favored by entropy with respect to the 4:1. Making use of our polymer model, we can also speculate that a long-enough poly-ubiquitin may be able to bind two NEMO dimers with a higher order (i.e. 4:1) stoichiometry with respect to the 2:1 observed for the Ub<sub>2</sub>-Ub<sub>4</sub> range. Given that the end-to-end distance of Ub<sub>2</sub> corresponds to one half of that of our NEMO construct (5.8 nm for Ub<sub>2</sub> and 11.4 nm for NEMO), and that the two Ub<sub>2</sub> units that bind the two NEMO should be allowed to be flexible, one can estimate the length of this poly-ubiquitin to be such that  $e2e(N - 2) \geq 11.4 \text{ nm}$ . This results in a minimum length of  $N=8$  ubiquitins. While this result will not be quantitative when considering a full-length NEMO, it suggests a possible need for such long chains in the assembly of the IKK complex.

## Conclusions

The combined use of experiments and molecular dynamics simulation is a powerful tool to investigate the structure and dynamics of biomolecules and provide a ground for the functional interpretation of protein dynamics. Here, we combined SAXS and coarse-grained Martini simulations to study the conformational dynamics of linear poly-ubiquitins and their binding to NEMO accurately and efficiently. The resulting conformational ensembles allowed us to propose that linear-poly-ubiquitin behave as self-avoiding polymer chains. This might also apply for poly-ubiquitins in general (with a different characteristic lengths). Combining structural studies with multiple biophysical experiments, we provide a systematic assessment of the effect of the poly-ubiquitin chain length in the molecular recognition of cognate proteins, suggesting that poly-ubiquitin may modulate the binding with their partners in a length-dependent manner.

## Methods

### *Coarse-grained molecular dynamics simulations*

Coarse-grained molecular dynamics simulations (CG-MD) were applied to investigate the dynamic of linear di- tri- and tetraubiquitin (Ub<sub>2</sub>, Ub<sub>3</sub>, Ub<sub>4</sub>), as well as Ub<sub>2</sub> with bound NEMO. In total, 11 different simulations have been performed with a total simulation length of 780  $\mu\text{s}$ . An overview of all simulations can be found in **Table S3**.

All CG-simulations were run using Gromacs 2016.3<sup>59</sup> and the Martini force field<sup>34-35</sup>. Additionally, an elastic network model was used to conserve the secondary and tertiary structure. In the case of poly-ubiquitin, the elastic network inside was only defined for the core region (from residue 1 to 70, 77 to 146, 153 to 222 or 229 to 298) and not between different domains nor for the linker region. All simulations were performed with periodic boundary conditions, and the systems were solvated with a 0.1 M NaCl solution and run as an isothermal-isobaric (NpT) with a temperature of 300 K and a pressure of 1 bar using 20 fs time steps. To control the temperature and pressure, the v-rescale thermostat<sup>60</sup> was used with a coupling constant  $\tau_t = 1.0$  ps together with the Parrinello-Rahman barostat<sup>61</sup> and a coupling constant of  $\tau_p = 20.0$  ps and compressibility of  $\chi = 3.0 \times 10^{-4}$  bar<sup>-1</sup>. Nonbonded interactions were treated with a dielectric constant of 15 and using a cutoff distance of 1.1 nm. VMD was used for visualization<sup>62</sup>.

For many simulations, the Martini 2.2. force field was modified to increase the protein-water (P-W) interaction, which was achieved by giving water beads their own atom type with a 5% larger C6 parameter in interactions with all other atom types resulting in around 5% higher protein-water interactions.

Parallel-biased metadynamics<sup>37, 63</sup>, as implemented in PLUMED2<sup>64</sup>, was utilized to enhance the sampling of the conformational space, together with the multiple walker approach<sup>65</sup> with 112 replicas for free poly-ubiquitin or 64 replicas for free NEMO as well as bound NEMO with Ub<sub>2</sub>, where each replica had a different starting conformation. The employed collective variables were the distances between the centers of the different ubiquitin cores, a torsion angle between the centers of residue 1 to 36 and 37 to 70 of two different ubiquitins, the radius of gyration (calculated only with backbone atoms) and the alphabeta collective variable describing the torsional angles for linkers between the ubiquitin pairs. In total, four collective variables for Ub<sub>2</sub>, 9 for Ub<sub>3</sub> and 16 for Ub<sub>4</sub> were used. In the case of simulations with NEMO, additional collective variables were employed for the shape of NEMO and distances between Ub<sub>2</sub> and NEMO. The bias factor of the well-tempered metadynamics<sup>66</sup> was set to 10, the frequency for the hill addition was 200 (every 4 ps), the height of gaussian hills was 0.1 kJ/mol for simulations with di- and Ub<sub>3</sub>, 0.075 kJ/mol for simulations with tetraubiquitin and 0.02 kJ/mol for the simulation with NEMO bound to Ub<sub>2</sub>. The flexible Gaussian approach<sup>67</sup> was used to determine the Gaussian width during the simulation.

Metainference<sup>36</sup>, a method based on Bayesian inference, was used to integrate experimental SAXS data into simulations and was coupled with metadynamics (M&M)<sup>45</sup>. The calculation of the SAXS intensities from a CG Martini representation is implemented in the PLUMED-ISDB module<sup>68-69</sup> using the parameters derived by Niebling et al<sup>70</sup> and the Debye equation. The SAXS data of the different systems were fitted with a 16<sup>th</sup>-degree polynomial to calculate points used for restraints. 21 equidistant points between for q between 0.017 to 0.24 nm<sup>-1</sup> were used for Ub<sub>2</sub>, Ub<sub>3</sub> and for free and bound NEMO, 19 points for q between 0.025 to 0.19 nm<sup>-1</sup> for Ub<sub>4</sub>. The range depends on the quality of the experimental data. An

initial scaling value was determined by comparing the calculated and experimental SAXS intensities for the lowest  $q$  value. Metainference was used with the outlier noise model<sup>36</sup> for each data point, and the restraints were applied every 5<sup>th</sup> step. A scaling factor and offset for the experimental data were sampled using a flat prior between 0.9 to 1.1 or -1 to 1. The error for calculating an average quantity  $\sigma_{\text{mean}}$  was determined automatically<sup>71</sup> from the maximum standard error over 2 ps of simulations.

Six different simulations with Ub<sub>2</sub> were performed using: Martini 2.2 with metadynamics, with increased P-W-interaction and metadynamics, with M&M, with M&M applied every step, and with Martini 3 beta. Notably, Martini 3 beta was not stable with metadynamics, therefore 112 replicas were run on a longer time scale. The SAXS data of Ub<sub>2</sub> were taken from Vincendeau et al<sup>23</sup>. Notably, the profile was measured with Ub<sub>2</sub> containing a HIS-tag on the N-terminus which was also modeled. Ub<sub>3</sub> and Ub<sub>4</sub> simulations were run with increased P-W interaction, metadynamics with and without Metainference and SAXS. All poly-ubiquitin simulations were run for at least 500 ns per replica. The simulations with free NEMO and NEMO bound to Ub<sub>2</sub> were performed with increased P-W interaction and M&M with SAXS for at least 100 ns per replica. The SAXS data of NEMO bound to Ub<sub>2</sub> were taken from Vincendeau et al<sup>23</sup>.

The plumed input files as well as the modified Martini topology files are deposited in PLUMED-NEST<sup>72</sup> as plumID:20.009.

### ***Protein expression and purification***

Human NEMO<sub>258-350</sub> C347S was expressed and purified as described in Vincendeau et al<sup>23</sup>. Protein concentration was determined by measuring the absorbance at 205 nm using specific absorbance for NEMO<sub>258-350</sub> C347S of 300990 M<sup>-1</sup> cm<sup>-1</sup>, respectively<sup>73</sup>.

The constructs for the expression of Ub<sub>3</sub> and Ub<sub>4</sub> were a kind gift of Dr. Paul Elliott and Dr. David Komander (MRC Laboratory of Molecular Biology, Cambridge, UK). The constructs were transformed into *E. coli* strain BL21 (DE3) and cultured at 20°C in 2-L flasks containing 500 ml ZYM 5052 auto-induction medium<sup>74</sup> and 100 µg/ml carbenicillin. Cells were harvested by centrifugation after reaching saturation, resuspended in 60 ml lysis buffer (50mM Tris-HCl, 300mM NaCl, 10mM MgCl<sub>2</sub>, 10 µg/ml DNaseI, 1mM AEBSF.HCl, 0.2% (v/v) NP-40, 1 mg/ml lysozyme, pH 8.0), and lysed by sonication. The lysate was clarified by centrifugation (40,000 x  $g$ ) and filtration (0.2 µm). The supernatant was heated in a water bath for 10-15 minutes at 60°C and the precipitate removed by centrifugation. The supernatant was dialyzed overnight against 2 L of buffer A (50 mM sodium acetate pH 4.5), clarified by centrifugation and applied to a 5-ml HiTrap SP HP column (GE Healthcare), equilibrated in buffer A. Bound proteins were eluted using a linear gradient (10 column volumes) from 0 to 1 M NaCl in buffer A using an Äkta Purifier (GE Healthcare). Elution fractions (1.6 ml) were collected

in wells containing 250  $\mu$ l 1M Tris-HCl pH 9.0. Fractions containing Ub<sub>3</sub> or Ub<sub>4</sub> were pooled, concentrated and applied to a HiLoad 16/600 Superdex 75 column (GE Healthcare), equilibrated in buffer B (50mM Tris-HCl, 100mM NaCl, pH 7.4). The main elution peak containing Ub<sub>3</sub> or Ub<sub>4</sub> was collected and concentrated to approx. 3-6 mg/ml, flash frozen and stored at -80°C. Protein concentrations were determined by measuring the absorbance at 205 nm using specific absorbance for Ub<sub>3</sub> and Ub<sub>4</sub> of 747790 and 997980 M<sup>-1</sup> cm<sup>-1</sup>, respectively<sup>73</sup>. The Ub<sub>4</sub> concentration values used on interaction studies of Ub<sub>4</sub> with NEMO were corrected by 30% based on the SEC/SLS results.

### ***Small angle X-ray scattering measurements***

SAXS measurements were performed on a Rigaku BIOSAXS1000 instrument with an HF007 microfocus generator equipped with a Cu-target at 40 kV and 30mA. Transmissions were measured with a photodiode beamstop, q-calibration was made by an Ag-behenate measurement. Absolute calibration was done with calibrated glassy carbon<sup>75</sup>. Measurements were done in four 900 second frames, which were averaged. Under these conditions, no radiation damage was detected. Circular averaging and background subtraction were done with the Rigaku SAXSLab software v 3.0.1r1.

Radii of gyration were calculated with the ATSAS package v 2.8.0<sup>76</sup>. Fits for the MW determination were made in Origin v 9. SAXS measurements were made at 293 K using a buffer contained 300mM NaCl and 50mM Tris-HCl at a pH of 8.0. Experiments on the free proteins were performed at the following concentrations (**Table S5**): NEMO at 2.34, 4.62 and 7.72 mg/ml, Ub<sub>3</sub> at 3.41, 6.72 and 11.17 mg/ml, Ub<sub>4</sub> at 4.5, 9.1 and 15.1 mg/ml. Experiments with NEMO and Ub<sub>3</sub> and Ub<sub>4</sub> at different ratios were performed at two concentrations (between 3-4 and 7-8 mg/ml) at the following ratios: NEMO:Ub<sub>3</sub> at 1.4:1 and 2.7:1 ratios, NEMO:Ub<sub>4</sub> at 1.0:1, 2.1:1 and 3.1:1 ratios. No concentration-dependent effects were detected.

### ***Isothermal Titration Calorimetry***

ITC measurements were carried out at 298 K using a PEAQ-ITC titration microcalorimeter (MicroCal, Malvern). The NEMO to Ub<sub>3</sub> calorimetric titration consisted of 19 injections of 2  $\mu$ l of a 2.13 mM NEMO solution, into the reaction cell containing 300  $\mu$ l of 94.71  $\mu$ M Ub<sub>3</sub>, at a stirring speed of 750 rpm. The NEMO to Ub<sub>4</sub> calorimetric titration consisted of 19 injections of 2  $\mu$ l of a 2.84 mM NEMO solution, into the reaction cell containing 300  $\mu$ l of 120.6  $\mu$ M Ub<sub>4</sub>, at a stirring speed of 750 rpm. Sample conditions were 50 mM sodium phosphate pH 7.0 and 50 mM NaCl. The heat of dilution was obtained by titrating NEMO into the sample cell containing only buffer. Experiments were done in triplicate. The ITC data were analyzed using the software MicroCal PEAQ-ITC analysis software. Parameters are presented as averages  $\pm$  standard errors.

### ***Size Exclusion Chromatography with Static Light Scattering***

Static light scattering (SLS) experiments were performed of NEMO mutant (C347S) in complex with tri- and tetra-ubiquitin at 30°C using a Viscotek TDA 305 triple array detector (Malvern Instruments) downstream to an Äkta Purifier (GE Healthcare) equipped with an analytical size exclusion column (Superdex 200 10/300 GL, GE Healthcare) at 4°C. The samples were run at approx. 8 mg/ml at a flow rate of 0.5 ml/min. The experiments were performed using a Tris buffer (50 mM Tris-HCl, 300 mM NaCl, pH 8.0) and a phosphate buffer (50 mM sodium phosphate, 50 mM NaCl, pH 7.0). The molecular masses of the samples were calculated from the refractive index and right-angle light-scattering signals using Omnisec (Malvern Instruments). The SLS detector was calibrated with a 4 mg/ml BSA solution using 66.4 kDa for the BSA monomer and a  $dn/dc$  value of 0.185 ml/g for all protein samples.

### ***Surface Plasmon Resonance measurements***

SPR measurements were performed at 25°C using a Pioneer FE instrument (FortéBio, Molecular Devices). Ub<sub>3</sub> and Ub<sub>4</sub> were covalently immobilized on to two different flow cell channels on a biosensor chip by amine coupling to 456 and 721 RU, respectively, using a 10 mM NaOAc pH 5 immobilization buffer. NEMO was injected in a two-fold concentration series over immobilized ubiquitins at 30 µL/min flow rate using a PBS running buffer (50 mM sodium phosphate, 50 mM NaCl, 0.005% tween, pH 7). The data were analyzed using Qdat Data Analysis Tool version 2.6.3.0 (FortéBio). The sensorgrams were corrected for buffer effects and unspecific binding to the chip matrix by subtraction of blank and reference surface (a blank flow cell channel activated by injection of EDC/NHS and inactivated by injection of ethanolamine). The equilibrium dissociation constants ( $K_D$ ) were estimated by plotting responses at equilibrium ( $R_{eq}$ ) against the injected concentration and curve fitted to a Langmuir (1:1) binding isotherm.

### ***Acknowledgments***

The authors acknowledge Paul Elliott and David Komander (MRC Laboratory of Molecular Biology, Cambridge, UK) for the Ub<sub>3</sub> and Ub<sub>4</sub> plasmids. We acknowledge Daniel Krappmann (Helmholtz Zentrum München, Germany) for reading the paper and providing useful feedback. AJ and CC acknowledge support by the Technische Universität München-Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under Grant Agreement No. 291763. SMØS. MS acknowledges funding by the DFG SFB1035, and AB acknowledge the support by the Lundbeck Foundation (Grant R190-2014-3710). We gratefully acknowledge the Gauss Centre for Supercomputing e.V. ([www.gauss-centre.eu](http://www.gauss-centre.eu)) for funding this project by providing computing time on the GCS

Super-computer SuperMUC at the Leibniz Supercomputing Center (LRZ, [www.lrz.de](http://www.lrz.de)). We acknowledge SAXS measurements at the facility of the SFB1035 at Department Chemie, Technische Universität München.

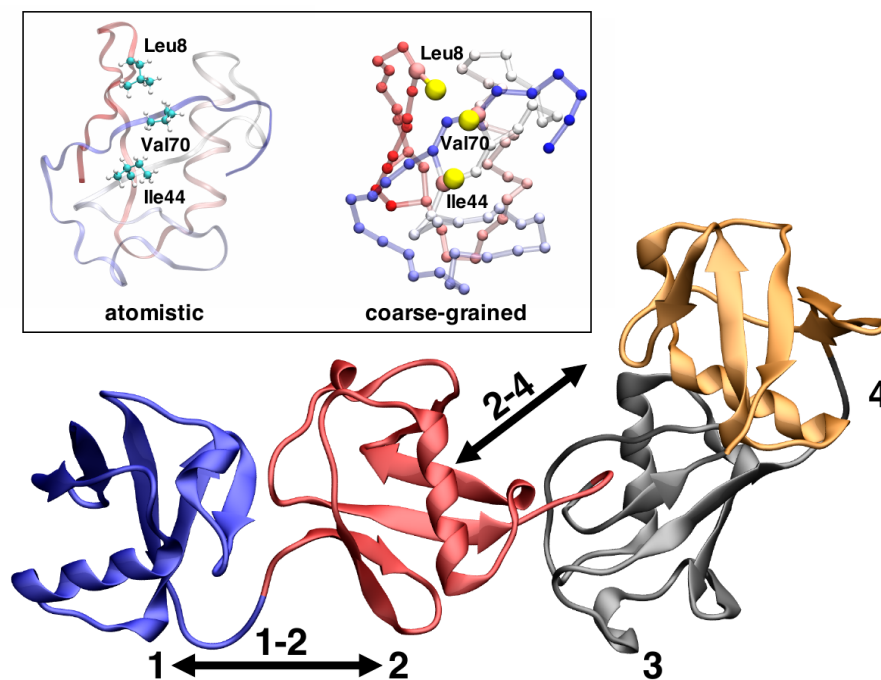
## References

- Haglund, K.; Dikic, I., Ubiquitylation and cell signaling. *EMBO J* **2005**, *24* (19), 3353-9.
- Deshaies, R. J.; Joazeiro, C. A., RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **2009**, *78*, 399-434.
- Li, W.; Ye, Y., Polyubiquitin chains: functions, structures, and mechanisms. *Cell Mol Life Sci* **2008**, *65* (15), 2397-406.
- Peng, J.; Schwartz, D.; Elias, J. E.; Thoreen, C. C.; Cheng, D.; Marsischky, G.; Roelofs, J.; Finley, D.; Gygi, S. P., A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **2003**, *21* (8), 921-6.
- Kirisako, T.; Kamei, K.; Murata, S.; Kato, M.; Fukumoto, H.; Kanie, M.; Sano, S.; Tokunaga, F.; Tanaka, K.; Iwai, K., A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J* **2006**, *25* (20), 4877-87.
- Iwai, K.; Tokunaga, F., Linear polyubiquitination: a new regulator of NF-kappaB activation. *EMBO Rep* **2009**, *10* (7), 706-13.
- Iwai, K.; Fujita, H.; Sasaki, Y., Linear ubiquitin chains: NF-kappaB signalling, cell death and beyond. *Nat Rev Mol Cell Biol* **2014**, *15* (8), 503-8.
- Komander, D.; Rape, M., The ubiquitin code. *Annu Rev Biochem* **2012**, *81*, 203-29.
- Zhao, S.; Ulrich, H. D., Distinct consequences of posttranslational modification by linear versus K63-linked polyubiquitin chains. *Proc Natl Acad Sci U S A* **2010**, *107* (17), 7704-9.
- Fushman, D.; Wilkinson, K. D., Structure and recognition of polyubiquitin chains of different lengths and linkage. *F1000 Biol Rep* **2011**, *3*, 26.
- Akutsu, M.; Dikic, I.; Bremm, A., Ubiquitin chain diversity at a glance. *J Cell Sci* **2016**, *129* (5), 875-80.
- Thrower, J. S.; Hoffman, L.; Rechsteiner, M.; Pickart, C. M., Recognition of the polyubiquitin proteolytic signal. *EMBO J* **2000**, *19* (1), 94-102.
- Tsuchiya, H.; Ohtake, F.; Arai, N.; Kaiho, A.; Yasuda, S.; Tanaka, K.; Saeki, Y., In Vivo Ubiquitin Linkage-type Analysis Reveals that the Cdc48-Rad23/Dsk2 Axis Contributes to K48-Linked Chain Specificity of the Proteasome. *Mol Cell* **2017**, *66* (4), 488-502 e7.
- von Delbruck, M.; Kniss, A.; Rogov, V. V.; Pluska, L.; Bagola, K.; Lohr, F.; Guntert, P.; Sommer, T.; Dotsch, V., The CUE Domain of Cue1 Aligns Growing Ubiquitin Chains with Ubc7 for Rapid Elongation. *Mol Cell* **2016**, *62* (6), 918-928.
- Pierce, N. W.; Kleiger, G.; Shan, S. O.; Deshaies, R. J., Detection of sequential polyubiquitylation on a millisecond timescale. *Nature* **2009**, *462* (7273), 615-9.
- Kovacev, J.; Wu, K.; Spratt, D. E.; Chong, R. A.; Lee, C.; Nayak, J.; Shaw, G. S.; Pan, Z. Q., A snapshot of ubiquitin chain elongation: lysine 48-tetra-ubiquitin slows down ubiquitination. *J Biol Chem* **2014**, *289* (10), 7068-81.
- Gerlach, B.; Cordier, S. M.; Schmukle, A. C.; Emmerich, C. H.; Rieser, E.; Haas, T. L.; Webb, A. I.; Rickard, J. A.; Anderton, H.; Wong, W. W.; Nachbur, U.; Gangoda, L.; Warnken, U.; Purcell, A. W.; Silke, J.; Walczak, H., Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature* **2011**, *471* (7340), 591-6.
- Wertz, I. E.; Newton, K.; Seshasayee, D.; Kusam, S.; Lam, C.; Zhang, J.; Popovych, N.; Helgason, E.; Schoeffler, A.; Jeet, S.; Ramamoorthi, N.; Kategaya, L.; Newman, R. J.; Horikawa, K.; Dugger, D.; Sandoval, W.; Mukund, S.; Zindal, A.; Martin, F.; Quan, C.; Tom, J.; Fairbrother, W. J.; Townsend, M.; Warming, S.; DeVoss, J.; Liu, J.; Dueber, E.; Caplazi, P.; Lee, W. P.; Goodnow, C. C.; Balazs, M.; Yu, K.; Kolumam, G.; Dixit, V. M., Phosphorylation and linear ubiquitin direct A20 inhibition of inflammation. *Nature* **2015**, *528* (7582), 370-5.
- Senft, D.; Qi, J.; Ronai, Z. A., Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer* **2018**, *18* (2), 69-88.
- Catici, D. A.; Horne, J. E.; Cooper, G. E.; Pudney, C. R., Polyubiquitin Drives the Molecular Interactions of the NF-kappaB Essential Modulator (NEMO) by Allosteric Regulation. *J Biol Chem* **2015**, *290* (22), 14130-9.
- Hauenstein, A. V.; Xu, G.; Kabaleeswaran, V.; Wu, H., Evidence for M1-Linked Polyubiquitin-Mediated Conformational Change in NEMO. *Journal of Molecular Biology* **2017**, *429* (24), 3793-3800.
- Hinz, M.; Scheidereit, C., The Ikb kinase complex in NF-kB regulation and beyond. *EMBO reports* **2014**, *15* (1), 46-61.
- Vincendeau, M.; Hadian, K.; Messias, A. C.; Brenke, J. K.; Halander, J.; Griesbach, R.; Greczmiel, U.; Bertossi, A.; Stehle, R.; Nagel, D.; Demski, K.; Velvarska, H.; Niessing, D.; Geerloff, A.; Sattler, M.; Krappmann, D., Inhibition of Canonical NF-kappaB Signaling by a Small Molecule Targeting NEMO-Ubiquitin Interaction. *Sci Rep* **2016**, *6*, 18934.
- 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., Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell* **2009**, *136* (6), 1098-109.

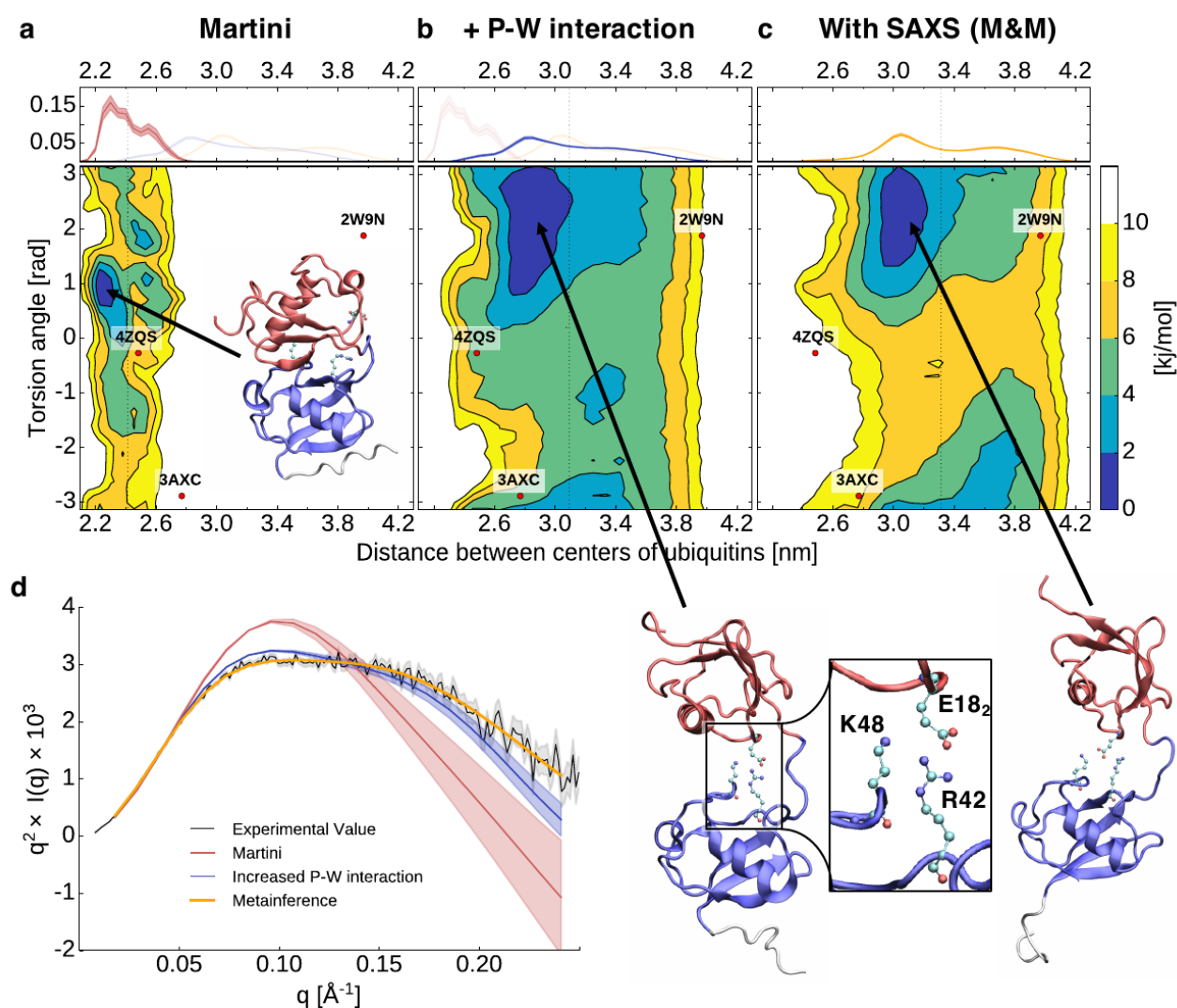
25. Ivins, F. J.; Montgomery, M. G.; Smith, S. J.; Morris-Davies, A. C.; Taylor, I. A.; Rittinger, K., NEMO oligomerization and its ubiquitin-binding properties. *Biochem J* **2009**, *421* (2), 243-51.
26. Polley, S.; Huang, D. B.; Hauenstein, A. V.; Fusco, A. J.; Zhong, X.; Vu, D.; Schrofelbauer, B.; Kim, Y.; Hoffmann, A.; Verma, I. M.; Ghosh, G.; Huxford, T., A structural basis for I $\kappa$ B kinase 2 activation via oligomerization-dependent trans auto-phosphorylation. *PLoS Biol* **2013**, *11* (6), e1001581.
27. DiDonato, J. A.; Hayakawa, M.; Rothwarf, D. M.; Zandi, E.; Karin, M., A cytokine-responsive I $\kappa$ B kinase that activates the transcription factor NF- $\kappa$ B. *Nature* **1997**, *388* (6642), 548-54.
28. Thach, T. T.; Shin, D.; Han, S.; Lee, S., New conformations of linear polyubiquitin chains from crystallographic and solution-scattering studies expand the conformational space of polyubiquitin. *Acta Crystallogr D Struct Biol* **2016**, *72* (Pt 4), 524-35.
29. Boldon, L.; Laliberte, F.; Liu, L., Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Rev* **2015**, *6*, 25661.
30. Zuckerman, D. M., Equilibrium sampling in biomolecular simulations. *Annu Rev Biophys* **2011**, *40*, 41-62.
31. Camilloni, C.; Pietrucci, F., Advanced simulation techniques for the thermodynamic and kinetic characterization of biological systems. *Advances in Physics: X* **2018**, *3* (1).
32. Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A., Coarse-Grained Protein Models and Their Applications. *Chem Rev* **2016**, *116* (14), 7898-936.
33. Bonomi, M.; Heller, G. T.; Camilloni, C.; Vendruscolo, M., Principles of protein structural ensemble determination. *Curr Opin Struct Biol* **2017**, *42*, 106-116.
34. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H., The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* **2007**, *111* (27), 7812-24.
35. de Jong, D. H.; Singh, G.; Bennett, W. F.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J., Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J Chem Theory Comput* **2013**, *9* (1), 687-97.
36. Bonomi, M.; Camilloni, C.; Cavalli, A.; Vendruscolo, M., Metainference: A Bayesian inference method for heterogeneous systems. *Sci Adv* **2016**, *2* (1), e1501177.
37. Laio, A.; Parrinello, M., Escaping free-energy minima. *Proc Natl Acad Sci U S A* **2002**, *99* (20), 12562-6.
38. Komander, D.; Reyes-Turcu, F.; Licchesi, J. D.; Odenwaelde, P.; Wilkinson, K. D.; Barford, D., Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep* **2009**, *10* (5), 466-73.
39. Rohaim, A.; Kawasaki, M.; Kato, R.; Dikic, I.; Wakatsuki, S., Structure of a compact conformation of linear diubiquitin. *Acta Crystallogr D Biol Crystallogr* **2012**, *68* (Pt 2), 102-8.
40. Stark, A. C.; Andrews, C. T.; Elcock, A. H., Toward optimized potential functions for protein-protein interactions in aqueous solutions: osmotic second virial coefficient calculations using the MARTINI coarse-grained force field. *J Chem Theory Comput* **2013**, *9* (9).
41. Javanainen, M.; Martinez-Seara, H.; Vattulainen, I., Excessive aggregation of membrane proteins in the Martini model. *PLoS One* **2017**, *12* (11), e0187936.
42. Yesylevskyy, S. O.; Schafer, L. V.; Sengupta, D.; Marrink, S. J., Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput Biol* **2010**, *6* (6), e1000810.
43. Gotz, A. W.; Bucher, D.; Lindert, S.; McCammon, J. A., Dipeptide Aggregation in Aqueous Solution from Fixed Point-Charge Force Fields. *J Chem Theory Comput* **2014**, *10* (4), 1631-1637.
44. Robustelli, P.; Piana, S.; Shaw, D. E., Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A* **2018**, *115* (21), E4758-E4766.
45. Bonomi, M.; Camilloni, C.; Vendruscolo, M., Metadynamic metainference: Enhanced sampling of the metainference ensemble using metadynamics. *Sci Rep* **2016**, *6*, 31232.
46. Bhattacharjee, S. M.; Giacometti, A.; Maritan, A., Flory theory for polymers. *J Phys Condens Matter* **2013**, *25* (50), 503101.
47. Paissoni, C.; Jussupow, A.; Camilloni, C., Determination of Protein Structural Ensembles by Hybrid-Resolution SAXS Restrained Molecular Dynamics. *J Chem Theory Comput* **2020**.
48. Castaneda, C. A.; Kashyap, T. R.; Nakasone, M. A.; Krueger, S.; Fushman, D., Unique structural, dynamical, and functional properties of k11-linked polyubiquitin chains. *Structure* **2013**, *21* (7), 1168-81.
49. Berlin, K.; Castaneda, C. A.; Schneidman-Duhovny, D.; Sali, A.; Nava-Tudela, A.; Fushman, D., Recovering a representative conformational ensemble from underdetermined macromolecular structural data. *J Am Chem Soc* **2013**, *135* (44), 16595-609.
50. Lin, S. M.; Lin, S. C.; Hong, J. Y.; Su, T. W.; Kuo, B. J.; Chang, W. H.; Tu, Y. F.; Lo, Y. C., Structural Insights into Linear Tri-ubiquitin Recognition by A20-Binding Inhibitor of NF- $\kappa$ B, ABIN-2. *Structure* **2017**, *25* (1), 66-78.
51. Liu, Z.; Tang, C., Ensemble structure description of Lys63-linked diubiquitin. *Data Brief* **2016**, *7*, 81-8.
52. Liu, Z.; Gong, Z.; Jiang, W. X.; Yang, J.; Zhu, W. K.; Guo, D. C.; Zhang, W. P.; Liu, M. L.; Tang, C., Lys63-linked ubiquitin chain adopts multiple conformational states for specific target recognition. *Elife* **2015**, *4*.

53. Wang, Y.; Tang, C.; Wang, E.; Wang, J., PolyUbiquitin chain linkage topology selects the functions from the underlying binding landscape. *PLoS Comput Biol* **2014**, *10* (7), e1003691.
54. Ye, Y.; Blaser, G.; Horrocks, M. H.; Ruedas-Rama, M. J.; Ibrahim, S.; Zhukov, A. A.; Orte, A.; Klenerman, D.; Jackson, S. E.; Komander, D., Ubiquitin chain conformation regulates recognition and activity of interacting proteins. *Nature* **2012**, *492* (7428), 266-70.
55. Bremm, A.; Freund, S. M.; Komander, D., Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nat Struct Mol Biol* **2010**, *17* (8), 939-47.
56. Tenno, T.; Fujiwara, K.; Tochio, H.; Iwai, K.; Morita, E. H.; Hayashi, H.; Murata, S.; Hiroaki, H.; Sato, M.; Tanaka, K.; Shirakawa, M., Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains. *Genes Cells* **2004**, *9* (10), 865-75.
57. Hadian, K.; Griesbach, R. A.; Dornauer, S.; Wanger, T. M.; Nagel, D.; Metlitzky, M.; Beisker, W.; Schmidt-Supprian, M.; Krappmann, D., NF-kappaB essential modulator (NEMO) interaction with linear and lys-63 ubiquitin chains contributes to NF-kappaB activation. *J Biol Chem* **2011**, *286* (29), 26107-17.
58. Crabtree, M. D.; Borchers, W.; Poosapati, A.; Shammass, S. L.; Daughdrill, G. W.; Clarke, J., Conserved Helix-Flanking Prolines Modulate Intrinsically Disordered Protein:Target Affinity by Altering the Lifetime of the Bound Complex. *Biochemistry* **2017**, *56* (18), 2379-2384.
59. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19-25.
60. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J Chem Phys* **2007**, *126* (1), 014101.
61. Parrinello, M.; Rahman, A., Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **1981**, *52* (12), 7182-7190.
62. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14* (1), 33-38.
63. Pfaffentner, J.; Bonomi, M., Efficient Sampling of High-Dimensional Free-Energy Landscapes with Parallel Bias Metadynamics. *J Chem Theory Comput* **2015**, *11* (11), 5062-7.
64. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **2014**, *185* (2), 604-613.
65. Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M., Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J Phys Chem B* **2006**, *110* (8), 3533-9.
66. Barducci, A.; Bussi, G.; Parrinello, M., Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett* **2008**, *100* (2), 020603.
67. Branduardi, D.; Bussi, G.; Parrinello, M., Metadynamics with Adaptive Gaussians. *J Chem Theory Comput* **2012**, *8* (7), 2247-54.
68. Bonomi, M.; Camilloni, C., Integrative structural and dynamical biology with PLUMED-ISDB. *Bioinformatics* **2017**, *33* (24), 3999-4000.
69. Paissoni, C.; Jussupow, A.; Camilloni, C., Martini bead form factors for nucleic acids and their application in the refinement of protein-nucleic acid complexes against SAXS data. *Journal of Applied Crystallography* **2019**, *52* (2), 394-402.
70. Niebling, S.; Bjorling, A.; Westenhoff, S., MARTINI bead form factors for the analysis of time-resolved X-ray scattering of proteins. *J Appl Crystallogr* **2014**, *47* (Pt 4), 1190-1198.
71. Löhr, T.; Jussupow, A.; Camilloni, C., Metadynamic meta-inference: Convergence towards force field independent structural ensembles of a disordered peptide. *The Journal of Chemical Physics* **2017**, *146* (16).
72. consortium, P., Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* **2019**, *16* (8), 670-673.
73. Anthis, N. J.; Clore, G. M., Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Sci* **2013**, *22* (6), 851-8.
74. Studier, F. W., Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **2005**, *41* (1), 207-234.
75. Fan, L.; Degen, M.; Bendle, S.; Grupido, N.; Ilavsky, J., The Absolute Calibration of a Small-Angle Scattering Instrument with a Laboratory X-ray Source. *Journal of Physics: Conference Series* **2010**, *247*.
76. Petoukhov, M. V.; Franke, D.; Shkumatov, A. V.; Tria, G.; Kikhney, A. G.; Gajda, M.; Gorba, C.; Mertens, H. D.; Konarev, P. V.; Svergun, D. I., New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* **2012**, *45* (Pt 2), 342-350.

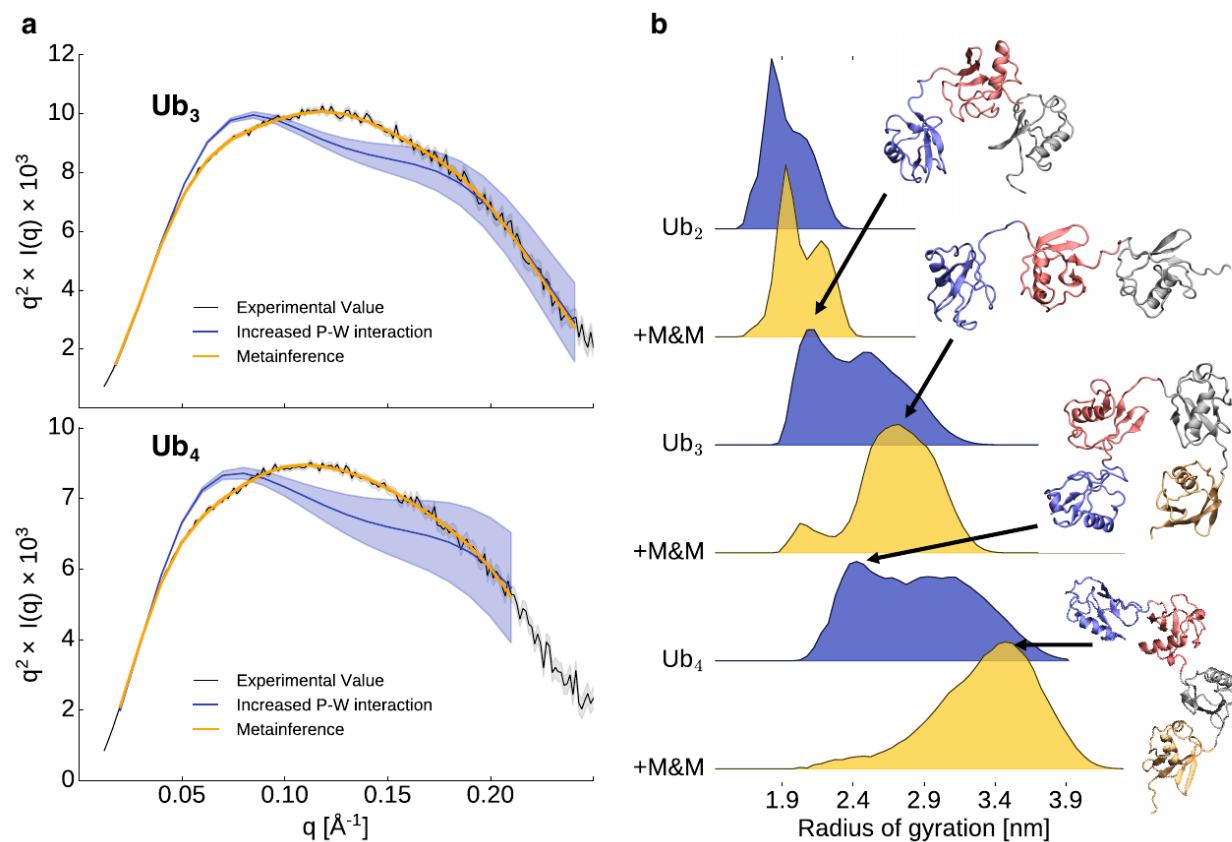




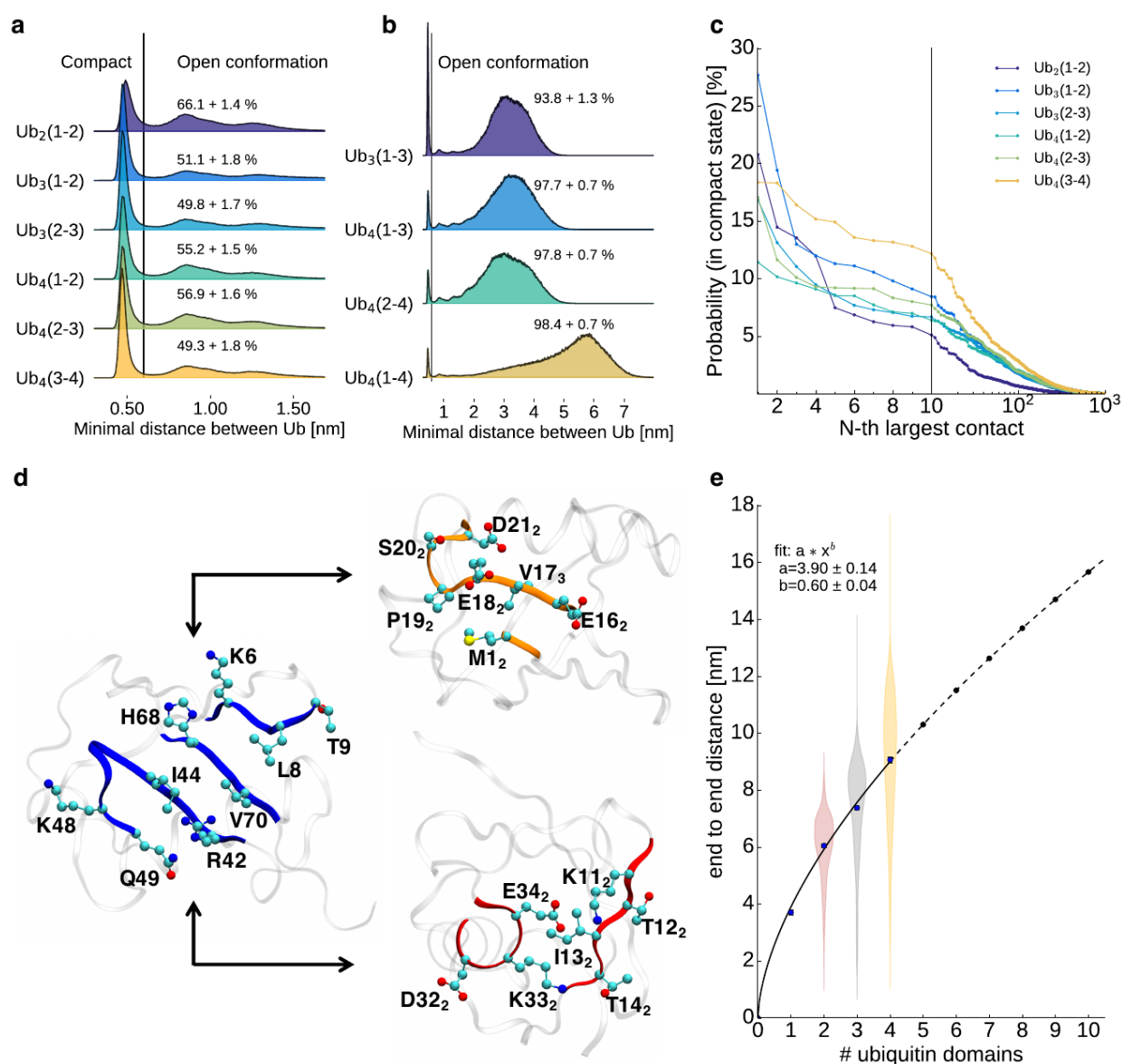
**Figure 1. Schematic illustrations linear poly-ubiquitins.** Cartoon representation of linear tetra-ubiquitin, the ubiquitin domains are numbered from the N-terminal to the C-terminal from 1 to 4. In the inset is shown an atomistic and coarse grained (Martini) ubiquitin domain highlighting the hydrophobic patch (Ile44, Val70 Leu8).



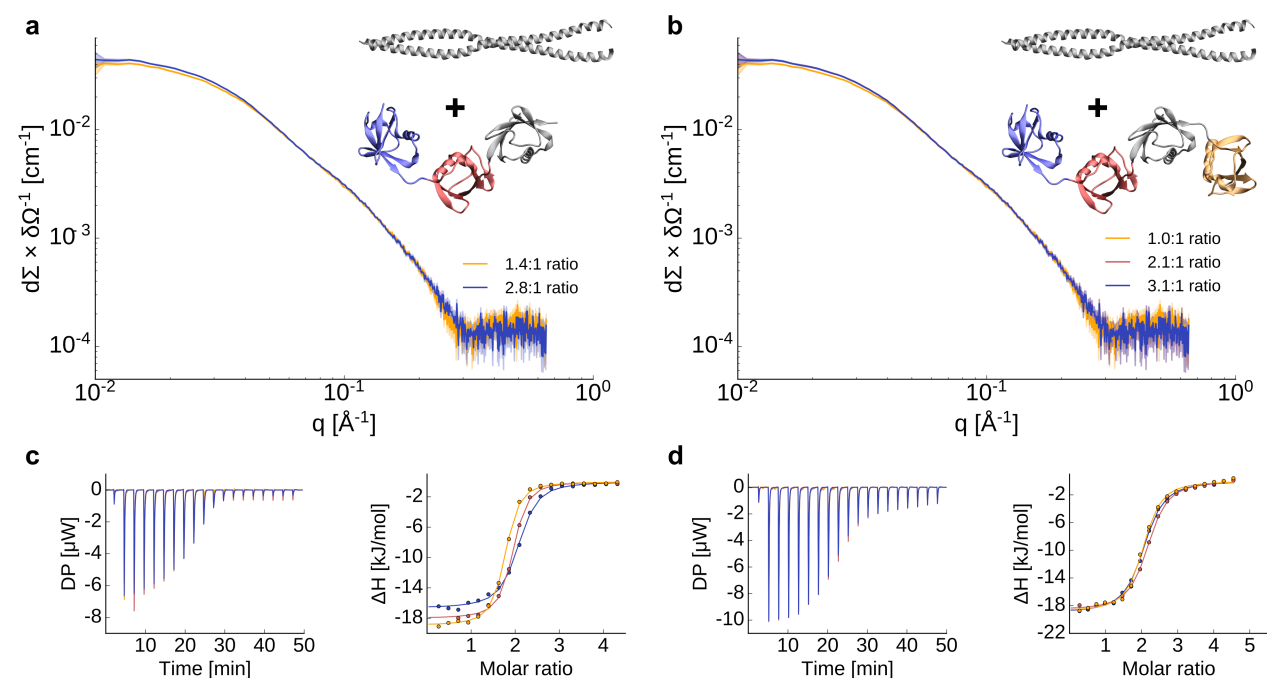
**Figure 2. Characterization of the dynamics of linear di-ubiquitin.** **a)-c)** Free energy landscapes (in kJ/mol) as a function of the distance between the center-of-mass of the two ubiquitin domains and their relative orientation (measured as the torsion angle between two axes defined using the first and second half of the sequence of each ubiquitin, see Methods). The dots represent the coordinates associated with the available di-ubiquitin crystal structures. On top is shown the probability distribution of the distance between the centers of the two ubiquitin domains. **d)** Experimental and from simulation calculated Kratky plot. The shaded area represents the error range.



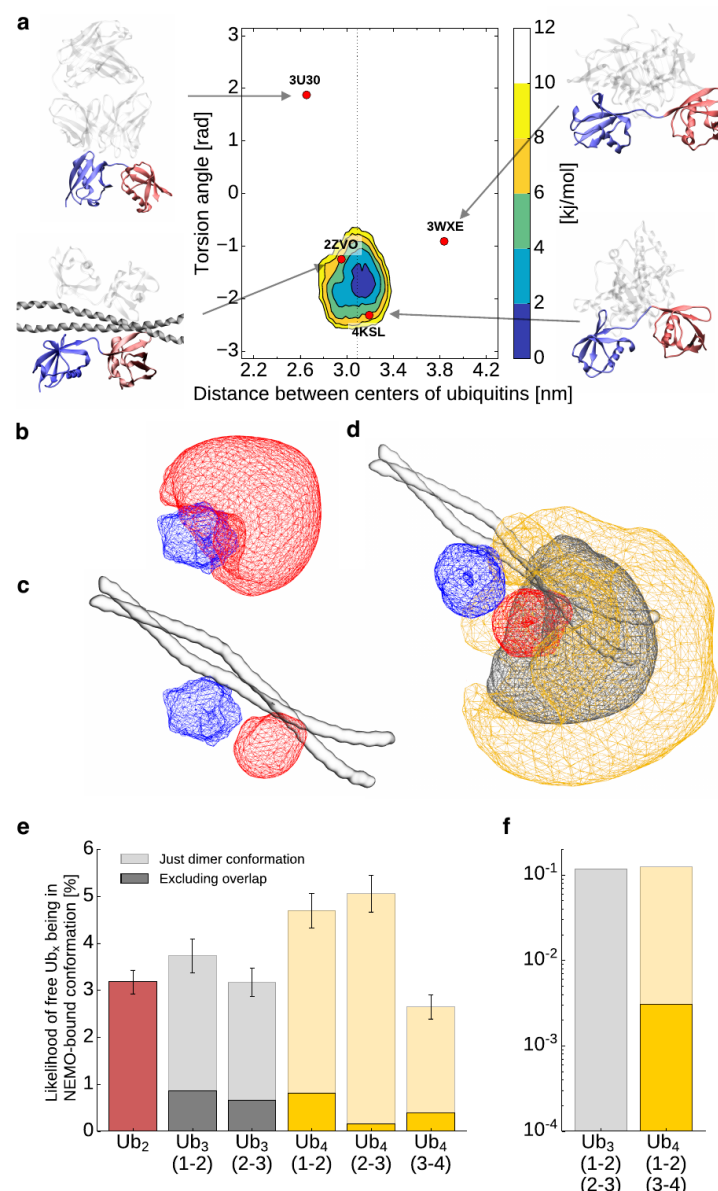
**Figure 3. Characterization of the dynamics of linear tri- and tetraubiquitin. a)** Experimental and from simulations calculated Kratky plot for tri- and tetraubiquitin. The shaded area represents the error range. **b)** Distribution of the radius of gyration from the ensemble with and without M&M.



**Figure 4. Intramolecular interactions of poly-ubiquitin.** a) Minimum distance distribution between two neighboring ubiquitin cores (residue 1 to 70, residue 77 to 146, and so forth). Structures with a distance larger than 0.6 are defined as open. b) Minimum distance distribution between two non-neighboring ubiquitin cores. Structures with a distance larger than 0.6 are defined as open. c) The probability of finding contacts between two amino acids of neighboring ubiquitin cores d) Interaction surface of two neighboring ubiquitins. Residues from the blue marked surface (first ubiquitin - left) are interacting with residues of the orange marked surface (middle) or red marked surface (right) of the second ubiquitin. e) Average end to end distance of a linear poly-ubiquitin chain.



**Figure 5. Effect of chain length on the binding of NEMO** a) and b) SAXS experiments for different ratios of NEMO and Ub<sub>3</sub> (a) and Ub<sub>4</sub> (b). c) and d) Isothermal titration calorimetry (ITC) measurement of the interaction of NEMO with -Ub<sub>3</sub> (c) and Ub<sub>4</sub> (d). NEMO was titrated into the poly-ubiquitin solutions. The experiment was repeated three times.



**Figure 6. Comparison between free and NEMO-bound poly-ubiquitin ensembles. a)** Free energy landscapes (in kJ/mol) as a function of the distance between the centers of the two ubiquitin domains and their relative orientation for Ub<sub>2</sub> bound to NEMO. The dots represent the coordinates associated with the available crystal structures with Ub<sub>2</sub> bound to different proteins. **b-c)** Conformational space of free (b) and NEMO-bound (c) ubiquitin pairs in Ub<sub>2</sub>. The blue area represents the first ubiquitin, while the red area shows the conformation of the second ubiquitin relative to the first one. **d)** Conformational space of third (gray area) and fourth (orange area) ubiquitin in Ub<sub>4</sub> with the first Ub-pair being in a NEMO-bound conformation. **e-f)** Probability of free Ub<sub>2</sub>, Ub<sub>3</sub>, and Ub<sub>4</sub> of being in a NEMO-bound conformation for one (e) or two (f) NEMO dimers. The transparent bars show the likelihood of the individual pairs being in the NEMO-bound conformation (RMSD < 6 Å compared to the average Ub<sub>2</sub> structure in the NEMO-bound simulation). The dark bars show

the probability of being in the NEMO-bound conformation, excluding structures with an overlap between the non-bound ubiquitins and NEMO.

**Table 1. Isothermal titration calorimetry (ITC) measurement of the NEMO interaction with Ub<sub>2</sub>, Ub<sub>3</sub>, and Ub<sub>4</sub>.**

NEMO was titrated into the ubiquitin solutions in 50 mM sodium phosphate pH 7, 50 mM NaCl. Values are averages  $\pm$  standard errors from three measurements. The individual ITC curves are shown in Fig. S8. \* Experiments taken from Vincendeau et al<sup>23</sup>. A stoichiometry of N=2 corresponds to one NEMO dimer binding to one poly-ubiquitin protein.

NEMO +	Ub <sub>2</sub> *	Ub <sub>3</sub>	Ub <sub>4</sub>
<b>N</b>	1.86 $\pm$ 0.01	1.84 $\pm$ 0.09	2.02 $\pm$ 0.04
<b>K<sub>D</sub> [<math>\mu</math>M]</b>	1.81 $\pm$ 0.12	1.57 $\pm$ 0.28	4.13 $\pm$ 0.30
<b><math>\Delta</math>H [kJ/mol]</b>	-16.9 $\pm$ 0.1	-17.9 $\pm$ 0.6	-18.8 $\pm$ 0.1
<b>-T<math>\Delta</math>S [kJ/mol]</b>	-15.8	-15.3 $\pm$ 0.4	-11.9 $\pm$ 0.2

**Table 2. Molecular weight determination.** SAXS and size exclusion chromatography (SEC) in combination with static light scattering (SLS) were used to determine the molecular weight of NEMO, Ub<sub>3</sub>, Ub<sub>4</sub>, NEMO:Ub<sub>3</sub> and NEMO:Ub<sub>4</sub>. The conditions were 50 mM Tris.HCl pH 8, 300 mM NaCl.

Protein	MW <sub>exp</sub> [kDa] (SAXS)	MW <sub>exp</sub> [kDa] (SEC/SLS)	MW <sub>calc</sub> [kDa]
<b>NEMO dimer</b>	24.4	20.6	22.0
<b>Ub<sub>3</sub></b>	20.5	24.1	25.7
<b>Ub<sub>4</sub></b>	29.9	32.3	34.2
<b>NEMO:Ub<sub>3</sub> 2:1</b>		41.1	47.7
<b>NEMO:Ub<sub>4</sub> 2:1</b>		56.2	56.2
<b>NEMO:Ub<sub>3</sub> (1.4:1)</b>	40.0		40.8
<b>NEMO:Ub<sub>3</sub> (2.7:1)</b>	38.7		40.6
<b>NEMO:Ub<sub>4</sub> (1.0:1)</b>	37.9		46.6
<b>NEMO:Ub<sub>4</sub> (3.1:1)</b>	45.1		44.0