# ELECTRONIC SUPPLEMENTARY INFORMATION

# One touch is all it takes: the supramolecular interaction between ubiquitin and lanthanide complexes revisited by paramagnetic NMR and molecular dynamics

Karen Dos Santos,[a,1] Alessio Bartocci,[b,c,d,1] Natacha Gillet,[e] Sandrine Denis-Quanquin,[e] Amandine Roux,[e,f] Eugene Lin,[a] Zeren Xu,[a] Raphael Finizola,[e] Pauline Chedozeau,[a] Xi Chen,[a] Cédric Caradeuc,[a] Mathieu Baudin,[a,g] Gildas Bertho,[a] François Riobé,[h] Olivier Maury,[e] Elise Dumont,[i,j*] and Nicolas Giraud.[a]*

[a] *Université Paris Cité, Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, UMR CNRS 8601, Paris, France*
[b] *Department of Physics, University of Trento, Via Sommarive 14, I-38123 Trento, Italy*
[c] *INFN-TIFPA, Trento Institute for Fundamental Physics and Applications, Via Sommarive 14, I-38123 Trento, Italy*
[d] *Institut de Chimie de Strasbourg, UMR 7177, CNRS, Universitè de Strasbourg, Strasbourg Cedex 67083, France*
[e] *Univ. Lyon, ENS de Lyon, CNRS UMR 5182, Université Claude Bernard Lyon 1, Laboratoire de Chimie, F69342, Lyon, France.*
[f] *Polyvalan SAS, Lyon, France*
[g] *Laboratoire des Biomolécules, LBM, Département de chimie, École Normale Supérieure, PSL Université, Sorbonne Université 45 Rue d'Ulm, 75005 Paris (France)*
[h] *Univ. Bordeaux, CNRS, Bordeaux INP, ICMCB UMR 5026, F-33600 Pessac, France*
[i] *Université Côte d'Azur, CNRS, Institut de Chimie de Nice, UMR 7272, 06108 Nice, France*
[j] *Institut Universitaire de France, 5 rue Descartes, 75005 Paris, France*

** Correspondence to elise.dumont@univ-cotedazur.fr; nicolas.giraud@u-paris.fr*
[1] *K. Dos Santos and A. Bartocci contributed equally to this work as first authors*

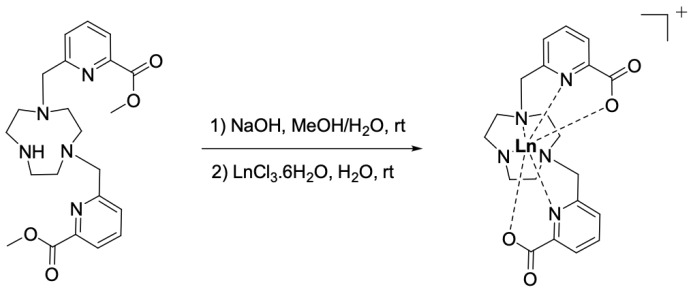
## Table of Contents

## Synthesis

### *General procedures*

High resolution mass spectrometry measurements were performed at Centre Commun de Spectrometrie de Masse (Villeurbanne, France). Starting materials were purchased from Sigma Aldrich®, Acros Organics® or Alfa Aesar® with the best available quality grade and used without further purification, otherwise noted. All reactions were routinely performed under argon atmosphere in anhydrous solvents when relevant. Preparative HPLC separations were performed on an Agilent Technologies 1260 Infinity apparatus at 5 mL/min, using Agilent Zorbax SB-C18 prepHT (5 µm, 21.2 x 100 mm) column. Mobile phase consisted in a gradient of solvent A (0.25 M ammonium formate in $H_2O$) and B ($CH_3CN$).

### *Synthesis of complexes*




**1** **[Gd-Xo4]$^+$** Ln = $Gd^{3+}$
**[Yb-Xo4]$^+$** Ln = $Yb^{3+}$

**[Yb-Xo4]$^+$**. Compound 1 [1] (136 mg, 0.32 mmol) was solubilized in a minimum of MeOH (2 mL) and 10 mL of $H_2O$. Then NaOH (35 mg, 0.88 mmol) was added, and the mixture was stirred at room temperature upon completion of the deprotection of the esters (which can be checked by LCMS and NMR). The mixture was then diluted with 10 mL of $H_2O$ and acidified using HCl 1M until pH 7. Then $YbCl_3 \cdot 6H_2O$ (136 mg, 0.35 mmol) was added and the pH adjusted again to 5.5 using NaOH 1M. The reaction mixture was stirred at room temperature for 36h. The crude product was purified by preparative RP-HPLC (gradient 5 to 100% B in 16 min) to yield complex [Yb-Xo4]$^+$ (45 mg, 24%) as a white powder after freeze-drying.

HRMS (ESI) calculated for $C_{20}H_{25}N_5O_5Yb.H_2O$ 589.1242. Exp 589.1241 $[M+H_2O]^+$.

**[Gd-Xo4]$^+$**. Compound 1 (182 mg, 0.40 mmol) was solubilized in a minimum of MeOH (6 mL) and 20 mL of $H_2O$. Then NaOH (43 mg, 1.06 mmol) was added, and the mixture was stirred at room temperature upon completion of the deprotection of the esters (which can be checked by LCMS and NMR). The mixture was then diluted with 10 mL of $H_2O$ and acidified using HCl 1M until pH 6.5. Then $GdCl_3 \cdot 6H_2O$ (176 mg, 0.47 mmol) was added and the pH adjusted again to 5.7 using NaOH 1M. The reaction mixture was stirred at room temperature overnight. The crude product was purified by preparative RP-HPLC (gradient 5 to 100% B in 16 min) to yield complex [Gd-Xo4]$^+$ (106 mg, 44%) as a pale yellow powder after freeze-drying.

HRMS (ESI) calculated for $C_{20}H_{23}N_5O_4Gd$ 555.0990. Exp 555.0994 $[M]^+$.

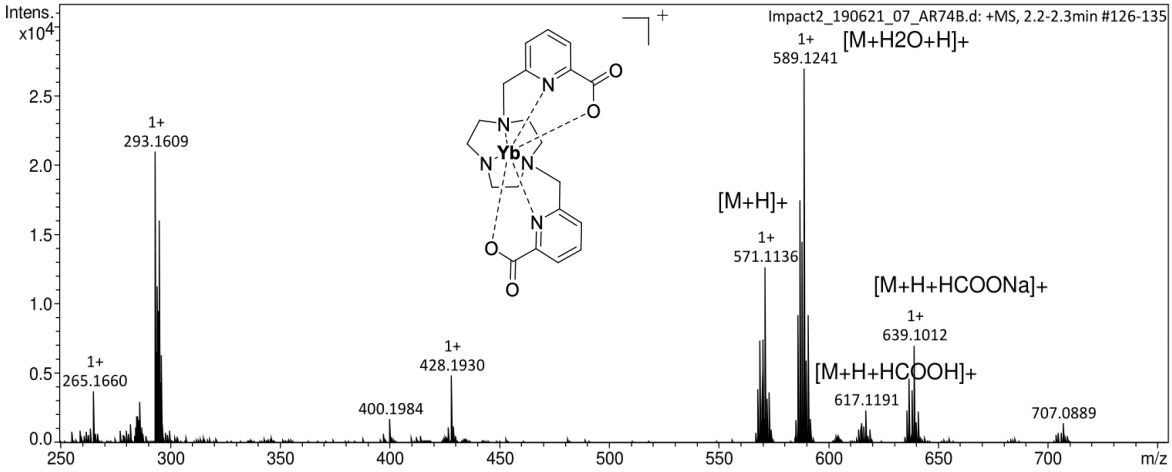
# CENTRE COMMUN DE SPECTROMETRIE DE MASSE

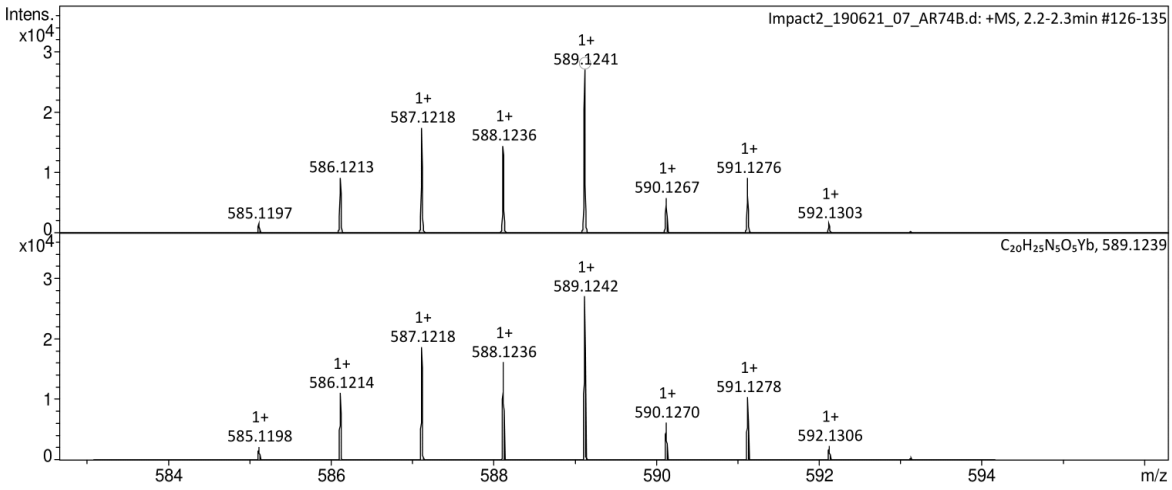
**Analysis Info**

| | | | |
|---|---|---|---|
| Analysis Name | Impact2_190621_07_AR74B.d | | |
| Method | Tune_pos_Standard.m | Acquisition Date | 6/21/2019 11:46:03 AM |
| Comment | | Instrument / Ser# | impact II 1825265.10081 |

**Acquisition Parameter**

| | | | | | |
|---|---|---|---|---|---|
| Source Type | ESI | Ion Polarity | Positive | Set Nebulizer | 0.3 Bar |
| Focus | Active | Set Capillary | 3500 V | Set Dry Heater | 200 °C |
| Scan Begin | 50 m/z | Set End Plate Offset | -500 V | Set Dry Gas | 4.0 l/min |
| Scan End | 1500 m/z | Set Collision Cell RF | 750.0 Vpp | Set Divert Valve | Source |




| Meas. m/z | Ion Formula | m/z | err [ppm] | mSigma |
|---|---|---|---|---|
| 571.1136 | C20H23N5O4Yb | 571.1133 | 0.0 | 62.2 |
| 589.1241 | C20H25N5O5Yb | 589.1239 | 0.2 | 38.7 |
| 617.1191 | C21H25N5O6Yb | 617.1188 | -0.0 | 43.7 |
| 639.1012 | C21H24N5NaO6Yb | 639.1007 | -0.2 | 43.6 |

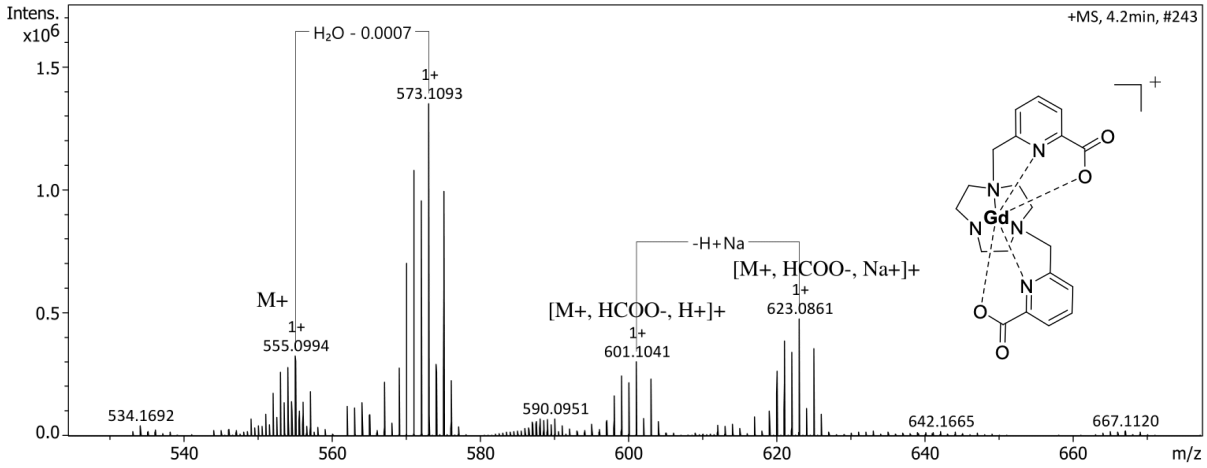
# CENTRE COMMUN DE SPECTROMETRIE DE MASSE

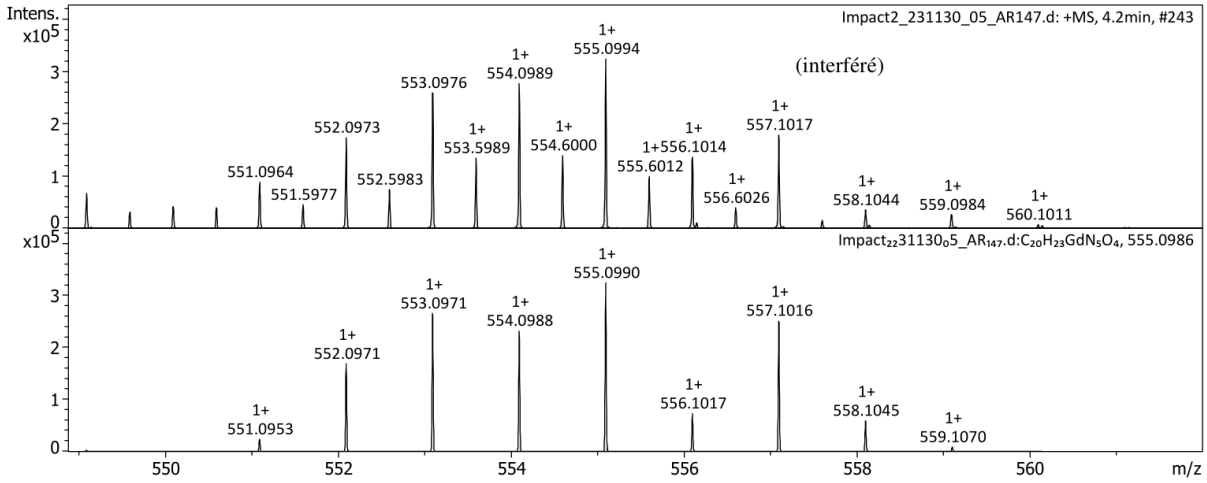
**Analysis Info**

| | | | |
|---|---|---|---|
| Analysis Name | Impact2_231130_05_AR147.d | | |
| Method | Tune_pos_Standard.m | Acquisition Date | 11/30/2023 11:02:01 AM |
| Comment | | Instrument / Ser# | impact II 1825265.10081 |

**Acquisition Parameter**

| | | | | | |
|---|---|---|---|---|---|
| Source Type | ESI | Ion Polarity | Positive | Set Nebulizer | 0.3 Bar |
| Focus | Active | Set Capillary | 4500 V | Set Dry Heater | 200 °C |
| Scan Begin | 50 m/z | Set End Plate Offset | -500 V | Set Dry Gas | 4.0 l/min |
| Scan End | 1000 m/z | Set Collision Cell RF | 750.0 Vpp | Set Divert Valve | Source |




| Meas. m/z | Ion Formula | Sum Formula | err [ppm] | mSigma | Adduct | z |
|---|---|---|---|---|---|---|
| 555.0994 | C20H23GdN5O4 | C20H23GdN5O4 | -0.7 | 132.1 | M | 1+ |
| 573.1093 | C20H25GdN5O5 | C20H25GdN5O5 | 0.6 | 46.6 | M | 1+ |
| 601.1041 | C21H25GdN5O6 | C21H24GdN5O6 | 0.6 | 46.1 | M+H | 1+ |
| 623.0861 | C21H24GdN5NaO6 | C21H24GdN5O6 | 0.6 | 48.6 | M+Na | 1+ |

## NMR Analyses

### *NMR assignment*

Ubiquitin was characterized in solution using 2D $^{1}H$-$^{1}H$ TOCSY, $^{1}H$-$^{1}H$ NOESY and $^{1}H$-$^{15}N$ HSQC NMR spectra. Experimental chemical shifts presented an overall good agreement with the literature, with absolute differences of NH and $H_{\alpha}$ < 0.1 ppm and $^{15}N$ < 0.5 ppm.[2] Except for M1, P19, P37, and P38, which are in line with the literature, each protein residue yielded a correlation signal on HSQC spectra. It is worth noting that T9, E24, G53 and G75 either had weak signals or could not be detected and, therefore, were not used (Table S 1).

Chemical shift assignment of NMR spectra was conducted using literature chemical shift values obtained from the Biological Magnetic Resonance Data Bank (BMRB) in CCPNMR Analysis.[3,4] Synthetic peak lists generated from BMRB entries 68,[5] 17769,[6] and 30693[2] were used to assign $^{1}H$-$^{1}H$ TOCSY and $^{1}H$-$^{15}N$ HSQC spectra, respectively. Synthetic NOESY peak lists were generated using an experimental chemical shift list created based on $^{1}H$-$^{1}H$ TOCSY peak assignments and the ubiquitin PDB structure (PDB: 1D3Z), with interproton distance thresholds of 2.5, 2.7, 3.0, 3.5, 4.0, and 4.5 Å. These synthetic NOESY peak lists were then utilized to help assign the NOESY spectrum and verify TOCSY assignments.

***Table S 1.*** *$^{1}H$-$^{15}N$ HSQC spectrum chemical shift assignment of the protein uniformly $^{15}N$-labelled Ubiquitin. Side chain protons in position delta and epsilon are shown as Hδ and Hε.*

| Residue | $\delta_H$ (ppm) | $\delta_N$ (ppm) | Residue | $\delta_H$ (ppm) | $\delta_N$ (ppm) |
|---|---|---|---|---|---|
| Q2 | 8.87 | 122.95 | L43 | 8.74 | 124.53 |
| Q2_Hε22 | 7.60 | 112.47 | I44 | 9.06 | 122.50 |
| Q2_Hε21 | 6.83 | 112.47 | F45 | 8.78 | 125.02 |
| I3 | 8.24 | 115.20 | A46 | 8.95 | 133.05 |
| F4 | 8.54 | 118.70 | G47 | 8.03 | 102.50 |
| V5 | 9.22 | 121.36 | K48 | 7.91 | 122.14 |
| K6 | 8.91 | 128.11 | Q49 | 8.59 | 123.27 |
| T7 | 8.67 | 115.47 | Q49_Hε22 | 6.67 | 112.02 |
| L8L | 9.04 | 121.30 | Q49_Hε21 | 7.66 | 112.02 |
| G10 | 7.77 | 109.34 | L50 | 8.50 | 125.78 |
| K11 | 7.21 | 121.95 | E51 | 8.28 | 123.21 |
| T12 | 8.57 | 120.65 | D52 | 8.10 | 120.49 |
| I13 | 9.49 | 127.74 | R54 | 7.41 | 119.43 |
| T14 | 8.68 | 121.67 | R55 | 8.77 | 108.91 |
| L15 | 8.67 | 125.25 | L56 | 8.09 | 118.14 |
| Q16 | 8.05 | 122.58 | S57 | 8.41 | 113.56 |
| V17 | 8.87 | 117.65 | D58 | 7.88 | 124.63 |
| Q18 | 8.60 | 119.43 | Y59 | 7.19 | 115.84 |
| S20 | 6.96 | 103.52 | N60 | 8.09 | 116.04 |
| D21 | 7.99 | 123.99 | N60_Hδ22 | 6.75 | 111.53 |
| T22 | 7.81 | 109.06 | N60_Hδ21 | 7.48 | 111.53 |
| I23 | 8.46 | 121.38 | I61 | 7.18 | 118.99 |
| N25 | 7.86 | 121.49 | Q62 | 7.55 | 125.03 |
| N25_Hδ21 | 6.82 | 109.92 | Q62_Hε21 | 7.22 | 112.42 |
| N25_Hδ22 | 7.78 | 109.92 | Q62_Hε22 | 6.75 | 112.44 |
| V26 | 8.05 | 122.27 | K63 | 8.43 | 120.66 |
| K27 | 8.49 | 119.03 | E64 | 9.25 | 114.72 |
| A28 | 7.91 | 123.51 | S65 | 7.60 | 115.02 |
| K29 | 7.80 | 120.32 | T66 | 8.64 | 117.54 |
| I30 | 8.22 | 121.43 | L67 | 9.34 | 127.92 |

| Residue | $\delta_H$ (ppm) | $\delta_N$ (ppm) | Residue | $\delta_H$ (ppm) | $\delta_N$ (ppm) |
|---|---|---|---|---|---|
| Q31 | 8.49 | 123.64 | H68 | 9.16 | 119.74 |
| Q31_Hε22 | 7.58 | 110.12 | L69 | 8.21 | 123.82 |
| Q31_Hε21 | 6.75 | 110.11 | V70 | 9.12 | 126.79 |
| D32 | 7.96 | 119.83 | L71 | 8.02 | 123.19 |
| K33 | 7.37 | 115.58 | R72 | 8.54 | 123.91 |
| E34 | 8.66 | 114.41 | L73 | 8.29 | 124.62 |
| G35 | 8.43 | 108.94 | R74 | 8.38 | 122.12 |
| I36 | 6.10 | 120.41 | G76 | 7.88 | 115.18 |
| D39 | 8.46 | 113.67 | | | |
| Q40 | 7.75 | 116.98 | | | |
| Q40_Hε21 | 7.61 | 111.11 | | | |
| Q40_Hε22 | 6.68 | 111.11 | | | |
| Q41 | 7.41 | 118.10 | | | |
| Q41_Hε21 | 6.13 | 104.34 | | | |
| Q41_Hε22 | 6.45 | 104.33 | | | |
| R42 | 8.46 | 123.20 | | | |

***Interaction between ubiquitin and $[Yb(DPA)_3]^{3-}$***

***Table S 2.*** *$^{1}H$-$^{15}N$ Experimental chemical shifts (ppm) of the Pseudo Contact Shifts (PCS) for the ubiquitin upon the addition of [Na]3[Yb(DPA)3]. Side chain protons in position delta and epsilon are shown as Hδ and Hε.*

| Residue | 0eq Yb | 0.1eq Yb | 0.2eq Yb | 0.4eq Yb | 0.8eq Yb | 2eq Yb | 4eq Yb | 8eq Yb | 10eq Yb |
|---|---|---|---|---|---|---|---|---|---|
| Q2 | 8.87 | 8.87 | 8.87 | 8.87 | 8.88 | 8.88 | 8.87 | 8.89 | 8.89 |
| Q2_Hε21 | 7.60 | 7.60 | 7.60 | 7.60 | 7.59 | 7.59 | 7.58 | 7.56 | 7.56 |
| Q2_Hε22 | 6.83 | 6.83 | 6.83 | 6.84 | 6.87 | 6.87 | 6.89 | 6.94 | 6.95 |
| I3 | 8.24 | 8.24 | 8.25 | 8.25 | 8.25 | 8.25 | 8.26 | 8.27 | 8.27 |
| F4 | 8.54 | 8.54 | 8.54 | 8.54 | 8.56 | 8.56 | 8.57 | 8.60 | 8.61 |
| V5 | 9.22 | 9.22 | 9.22 | 9.22 | 9.22 | 9.22 | 9.23 | 9.23 | 9.23 |
| K6 | 8.91 | 8.91 | 8.91 | 8.91 | 8.91 | 8.91 | 8.91 | 8.90 | 8.90 |
| T7 | 8.67 | 8.67 | 8.67 | 8.66 | 8.66 | 8.66 | 8.65 | 8.63 | 8.63 |
| L8 | 9.04 | 9.04 | 9.04 | 9.04 | 9.05 | 9.04 | 9.05 | 9.05 | 9.05 |
| G10 | 7.77 | 7.77 | 7.77 | 7.77 | 7.76 | 7.76 | 7.76 | 7.75 | 7.75 |
| K11 | 7.21 | 7.21 | 7.20 | 7.20 | 7.20 | 7.20 | 7.19 | 7.17 | 7.17 |
| T12 | 8.57 | 8.57 | 8.57 | 8.57 | 8.56 | 8.56 | 8.56 | 8.55 | 8.54 |
| I13 | 9.49 | 9.49 | 9.49 | 9.49 | 9.49 | 9.49 | 9.48 | 9.48 | 9.48 |
| T14 | 8.68 | 8.68 | 8.68 | 8.68 | 8.68 | 8.68 | 8.68 | 8.66 | 8.66 |
| L15 | 8.67 | 8.67 | 8.67 | 8.67 | 8.68 | 8.67 | 8.68 | 8.68 | 8.68 |
| E16 | 8.05 | 8.05 | 8.05 | 8.05 | 8.05 | 8.05 | 8.05 | 8.06 | 8.07 |
| V17 | 8.87 | 8.87 | 8.87 | 8.87 | 8.87 | 8.87 | 8.88 | 8.88 | 8.88 |
| E18 | 8.60 | 8.60 | 8.60 | 8.60 | 8.60 | 8.60 | 8.61 | 8.61 | 8.61 |
| S20 | 6.96 | 6.96 | 6.96 | 6.96 | 6.96 | 6.96 | 6.97 | 6.97 | 6.97 |
| D21 | 7.99 | 7.99 | 7.99 | 7.99 | 7.99 | 7.99 | 7.99 | 8.00 | 8.00 |
| T22 | 7.81 | 7.81 | 7.81 | 7.81 | 7.81 | 7.81 | 7.81 | 7.81 | 7.81 |
| I23 | 8.46 | 8.46 | 8.46 | 8.46 | 8.46 | 8.46 | 8.46 | 8.46 | 8.46 |
| N25 | 7.86 | 7.86 | 7.86 | 7.86 | 7.86 | 7.86 | 7.86 | 7.86 | 7.86 |
| N25_Hδ21 | 6.82 | 6.82 | 6.82 | 6.82 | 6.82 | 6.82 | 6.82 | 6.82 | 6.82 |
| N25_Hδ22 | 7.78 | 7.78 | 7.78 | 7.78 | 7.78 | 7.78 | 7.78 | 7.78 | 7.78 |
| V26 | 8.05 | 8.05 | 8.05 | 8.05 | 8.05 | 8.05 | 8.05 | 8.06 | 8.06 |
| K27 | 8.49 | 8.49 | 8.49 | 8.49 | 8.50 | 8.49 | 8.50 | 8.50 | 8.50 |
| A28 | 7.91 | 7.91 | 7.91 | 7.91 | 7.91 | 7.91 | 7.91 | 7.91 | 7.91 |
| K29 | 7.80 | 7.80 | 7.80 | 7.80 | 7.80 | 7.80 | 7.81 | 7.81 | 7.81 |
| I30 | 8.22 | 8.22 | 8.22 | 8.22 | 8.22 | 8.22 | 8.23 | 8.23 | 8.23 |
| Q31 | 8.49 | 8.49 | 8.49 | 8.49 | 8.49 | 8.49 | 8.49 | 8.49 | 8.48 |
| Q31_Hε22 | 7.58 | 7.58 | 7.58 | 7.58 | 7.58 | 7.58 | 7.58 | 7.58 | 7.58 |

| Residue | 0eq Yb | 0.1eq Yb | 0.2eq Yb | 0.4eq Yb | 0.8eq Yb | 2eq Yb | 4eq Yb | 8eq Yb | 10eq Yb |
|---|---|---|---|---|---|---|---|---|---|
| Q31_Hε21 | 6.75 | 6.75 | 6.75 | 6.75 | 6.75 | 6.74 | 6.74 | 6.74 | 6.74 |
| D32 | 7.96 | 7.96 | 7.96 | 7.96 | 7.96 | 7.96 | 7.96 | 7.96 | 7.96 |
| K33 | 7.37 | 7.37 | 7.37 | 7.37 | 7.37 | 7.37 | 7.37 | 7.37 | 7.37 |
| E34 | 8.66 | 8.66 | 8.66 | 8.66 | 8.65 | 8.65 | 8.65 | 8.65 | 8.66 |
| G35 | 8.43 | 8.43 | 8.43 | 8.43 | 8.43 | 8.43 | 8.43 | 8.42 | 8.42 |
| I36 | 6.10 | 6.10 | 6.10 | 6.10 | 6.10 | 6.10 | 6.09 | 6.08 | 6.08 |
| D39 | 8.46 | 8.46 | 8.46 | 8.46 | 8.47 | 8.47 | 8.47 | 8.47 | 8.47 |
| Q40 | 7.75 | 7.75 | 7.75 | 7.75 | 7.75 | 7.75 | 7.75 | 7.76 | 7.76 |
| Q40_Hε21 | 7.61 | 7.61 | 7.61 | 7.61 | 7.61 | 7.61 | 7.61 | 7.61 | 7.61 |
| Q40_Hε22 | 6.68 | 6.68 | 6.68 | 6.68 | 6.68 | 6.68 | 6.68 | 6.67 | 6.67 |
| Q41 | 7.41 | 7.41 | 7.41 | 7.41 | 7.41 | 7.41 | 7.48 | 7.45 | 7.45 |
| Q41_Hε21 | 6.13 | 6.13 | 6.13 | 6.13 | 6.13 | 6.13 | 6.12 | 6.12 | 6.11 |
| Q41_Hε22 | 6.45 | 6.45 | 6.45 | 6.45 | 6.45 | 6.45 | 6.45 | 6.45 | 6.44 |
| R42 | 8.46 | 8.46 | 8.45 | 8.46 | 8.46 | 8.46 | 8.46 | 8.48 | 8.48 |
| L43 | 8.74 | 8.74 | 8.74 | 8.74 | 8.74 | 8.74 | 8.74 | 8.73 | 8.73 |
| I44 | 9.06 | 9.06 | 9.06 | 9.06 | 9.04 | 9.04 | 9.02 | 8.98 | 8.97 |
| F45 | 8.78 | 8.79 | 8.79 | 8.79 | 8.77 | 8.77 | 8.76 | 8.75 | 8.73 |
| A46 | 8.95 | 8.96 | 8.96 | 9.04 | 9.03 | 9.02 | 9.04 | 9.04 | 9.03 |
| G47 | 8.03 | 8.03 | 8.02 | 8.01 | 7.94 | 7.94 | 7.86 | 7.73 | 7.68 |
| K48 | 7.91 | 7.91 | 7.90 | 7.91 | 7.87 | 7.87 | 7.84 | 7.79 | 7.77 |
| Q49 | 8.59 | 8.59 | 8.58 | 8.58 | 8.54 | 8.54 | 8.50 | 8.45 | 8.42 |
| Q49_Hε22 | 6.67 | 6.67 | 6.66 | 6.66 | 6.67 | 6.66 | 6.66 | 6.66 | 6.66 |
| Q49_Hε21 | 7.66 | 7.66 | 7.66 | 7.66 | 7.66 | 7.66 | 7.66 | 7.66 | 7.66 |
| L50 | 8.50 | 8.50 | 8.50 | 8.50 | 8.49 | 8.49 | 8.48 | 8.46 | 8.45 |
| E51 | 8.32 | 8.31 | 8.32 | 8.32 | 8.32 | 8.31 | 8.31 | 8.30 | 8.29 |
| D52 | 8.10 | 8.10 | 8.10 | 8.10 | 8.10 | 8.09 | 8.09 | 8.08 | 8.07 |
| R54 | 7.41 | 7.41 | 7.41 | 7.41 | 7.41 | 7.41 | 7.41 | 7.40 | 7.40 |
| T55 | 8.77 | 8.77 | 8.77 | 8.77 | 8.76 | 8.76 | 8.76 | 8.77 | 8.76 |
| L56 | 8.09 | 8.09 | 8.09 | 8.09 | 8.09 | 8.09 | 8.10 | 8.10 | 8.10 |
| S57 | 8.41 | 8.41 | 8.41 | 8.41 | 8.42 | 8.42 | 8.42 | 8.43 | 8.43 |
| D58 | 7.88 | 7.88 | 7.88 | 7.88 | 7.88 | 7.88 | 7.88 | 7.88 | 7.88 |
| Y59 | 7.19 | 7.19 | 7.19 | 7.19 | 7.19 | 7.19 | 7.20 | 7.20 | 7.21 |
| N60 | 8.09 | 8.09 | 8.09 | 8.09 | 8.10 | 8.09 | 8.11 | 8.12 | 8.12 |
| N60_Hδ22 | 6.75 | 6.75 | 6.75 | 6.76 | 6.76 | 6.76 | 6.77 | 6.78 | 6.79 |
| N60_Hδ21 | 7.48 | 7.48 | 7.48 | 7.49 | 7.50 | 7.50 | 7.51 | 7.53 | 7.54 |
| I61 | 7.18 | 7.18 | 7.18 | 7.19 | 7.21 | 7.22 | 7.22 | 7.25 | 7.26 |
| Q62 | 7.55 | 7.55 | 7.55 | 7.53 | 7.59 | 7.58 | 7.61 | 7.66 | 7.68 |

| Residue | 0eq Yb | 0.1eq Yb | 0.2eq Yb | 0.4eq Yb | 0.8eq Yb | 2eq Yb | 4eq Yb | 8eq Yb | 10eq Yb |
|---|---|---|---|---|---|---|---|---|---|
| Q62_Hε21 | 7.22 | 7.22 | 7.22 | 7.23 | 7.24 | 7.24 | 7.25 | 7.27 | 7.28 |
| Q62_Hε22 | 6.75 | 6.75 | 6.75 | 6.76 | 6.76 | 6.76 | 6.77 | 6.79 | 6.79 |
| K63 | 8.43 | 8.43 | 8.43 | 8.43 | 8.43 | 8.43 | 8.44 | 8.45 | 8.46 |
| E64 | 9.25 | 9.25 | 9.25 | 9.25 | 9.26 | 9.26 | 9.27 | 9.29 | 9.29 |
| S65 | 7.60 | 7.60 | 7.60 | 7.60 | 7.63 | 7.63 | 7.65 | 7.69 | 7.70 |
| T66 | 8.64 | 8.64 | 8.65 | 8.64 | 8.74 | 8.73 | 8.82 | 8.95 | 8.99 |
| L67 | 9.34 | 9.34 | 9.34 | 9.36 | 9.37 | 9.37 | 9.39 | 9.43 | 9.45 |
| H68 | 9.16 | 9.16 | 9.16 | 9.16 | 9.15 | 9.15 | 9.14 | 9.13 | 9.13 |
| L69 | 8.21 | 8.21 | 8.21 | 8.20 | 8.20 | 8.20 | 8.19 | 8.17 | 8.16 |
| V70 | 9.12 | 9.12 | 9.12 | 9.12 | 9.10 | 9.11 | 9.09 | 9.07 | 9.07 |
| L71 | 8.02 | 8.02 | 8.02 | 8.02 | 8.03 | 8.03 | 8.03 | 8.03 | 8.03 |
| R72 | 8.54 | 8.55 | 8.54 | 8.55 | 8.57 | 8.56 | 8.57 | 8.59 | 8.60 |
| L73 | 8.29 | 8.29 | 8.29 | 8.24 | 8.31 | 8.31 | 8.33 | 8.36 | 8.36 |
| R74 | 8.38 | 8.38 | 8.37 | 8.37 | 8.36 | 8.35 | 8.33 | 8.31 | 8.30 |
| G76 | 7.88 | 7.88 | 7.88 | 7.88 | 7.86 | 7.86 | 7.84 | 7.82 | 7.81 |

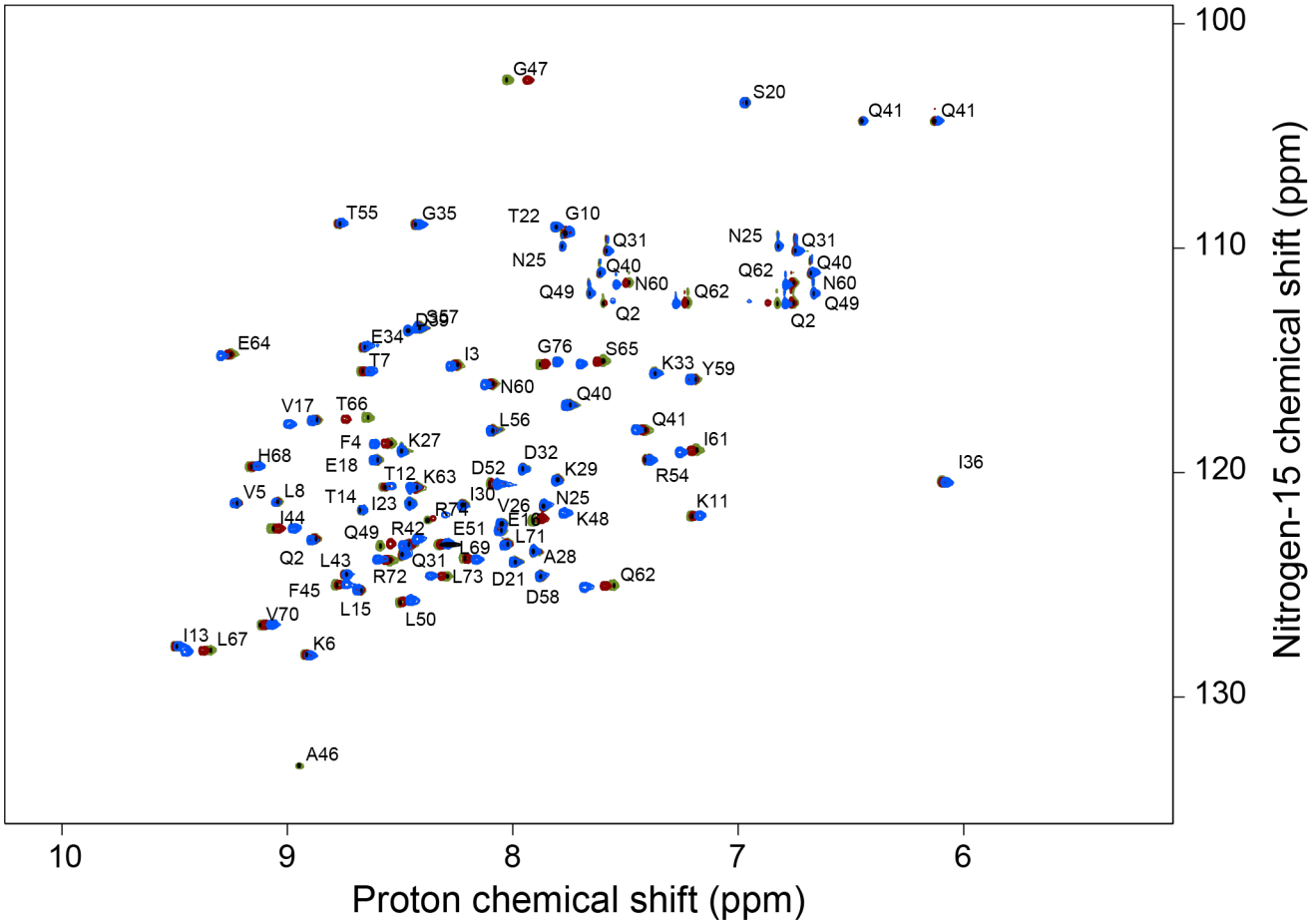



***Figure S 1:*** *$^{1}$H-$^{15}$N HSQC spectrum of the ubiquitin without (green), and with 2 (red) and 10 (blue) equivalents of $[Na]_3[Yb(DPA)_3]$.*

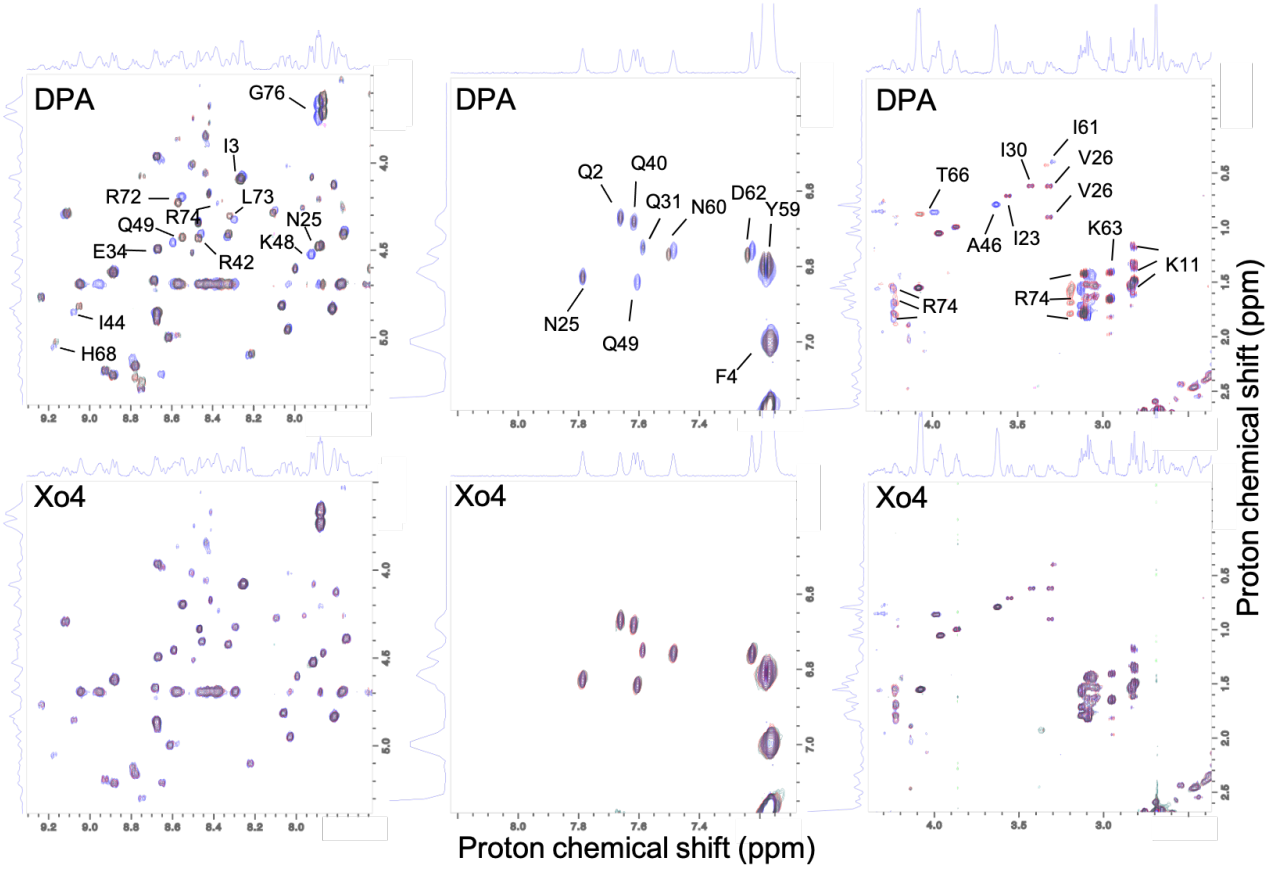



***Figure S 2:*** *Selected regions of $^{1}H$-$^{1}H$ TOCSY spectra recorded on ubiquitin sample. Upper panel: without (blue), and with 0.8 (red) 2 (green) equivalents of $[Yb(DPA)_3]^{3-}$. Lower panel: Spectra without (blue) , and with 0.8 (red) and 2 (green) equivalents of $[YbXo4]^{+}$. Residues whose correlation undergoes a significant shift upon addition of $[Yb(DPA)_3]^{3-}$ are labeled.*

***Titration Curves***

For a system undergoing rapid exchange between a free and bound state, the chemical shifts observed are an average of both different nuclear spin populations and follow the equation:

$$\delta_{obs} = \delta_{free} + \left(\delta_{bound} - \delta_{free}\right)\frac{K_d+P_T+nP_T-\sqrt{-4nP_T{}^2+(\ [\![K_d+P_T+nP_T)]\!]\ ^2}}{2P_T} \qquad \text{(S1)}$$

where $\delta_{obs}$ is the observed chemical shift, $\delta_{free}$ is the free state chemical shift, $\delta_{bound}$ is the chemical shift of the bound state, $K_d$ is the equilibrium dissociation constant, PT is the protein concentration, and n is the number of equivalents ($n_{eq}$= $[Ln(DPA)_3{}^{3-}]/[ubiquitin]_{initial}$). This equation was used to fit the experimental data and extract an estimate of $K_d$ and $\delta_{bound}$. The quality of the extracted values was then evaluated by Monte Carlo simulation to estimate the uncertainty of the data fitting (Figure S3).

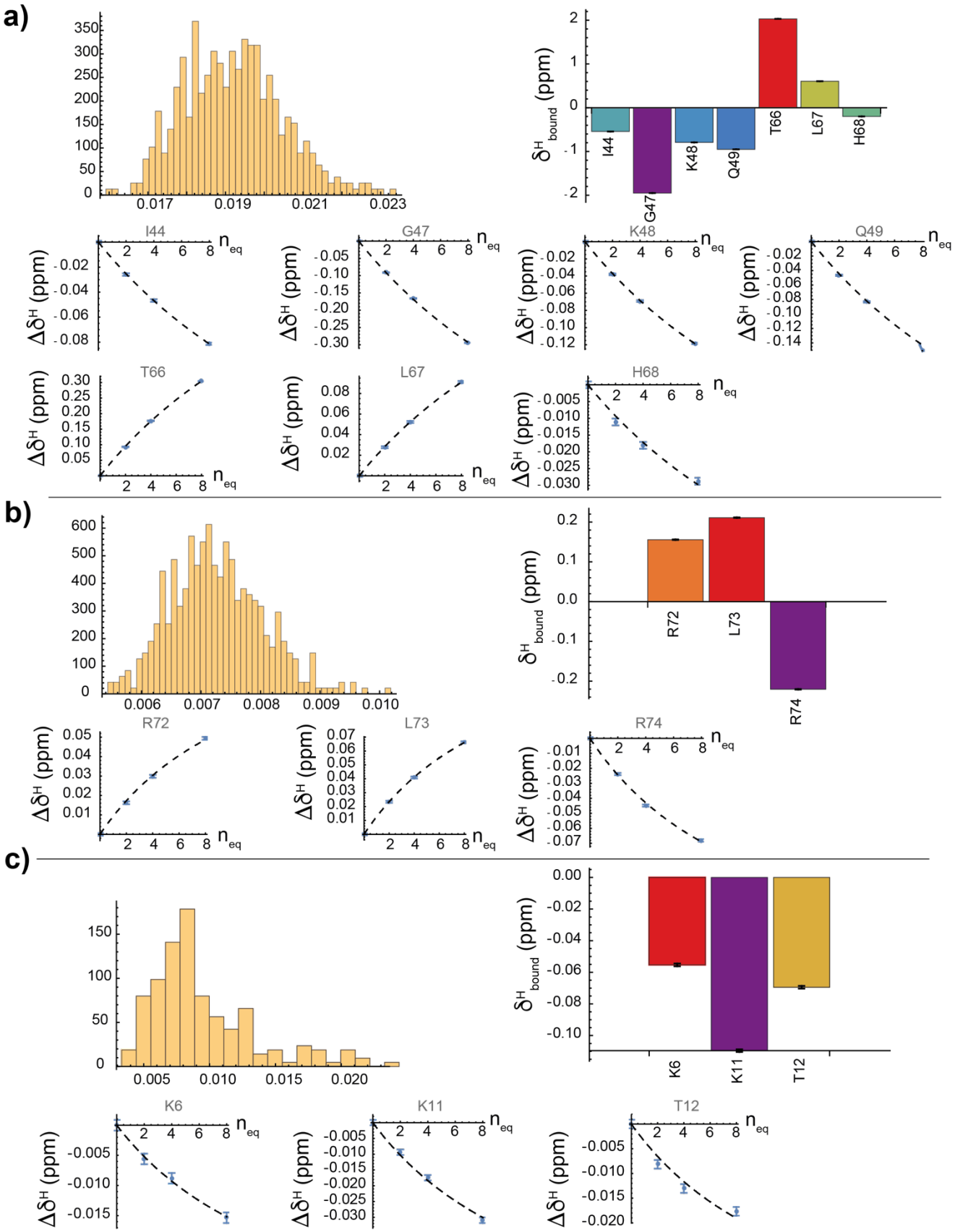



***Figure S 3.*** *The $K_d$ distributions resulting from the Monte Carlo analysis of the NMR titration experiment performed to monitor the interaction between ubiquitin and $[Yb(DPA)_3]^{3-}$ are shown together with the values of the $^1H$ chemical shifts $\delta_{bound}$ that were adjusted simultaneously for the selected residues that have been identified by MD simulation as belonging to each interaction site. (a) Residues that could form part of the H68 binding site. (b) Residues that could form part of the R42-R74 binding site. (c) Residues*

*that could form part of the third binding site. The resulting best fits are shown for the titration curve of each selected residue.*

***Interaction between ubiquitin and [YbXo4]⁺***

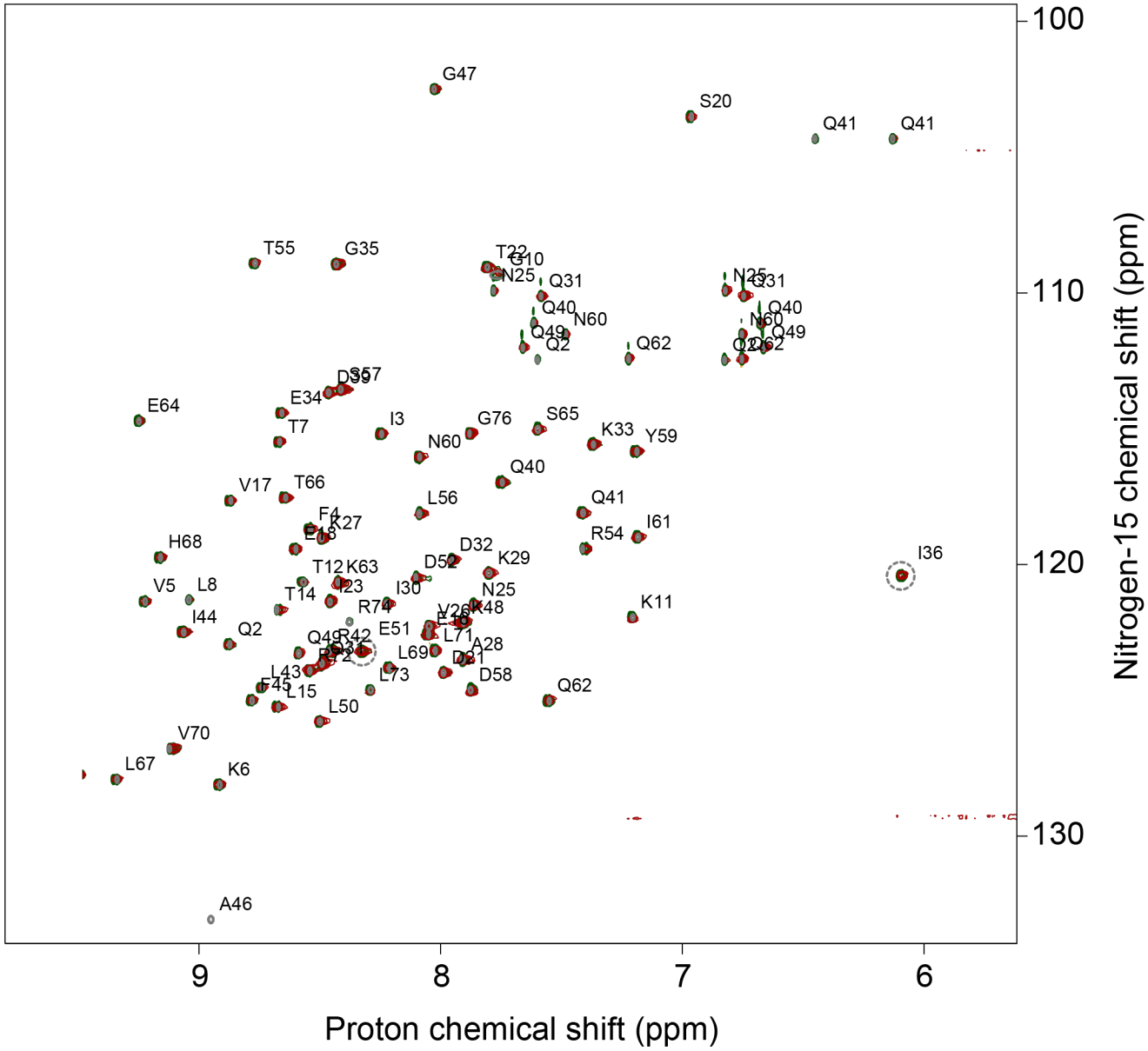



***Figure S 4.*** *$^{1}$H-$^{15}$N HSQC spectrum of ubiquitin without (green), and with 4 equivalents (red) of [YbXo4]⁺.*

## Molecular dynamics

Our molecular dynamics simulations were performed with our previous MCPB.py-based parametrization of the two lanthanide complexes (reference 17 of the main text), where the lanthanide center was replaced by an yttrium atom, which presents the same coordination and van der Waals parameters. No further electronic description was required to model the binding process and we remain at the classical level of molecular dynamics to describe the diffusion-then-binding process.

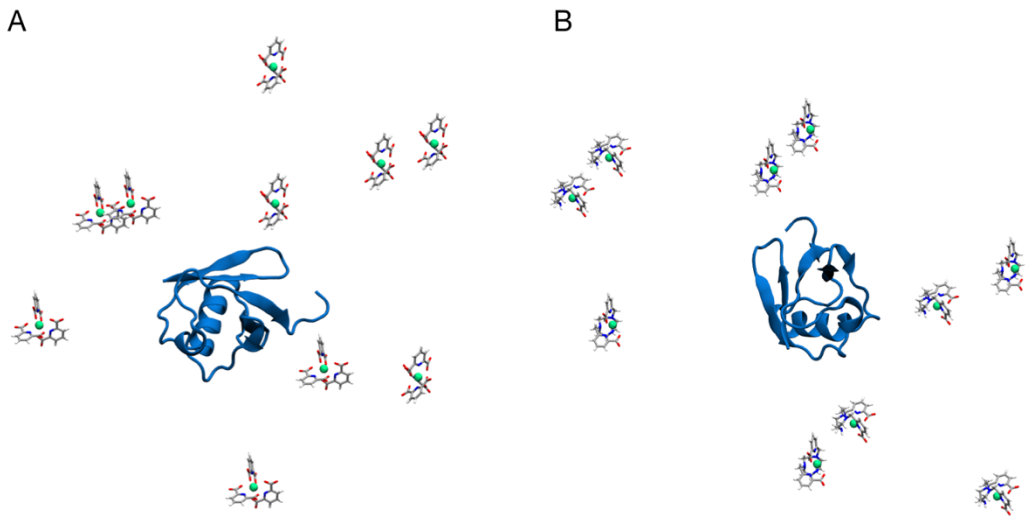



***Figure S 5.*** *Ubiquitin:$[Ln(DPA)_3]^{3-}$ (panel A) and Ubiquitin:$[Ln\text{-}Xo4]^{+}$ (panel B) initial systems. The protein is reported in blue cartoon, while ligands in sticks and Ln central atom as green sphere.*

***Density maps***

2D-density maps of $[Ln(DPA)_3]^{3-}$ and $[Ln\text{-}Xo4]^{+}$ ligands (both enantiomers), around the protein, were obtained by using *gmx densmap* tool of *Gromacs 2021*[7] suite of programs. *cpptraj* and *parmed* programs of *AmberTools18* were used to convert the amber trajectories and the topologies into Gromacs readable formats (gro, top and xtc), respectively. The trajectories were fitted into a reference structure (Figure S 6) using the protein backbone heavy atoms (CA, C, N and O). 2D maps were computed for the "Upper" and "Lower" protein regions (Figure S 6 and Figure S 7) defined as the volume slices included between the protein center of mass and the highest and lowest residue along the z-axis, respectively. 3D-density maps of $[Ln(DPA)_3]^{3-}$ and $[Ln\text{-}Xo4]^{+}$ ligands (both Δ and Λ enantiomers), around the protein, were computed using the *volmap* tool of VMD [8] (Figure 5). 1 Å was used as volume cube side length (-res option).

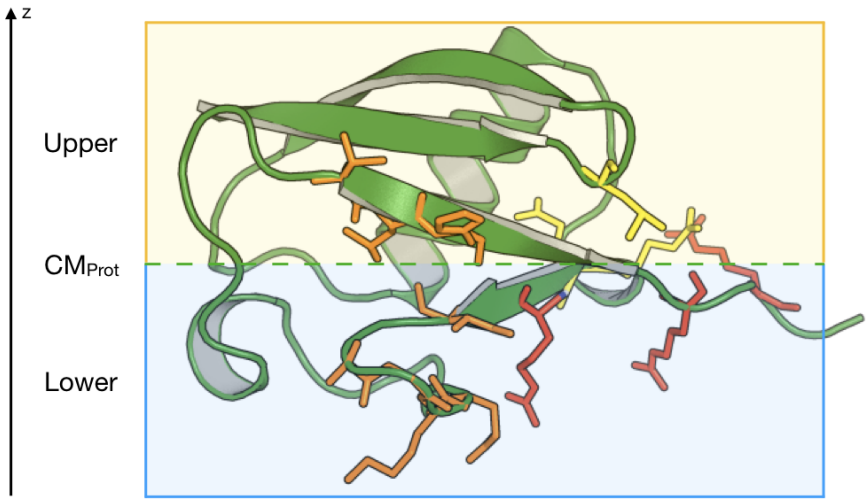



***Figure S 6.*** *Ubiquitin protein is represented as green cartoon, while H68, R42–R74 and Q41 binding site residues as orange, red and yellow lines, respectively. "Upper" and "Lower" regions are defined as the protein regions, along the z axis perpendicular to the protein, between the protein center of mass ($CM_{Prot}$, dashed green line) and the highest and lowest residue along the z axis (light orange and light blue rectangles), respectively.*

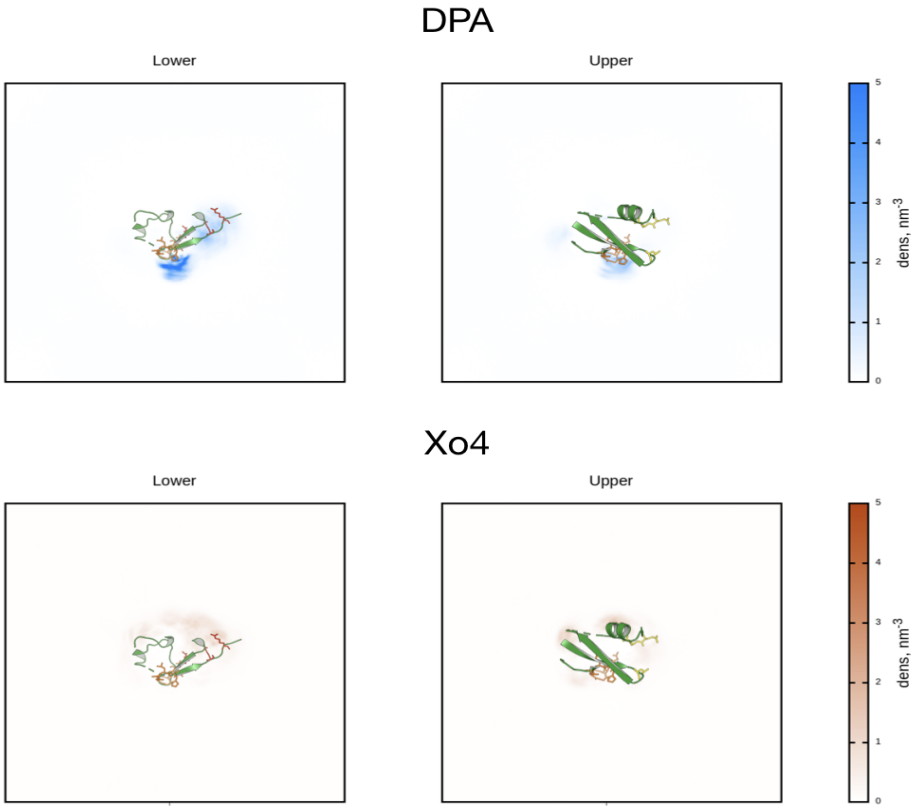



***Figure S 7.*** *2D densities, along the $\vec{z}$ direction (see Figure 2), of $[Ln(DPA)_3]^{3-}$ (DPA, upper panels) and [Ln-Xo4]$^{+}$ (Xo4, lower panels) molecules around the protein in the "Lower" and "Upper" protein regions.*

***Contact probabilities***

The contact probability per–residue was expressed as the probability of the ligand (lig) to be in contact with a specific protein residue for a specific amount of time $\Delta t_{res}$ and considering the whole trajectory time lapse. It is reported in Eq. S2:

$$\rho_{con}^{res} = \frac{1}{t_{tot}} \sum_{i=1}^{N_{lig}} \Delta t_{res}$$

$$Prob_{cont} = \frac{\rho_{con}^{res}}{\sum_{i=1}^{N_{lig}} \rho_{con}^{i}} \quad \text{(S2)}$$

where $\Delta t_{res}$ is the time spent by the ligand interacting within a distance cutoff from residue res and $t_{tot}$ is the total analyzed simulation time (3.8 μs). The cutoff distances were chosen by reflecting the principal non-covalent interaction that could take place between the residue and the ligands (see Table S 3). The ligand is in contact with the residue if at least one of the different interaction patterns or the distance between the residue center of mass and the ligand metal atom Ln atom is less equal to the cutoffs. Distances were calculated using cpptraj. The normalized values are reported in Figure 5A in the main text, considering "ligand–residue" contacts, and in Figure 5B considering "ligand–peptidic -NH groups" ($d_{cutoff}$ =6 Å between the ligand center of mass and -NH).

***Table S 3.*** *Distances used as cutoff ($d_{cutoff}$) for evaluating the contact probability between a protein residue (Res) and the ligand molecule ($[Ln(DPA)_3]^{3-}$ and $[Ln\text{-}Xo4]^{+}$). Interaction defines the "non-covalent" interaction involved within the cutoff distance. For the other residues not specified, a standard 5.00 Å distance cutoff has been used only between the centers of mass of the two interacting partners.*

| | **Ln-Xo4** | | **$[Ln(DPA)_3^{3-}]$** | |
|---|---|---|---|---|
| Residue | **Interaction** | **$d_{cutoff}$ (A)** | **Interaction** | **$d_{cutoff}$ (A)** |
| F | | | π–π stacking | 5.00 |
| Y | | | OH – C $COO^-$ | 5.00 |
| H | π–π stacking, CH–π | 5.00 | | |
| W | | | π–π stacking | 5.00 |
| N | | | | |
| Q | $N_{NH}$–$COO^-$ , $C_{C=O}$–Ln | 4.00 | $N_{NH}$–$COO^-$ | 4.00 |
| D | | | | |
| E | $C_{COO^-}$–Ln | 3.50 | $C_{COO^-}$–Ln | 3.50 |
| R | $C_{guad.}$–$C_{COO^-}$, cation–π | 5.00 | $C_{guad.}$–$C_{COO^-}$ | 5.00 |
| | $N_{NH3^+}$ –$COO^-$ , CH–π | 4.00 | $^{N}{}_{NH3^+}$ –$COO^-$ ,CH–π | 4.00 |
| K | | | | |

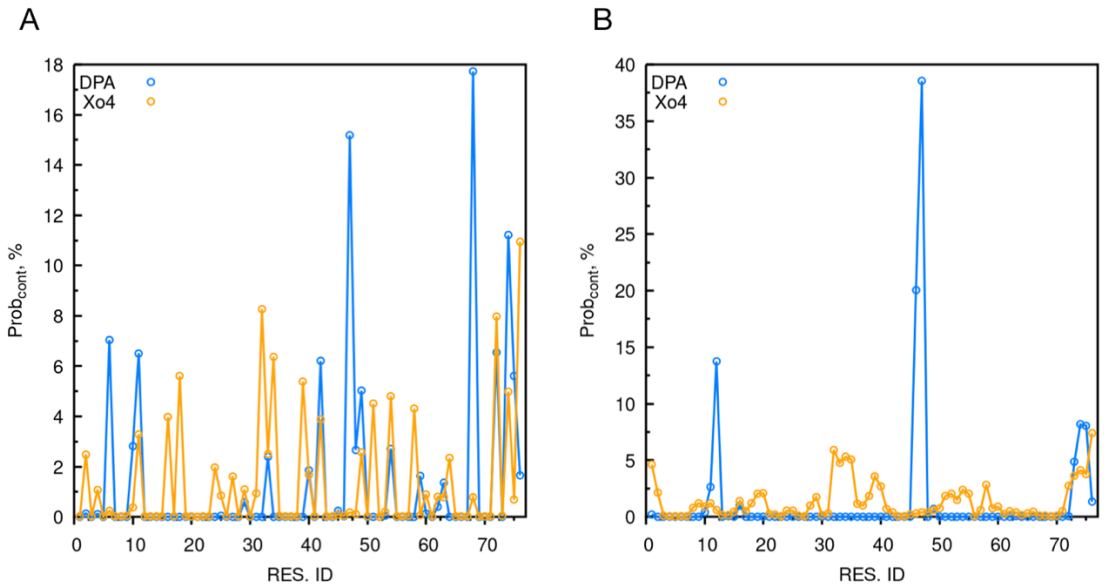



***Figure S 8.*** *Normalized contact probability (Prob$_{cont}$ ) considering ligand–residue (panel A) and ligand-NH peptidic groups of the residue (panel B) contacts from MD simulations. [Ln(DPA)$_3$]$^{3-}$ (DPA) ligand is shown in blue, while [Ln-Xo4]$^+$ (Xo4) in orange.*

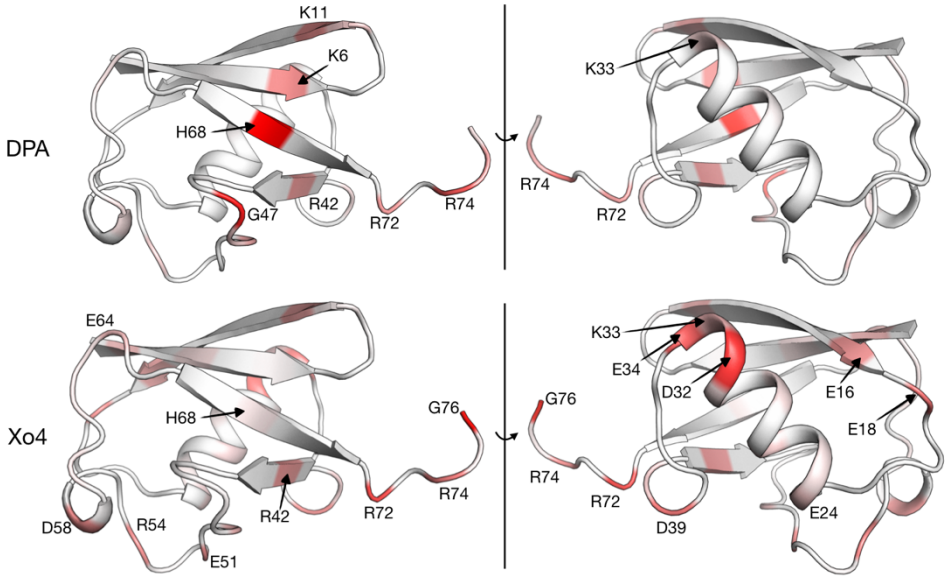



***Figure S 9.*** *Normalized 3D color maps of contact probabilities reported in Figure S8 for [Ln(DPA)$_3$]$^{3-}$ (DPA, upper panels) and [Ln-Xo4]$^+$ (Xo4, lower panels).*

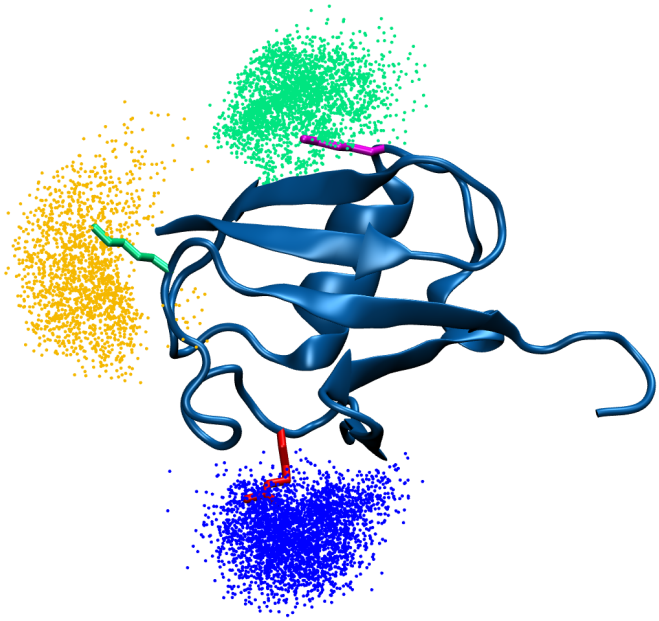


***Figure S 10.*** *Ubiquitin residues that can act as transient interaction sites for $[Ln(DPA)_3]^{3-}$. K33, R54 and K63 are colored as magenta, red and green lines, while the $[Ln(DPA)_3]^{3-}$ ligands, within 0.6 nm of the residue, are represented as points every 50 frames of the analyzed trajectory.*

***Radial Distribution Function and Binding Stoichiometry***

Radial distribution function (g(r)) was calculated using the gromacs *gmx rdf* tool. Ubiquitin center of mass was chosen as reference point, while the distribution was calculated for the metal atom. The average number of metal atoms in function of the distance was also calculated and reported as coordination number η (see Figure S11).

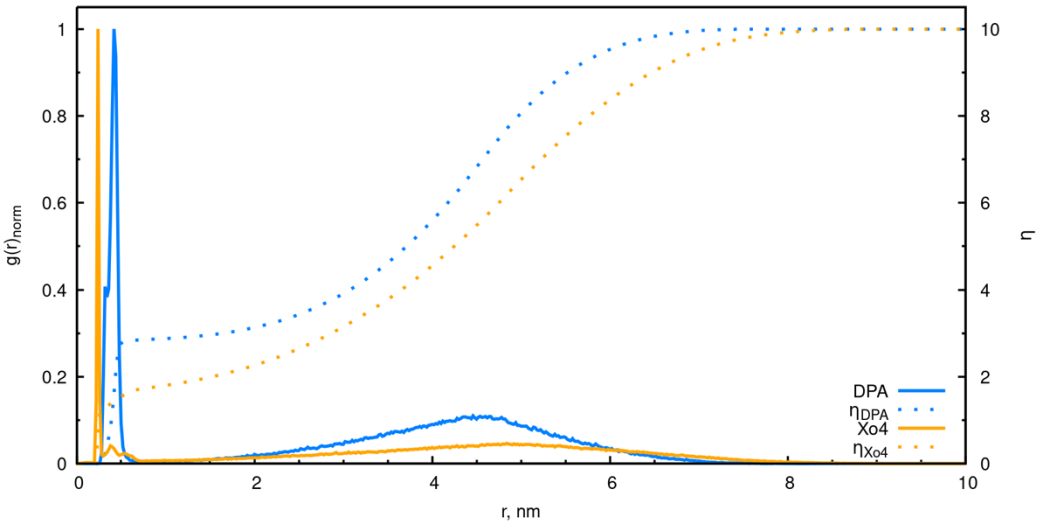



***Figure S 11.*** *Radial distribution function g(r) in function of the distance (in nm) between the protein center of mass and the metal atom of $[Ln(DPA)_3]^{3-}$ (DPA) and $[Ln\text{-}Xo4]^{+}$ (Xo4) complexes. The dashed lines indicate the coordination number η, i.e. the average number of metal atoms within a distance "r".*

The average number of ligands in contact with the protein at the same time was calculated using the *gromacs gmx mindist* tool. To make a protein-metal complex contact, the distance between the two centers of mass must be less equal than the cutoff, chosen as 0.6 nm. At this point, a contact dual-cutoff scheme was used to reduce the statistical noise: a cutoff of 15 contacts ($cutoff_b$) was used to identify the beginning of a binding event, while a cutoff of 10 contacts ($cutoff_{unb}$) to identify ligand unbinding. By post-processing the data considering as block-averaging each replicate, the probability is obtained as the length of the binding event in which the protein is in contact with N ligand(s) divided by the total length of each replicate (190 ns, see main text for details). An average probability is then calculated by averaging all over the 20 replicates, while the error bars correspond to the standard deviation, as reported in Figure S12.

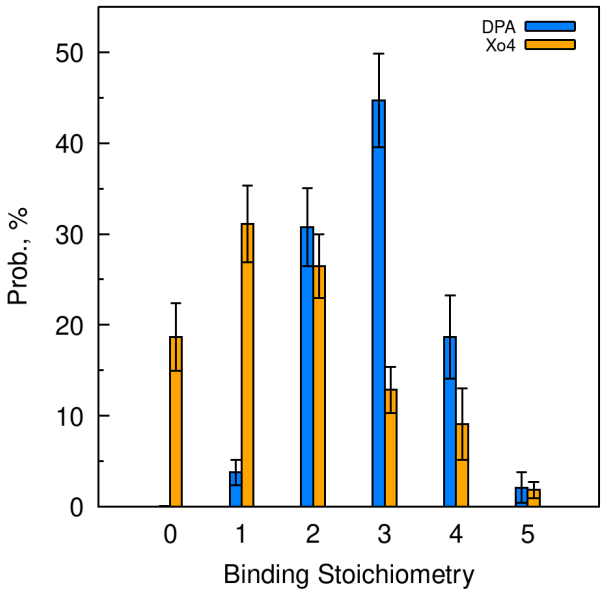



***Figure S 12.*** *Probability of formation of ubiquitin:$[Ln(DPA)_3]^{3-}$ (DPA) or ubiquitin:$[Ln\text{-}Xo4]^{+}$ (Xo4) complex. The binding stoichiometry corresponds to the average number of ligands in contact with the protein at the same time.*

***Cluster Analysis***

To obtain a good qualitative representation of $[Ln(DPA)_3]^{3-}$ and $[Ln\text{-}Xo4]^{+}$ ligands interacting with ubiquitin, cluster analysis has been performed on the whole trajectory obtained by concatenating the last 190 ns of each replicate. Clustering of the MD trajectories was carried out using the hierarchical average-linkage clustering algorithm [9] implemented in the *cpptraj* module of *AmberTools18*. The most representative structures of are shown in Figure 7 and 8 in the main text.

***RMSD Analysis***

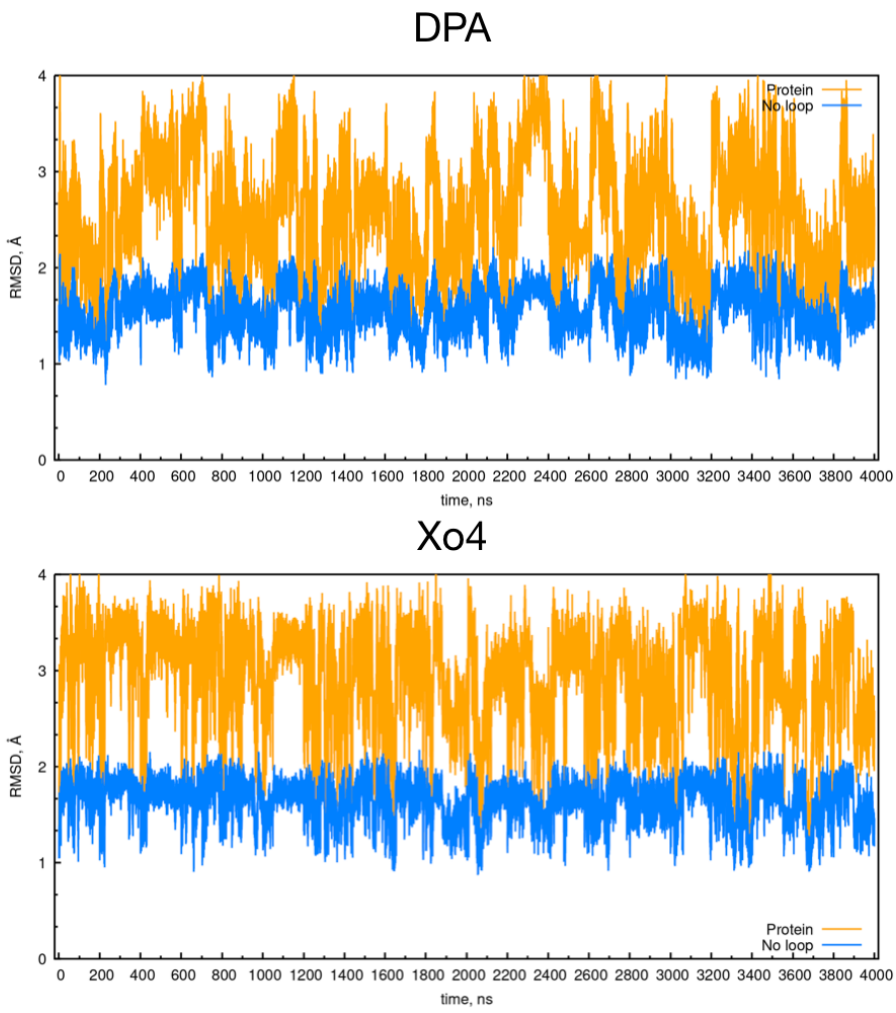



***Figure S 13.*** *Ubiquitin RMSD plots in function of the simulation time for ubiquitin:$[Ln(DPA)_3]^{3-}$ (DPA, upper panel) and ubiquitin:$[Ln\text{-}Xo4]^{+}$ (Xo4, lower panel) complex. As RMSD reference frame, the x-ray protein structure was taken. Orange lines indicate the RMSD behavior for the whole protein backbone atoms, while the blue ones removing the flexible loop (residues 72-76).*

***Force field parameters***

$[Ln\text{-}Xo4]^{+}$ Δ and Λ enantiomers bonded terms, i.e. bond, angle and dihedral equilibrium values and their corresponding force constants, were obtained through the use of the "Metal Center Parameter Builder" (MCPB)[10] python code in AmberTools18. The geometry of $[Ln\text{-}Xo4]^{+}$ ligand was optimized at the DFT-B3LYP/6-31G(d) level of theory, and the atomic charges, assigned on the electrostatic potential (ESP) calculated at the same level of theory, were obtained using the Merz–Kollman restrained electrostatic potential (RESP) scheme [11,12], leading to a partial charge of 1.54e on the Ln atom. All DFT calculations were performed using the Gaussian16 revision B.01 series of programs.[13] $[Ln(DPA)_3]^{3-}$ and $[Ln\text{-}Xo4]^{+}$ ligands were described with GAFF2 [11] parameters, while the protein was described with the AMBER/ff14SB [12] force field. $[Ln(DPA)_3]^{3-}$ and $[Ln\text{-}Xo4]^{+}$ complexes overall charges are -3e and 1e, respectively. This parametrization has been previously tested on various systems, and in close in philosophy to the one proposed by Ramos and coworkers. [14]

***Molecular dynamics protocol***

The systems were initially minimized for 10000 steps (5000 of steepest descent and 5000 of conjugate gradient), and then heated up from 0 K to 300 K for 30 ps in the isothermal–isochoric ensemble (NVT).

Equilibration was carried out for 1 ns, using an integration time step of 2 fs in the isothermal–isobaric (NPT) ensemble (P = 1 atm and T = 300 K). Pressure control in NPT simulations was achieved using the Berendsen barostat. For both systems, NPT replica production runs (20) of 200 ns each and with different initial velocities followed, giving a total simulation time per system of 4 μs (integration time step of 2 fs). During all the MD simulation steps, temperature control was achieved using the Langevin thermostat ($\gamma_{coll}$ = 1 $ps^{-1}$). In the equilibration and production runs, a cutoff of 10 Å was applied for the van der Waals, for electrostatic interactions and for the real space of the electrostatic interaction, while was used of 8 Å during the minimization and of 9 Å during the heating processes. Long-range electrostatic interactions were computed using Particle Mesh Ewald (PME) algorithm. [14,15] The bonds involving hydrogen were treated with SHAKE constraints algorithm. Post-process analysis was performed on the entire trajectory made by concatenating all the replicates, whose first 10 ns were discarded, giving a total of 3.8 μs.

## References


1 S. Engilberge, F. Riobé, S. Di Pietro, L. Lassalle, N. Coquelle, C.-A. Arnaud, D. Pitrat, J.-C. Mulatier, D. Madern, C. Breyton, O. Maury and E. Girard, *Chem. Sci.*, 2017, **8**, 5909–5917.

2 C. A. Smith, A. Mazur, A. K. Rout, S. Becker, D. Lee, B. L. de Groot and C. Griesinger, *J Biomol NMR*, 2020, **74**, 27–43.

3 W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, *Proteins*, 2005, **59**, 687–696.

4 S. P. Skinner, R. H. Fogh, W. Boucher, T. J. Ragan, L. G. Mureddu and G. W. Vuister, *J Biomol NMR*, 2016, **66**, 111–124.

5 P. L. Weber, S. C. Brown and L. Mueller, *Biochemistry*, 1987, **26**, 7282–7290.

6 G. Cornilescu, J. L. Marquardt, M. Ottiger and A. Bax, *J. Am. Chem. Soc.*, 1998, **120**, 6836–6837.

7 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.

8 W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38.

9 J. Shao, S. W. Tanner, N. Thompson and T. E. Cheatham, *J. Chem. Theory Comput.*, 2007, **3**, 2312–2334.

10 P. Li and K. M. Merz, *J. Chem. Inf. Model.*, 2016, **56**, 599–604.

11 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.

12 J. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049–1074.

13 Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; et al., Gaussian 16 Revision B.01 Gaussian Inc., Wallingford CT 2016.

14 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *The Journal of Chemical Physics*, 1995, **103**, 8577–8593.

15 T. Darden, D. York and L. Pedersen, *The Journal of Chemical Physics*, 1993, **98**, 10089–10092.