

## Supporting Information

### Assessment of the components of the electrostatic potential of proteins in solution: Comparing experiment and theory

Chuanying Chen,<sup>†</sup> Binhan Yu,<sup>†</sup> Raziye Yousefi, Junji Iwahara, B. Montgomery Pettitt\*

Department of Biochemistry and Molecular Biology, Sealy Center for Structural Biology and  
Molecular Biophysics, University of Texas Medical Branch,  
Galveston, TX 77555

<sup>†</sup> Co-first authors

\* corresponding author: [mpettitt@utmb.edu](mailto:mpettitt@utmb.edu)

## Electrostatics via the Ewald Potential

In simulations the electrostatic potential is not usually defined from a simple superposition of Coulomb potentials due to the cut-off problem. For our MD simulations, a grid potential  $\varphi(\mathbf{x}_p)$  at any grid point,  $\mathbf{p}$ , in space was decomposed into  $\varphi^{real}(\mathbf{x}_p)$  in real space and  $\varphi^{image}(\mathbf{x}_p)$  in Fourier space by using the standard Ewald summation<sup>1-2</sup> under consideration of periodic boundary condition:

$$\varphi(\mathbf{x}_p) = \varphi^{real}(\mathbf{x}_p) + \varphi^{image}(\mathbf{x}_p) + \varphi^{self}(\mathbf{x}_p) \quad [1-1]$$

$$\varphi^{real}(\mathbf{x}_p) = \frac{1}{4\pi\epsilon_0} \sum_{i \neq j}^N q_j \frac{\text{erfc}(\alpha r_{pj})}{r_{pj}} \quad [1-2]$$

$$\varphi^{image}(\mathbf{x}_p) = \frac{1}{V\epsilon_0} \sum_{\mathbf{k} \neq 0} \frac{e^{-k^2/4\alpha}}{k^2} \sum_j^N q_j e^{-i\mathbf{k} \cdot \mathbf{r}_{pj}} \quad [1-3]$$

$$\varphi^{self}(\mathbf{x}_p) = -\frac{\alpha}{4\pi\epsilon_0 \sqrt{\pi}} q_i, \quad [1-4]$$

in which  $\mathbf{x}_p = (x,y,z)$  of a grid point,  $\mathbf{p}$ , in space,  $\varphi^{self}(\mathbf{x}_p)$  is self-interaction term due to the decomposition,  $q_j$  is the charge of an interest atom  $j$ ,  $q_i$  is the charge of an atom at point  $\mathbf{p}$ ,  $r_{pj}$  is the distance between charge  $i$  and  $j$ ; the  $\mathbf{k}$  vectors are given by  $\mathbf{k} = \frac{2\pi}{L} \hat{\mathbf{l}}$ , and  $L$  is the box length.  $\epsilon_0$  is permittivity of vacuum,  $\alpha$  is a screening parameter of  $0.257952 \text{ \AA}^{-1}$ , and we used 1418  $\mathbf{k}$ -vectors for  $\varphi^{image}(\mathbf{x}_p)$ . The Ewald screening parameter was the same as used in the MD simulation and gives a relative error in the energy of no greater than  $10^{-6}$  in the real space sum.

## Structural stability in MD simulations

We monitored the conformational stability of the protein in all simulations to understand the contribution to the electrostatic field. The C-terminal tail is very flexible in all four different ionic strength solutions (Figure S4). In the salt-free solution, the protein core remains stable (State 1, near the x-ray structure) in the first 600 ns, and then switches to State 2 that stays stable after 700 ns. Excluding the C-terminal tail, the averaged structure in State 1 has an RMSD of  $\sim 1.0 \pm 0.1 \text{ \AA}$  from the crystal structure, and in State 2 has a higher RMSD of  $1.6 \pm 0.1 \text{ \AA}$  that is mainly caused by the movement of residues 34-40 in a loop. The results are consistent with  $\sim 1$  millisecond simulation study<sup>3</sup>, in which two or more populated conformational states were identified.

In the presence of KCl salt, the protein is relatively stable during the entire simulation in State 1 (excluding the C-terminal tail), but low populated states were observed with backbone RMSDs of  $\sim 1.6$  Å. Although these low population states are more flexible in residue 34-40, residue 50-54 and/or have slightly shifted or tilted helix, they are able to switch back to State 1. The conformational dynamics of ubiquitin were also observed in experiments<sup>4-9</sup> as well as in other MD simulations on a time scale from a few nanoseconds to milliseconds.<sup>3, 10-12</sup> Thus, the force field chosen apparently provides a reasonable description of the protein.

**Table S1.** Systems set up for the MD simulations on ubiquitin in different KCl salt solutions at pH = 7.5 and temperature of 298.15 K.

	Number of KCl	Number of all atoms	Box size (Å)
0 mM	0	20053	58.110
130 mM	15	19993	58.071
345 mM	40	19893	58.016
745 mM	87	19705	57.910

**Table S2.** PRE  $\Gamma_2$  rates for backbone  $^1\text{H}_\text{N}$  nuclei of ubiquitin at the ionic strength of 345 mM

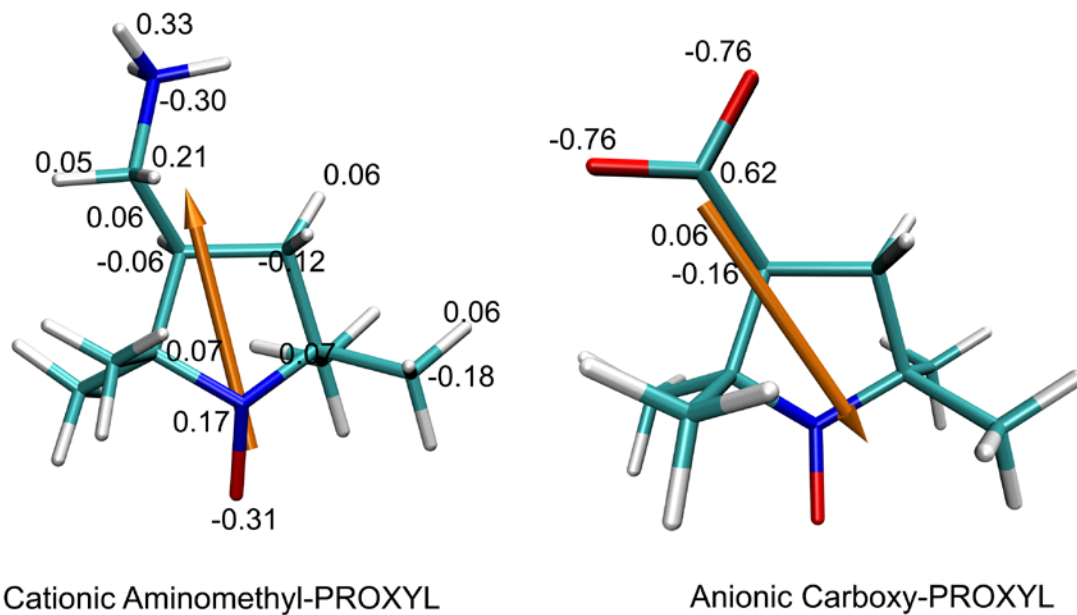
Residue	$\Gamma_{2,+}$ (s $^{-1}$ ) <sup>a</sup>	$\Gamma_{2,-}$ (s $^{-1}$ ) <sup>b</sup>	Residue	$\Gamma_{2,+}$ (s $^{-1}$ ) <sup>a</sup>	$\Gamma_{2,-}$ (s $^{-1}$ ) <sup>b</sup>
Q 2	4.55 ± 0.17	4.47 ± 0.17	Q 41	4.27 ± 0.17	4.65 ± 0.18
I 3	2.95 ± 0.21	3.06 ± 0.21	R 42	5.55 ± 0.19	8.63 ± 0.20
F 4	3.93 ± 0.24	3.83 ± 0.25	L 43	5.14 ± 0.22	7.04 ± 0.23
V 5	4.24 ± 0.21	4.60 ± 0.21	I 44	10.49 ± 0.23	17.92 ± 0.26
K 6	5.81 ± 0.20	8.62 ± 0.21	F 45	10.04 ± 0.22	17.18 ± 0.25
T 7	6.55 ± 0.17	8.56 ± 0.17	A 46	n.d. <sup>3</sup>	n.d. <sup>3</sup>
L 8	73.51 ± 8.33	72.42 ± 6.82	G 47	59.41 ± 1.60	39.87 ± 2.59
T 9	n.d. <sup>1</sup>	n.d. <sup>1</sup>	K 48	14.37 ± 0.16	25.17 ± 0.20
G 10	8.24 ± 0.76	12.78 ± 0.87	Q 49	68.19 ± 1.77	82.83 ± 3.57
K 11	17.89 ± 0.45	12.85 ± 0.40	L 50	6.12 ± 0.21	9.09 ± 0.22
T 12	31.98 ± 3.03	31.47 ± 3.55	E 51	6.23 ± 0.20	7.83 ± 0.21
I 13	4.39 ± 0.22	4.55 ± 0.22	D 52	40.56 ± 0.32	78.08 ± 1.03
T 14	12.21 ± 0.22	19.15 ± 0.27	G 53	n.d. <sup>1</sup>	n.d. <sup>1</sup>
L 15	2.52 ± 0.19	2.69 ± 0.20	R 54	5.68 ± 0.16	3.86 ± 0.16
E 16	20.21 ± 0.21	16.69 ± 0.20	T 55	5.39 ± 0.21	4.97 ± 0.21
V 17	2.95 ± 0.17	2.20 ± 0.17	L 56	1.91 ± 0.19	1.86 ± 0.20
E 18	3.98 ± 0.19	2.91 ± 0.19	S 57	3.14 ± 0.14	2.97 ± 0.14
S 20	4.54 ± 0.18	2.73 ± 0.18	D 58	2.97 ± 0.16	3.44 ± 0.16
D 21	2.98 ± 0.14	1.89 ± 0.14	Y 59	4.41 ± 0.20	5.89 ± 0.20
T 22	7.84 ± 0.17	3.60 ± 0.17	N 60	5.30 ± 0.19	9.67 ± 0.20
I 23	2.71 ± 0.22	2.18 ± 0.22	I 61	4.03 ± 0.20	4.95 ± 0.20
E 24	n.d. <sup>1</sup>	n.d. <sup>1</sup>	Q 62	5.51 ± 0.17	5.90 ± 0.17
N 25	6.87 ± 0.16	3.12 ± 0.16	K 63	46.85 ± 0.58	61.16 ± 1.03
V 26	4.14 ± 0.17	2.48 ± 0.17	E 64	3.53 ± 0.23	3.31 ± 0.23
K 27	3.80 ± 0.18	2.68 ± 0.19	S 65	2.85 ± 0.15	2.74 ± 0.15
A 28	9.08 ± 0.15	5.60 ± 0.15	T 66	48.18 ± 0.54	42.91 ± 0.47
K 29	10.98 ± 0.17	5.91 ± 0.16	L 67	4.14 ± 0.23	6.06 ± 0.24
I 30	4.84 ± 0.19	3.82 ± 0.19	H 68	11.67 ± 0.23	21.00 ± 0.26
Q 31	5.56 ± 0.18	4.46 ± 0.18	L 69	9.13 ± 0.20	16.63 ± 0.22
D 32	9.00 ± 0.14	5.60 ± 0.14	V 70	10.20 ± 0.22	17.41 ± 0.25
K 33	6.83 ± 0.14	4.28 ± 0.14	L 71	46.48 ± 0.41	62.21 ± 0.70
E 34	7.28 ± 0.21	5.00 ± 0.21	R 72	11.13 ± 0.19	16.70 ± 0.22
G 35	15.86 ± 0.24	6.80 ± 0.22	L 73	44.56 ± 1.69	68.46 ± 4.91
I 36	11.92 ± 0.17	10.54 ± 0.17	R 74	n.d. <sup>1</sup>	n.d. <sup>1</sup>
D 39	7.28 ± 0.18	2.38 ± 0.17	G 75	n.d. <sup>1</sup>	n.d. <sup>1</sup>
Q 40	5.18 ± 0.17	5.21 ± 0.17	G 76	12.23 ± 0.15	11.67 ± 0.15

<sup>a</sup>Solvent PRE arising from 25 mM amino-methyl-PROXYL; <sup>b</sup>Solvent PRE arising from 25 mM carboxyl-PROXYL; <sup>1</sup>The signal was too broad; <sup>2</sup>PRE was too large to measure; <sup>3</sup>Error was too large.

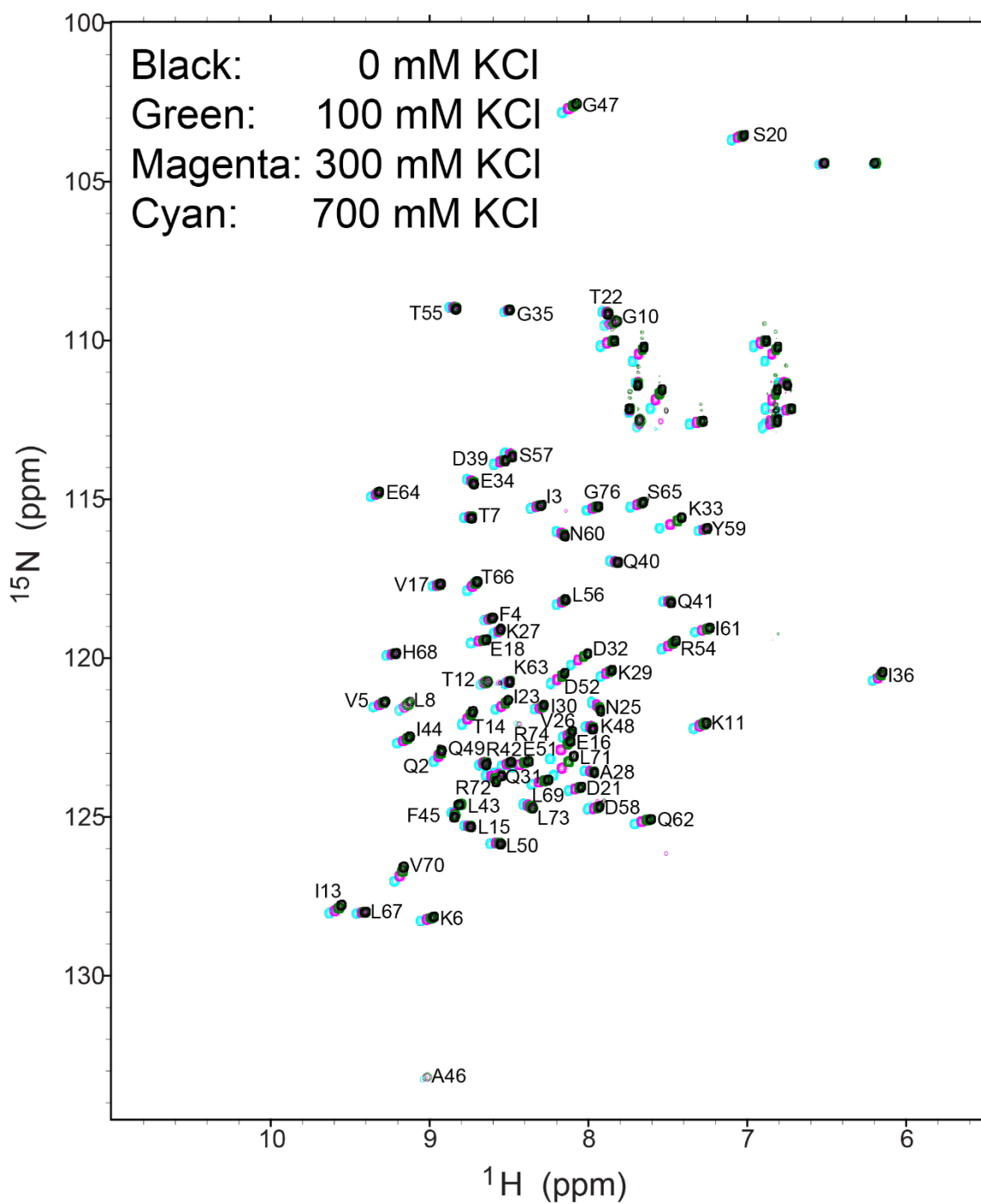
**Table S3.** PRE  $\Gamma_2$  rates for backbone  $^1\text{H}_\text{N}$  nuclei of ubiquitin at the ionic strength of 745 mM

Residue	$\Gamma_{2,+}$ (s $^{-1}$ ) <sup>a</sup>	$\Gamma_{2,-}$ (s $^{-1}$ ) <sup>b</sup>	Residue	$\Gamma_{2,+}$ (s $^{-1}$ ) <sup>a</sup>	$\Gamma_{2,-}$ (s $^{-1}$ ) <sup>b</sup>
Q 2	4.77 ± 0.20	4.11 ± 0.18	Q 41	4.73 ± 0.19	4.47 ± 0.17
I 3	3.21 ± 0.23	3.01 ± 0.21	R 42	7.01 ± 0.21	7.94 ± 0.20
F 4	4.22 ± 0.27	3.62 ± 0.25	L 43	5.99 ± 0.24	6.60 ± 0.23
V 5	4.01 ± 0.23	3.67 ± 0.21	I 44	13.82 ± 0.27	17.00 ± 0.26
K 6	7.07 ± 0.22	8.63 ± 0.21	F 45	13.84 ± 0.25	16.29 ± 0.24
T 7	8.31 ± 0.19	8.55 ± 0.18	A 46	n.d. <sup>3</sup>	n.d. <sup>3</sup>
L 8	n.d. <sup>3</sup>	n.d. <sup>3</sup>	G 47	95.55 ± 6.18	121.23 ± 9.11
T 9	n.d. <sup>1</sup>	n.d. <sup>1</sup>	K 48	18.68 ± 0.21	23.62 ± 0.20
G 10	8.78 ± 0.80	11.12 ± 0.82	Q 49	49.95 ± 2.10	33.54 ± 3.88
K 11	24.01 ± 0.61	13.72 ± 0.40	L 50	7.28 ± 0.23	8.23 ± 0.22
T 12	38.92 ± 3.95	41.29 ± 3.79	E 51	6.88 ± 0.22	7.60 ± 0.21
I 13	4.68 ± 0.24	4.09 ± 0.22	D 52	53.26 ± 0.66	74.88 ± 1.14
T 14	14.13 ± 0.29	18.42 ± 0.29	G 53	n.d. <sup>1</sup>	n.d. <sup>1</sup>
L 15	3.24 ± 0.21	2.57 ± 0.20	R 54	5.84 ± 0.17	4.21 ± 0.15
E 16	21.45 ± 0.28	18.80 ± 0.22	T 55	5.40 ± 0.22	4.87 ± 0.21
V 17	2.11 ± 0.19	1.46 ± 0.17	L 56	2.04 ± 0.20	1.72 ± 0.19
E 18	3.91 ± 0.20	2.98 ± 0.18	S 57	2.61 ± 0.15	2.46 ± 0.14
S 20	4.06 ± 0.19	2.77 ± 0.17	D 58	3.12 ± 0.17	3.21 ± 0.15
D 21	2.74 ± 0.15	1.71 ± 0.14	Y 59	4.38 ± 0.21	5.58 ± 0.20
T 22	7.28 ± 0.19	3.85 ± 0.17	N 60	6.18 ± 0.22	10.40 ± 0.22
I 23	2.27 ± 0.23	1.64 ± 0.22	I 61	3.73 ± 0.21	4.86 ± 0.20
E 24	n.d. <sup>1</sup>	n.d. <sup>1</sup>	Q 62	7.78 ± 0.18	5.61 ± 0.17
N 25	6.11 ± 0.18	2.99 ± 0.16	K 63	59.89 ± 1.25	74.75 ± 1.63
V 26	3.96 ± 0.18	2.38 ± 0.17	E 64	3.70 ± 0.25	3.29 ± 0.23
K 27	3.60 ± 0.20	2.49 ± 0.19	S 65	3.36 ± 0.16	2.32 ± 0.15
A 28	8.95 ± 0.16	5.31 ± 0.15	T 66	57.05 ± 1.00	44.77 ± 0.57
K 29	10.45 ± 0.18	6.74 ± 0.16	L 67	5.55 ± 0.25	5.98 ± 0.24
I 30	4.43 ± 0.20	3.69 ± 0.19	H 68	15.23 ± 0.27	19.96 ± 0.27
Q 31	5.59 ± 0.18	4.41 ± 0.17	L 69	12.18 ± 0.24	16.31 ± 0.23
D 32	8.41 ± 0.15	6.08 ± 0.14	V 70	13.14 ± 0.26	16.69 ± 0.25
K 33	6.47 ± 0.14	4.28 ± 0.13	L 71	61.70 ± 0.92	63.81 ± 0.84
E 34	7.42 ± 0.22	5.46 ± 0.21	R 72	16.17 ± 0.24	17.52 ± 0.22
G 35	14.60 ± 0.26	6.52 ± 0.21	L 73	54.58 ± 2.84	75.64 ± 4.60
I 36	12.29 ± 0.18	10.85 ± 0.17	R 74	n.d. <sup>1</sup>	n.d. <sup>1</sup>
D 39	6.11 ± 0.22	2.35 ± 0.18	G 75	n.d. <sup>1</sup>	n.d. <sup>1</sup>
Q 40	5.52 ± 0.19	4.93 ± 0.17	G 76	12.72 ± 0.20	10.85 ± 0.17

<sup>a</sup>Solvent PRE arising from 25 mM amino-methyl-PROXYL; <sup>b</sup>Solvent PRE arising from 25 mM carboxyl-PROXYL; <sup>1</sup>The signal was too broad; <sup>2</sup>PRE was too large to measure; <sup>3</sup>Error was too large.

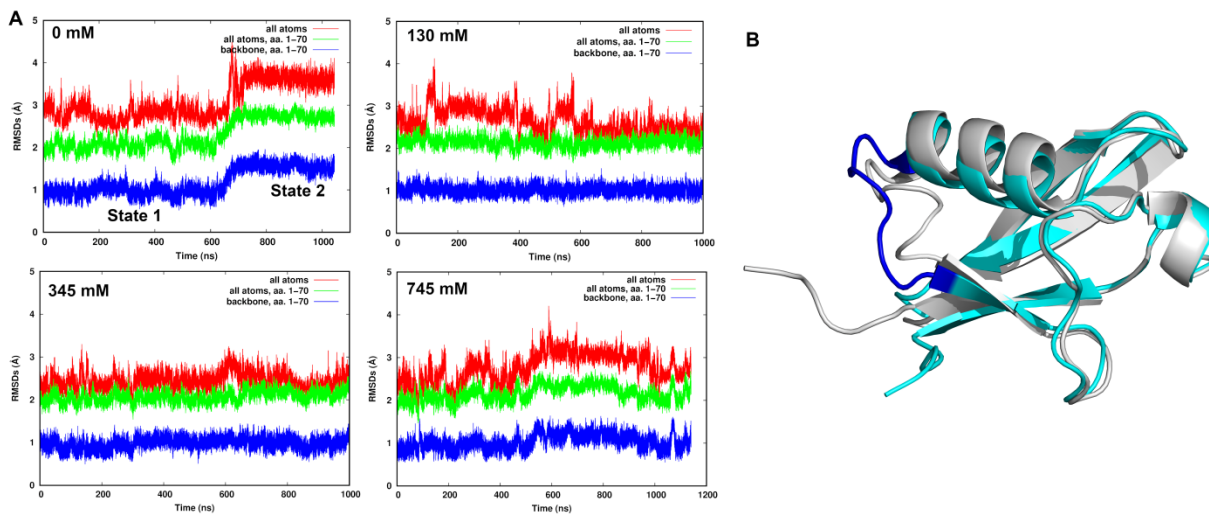


**Figure S1.** The partial charges of amino-methyl-PROXYL and carboxy-PROXYL. The arrow in orange demonstrates the direction of the dipole of each probe molecule.

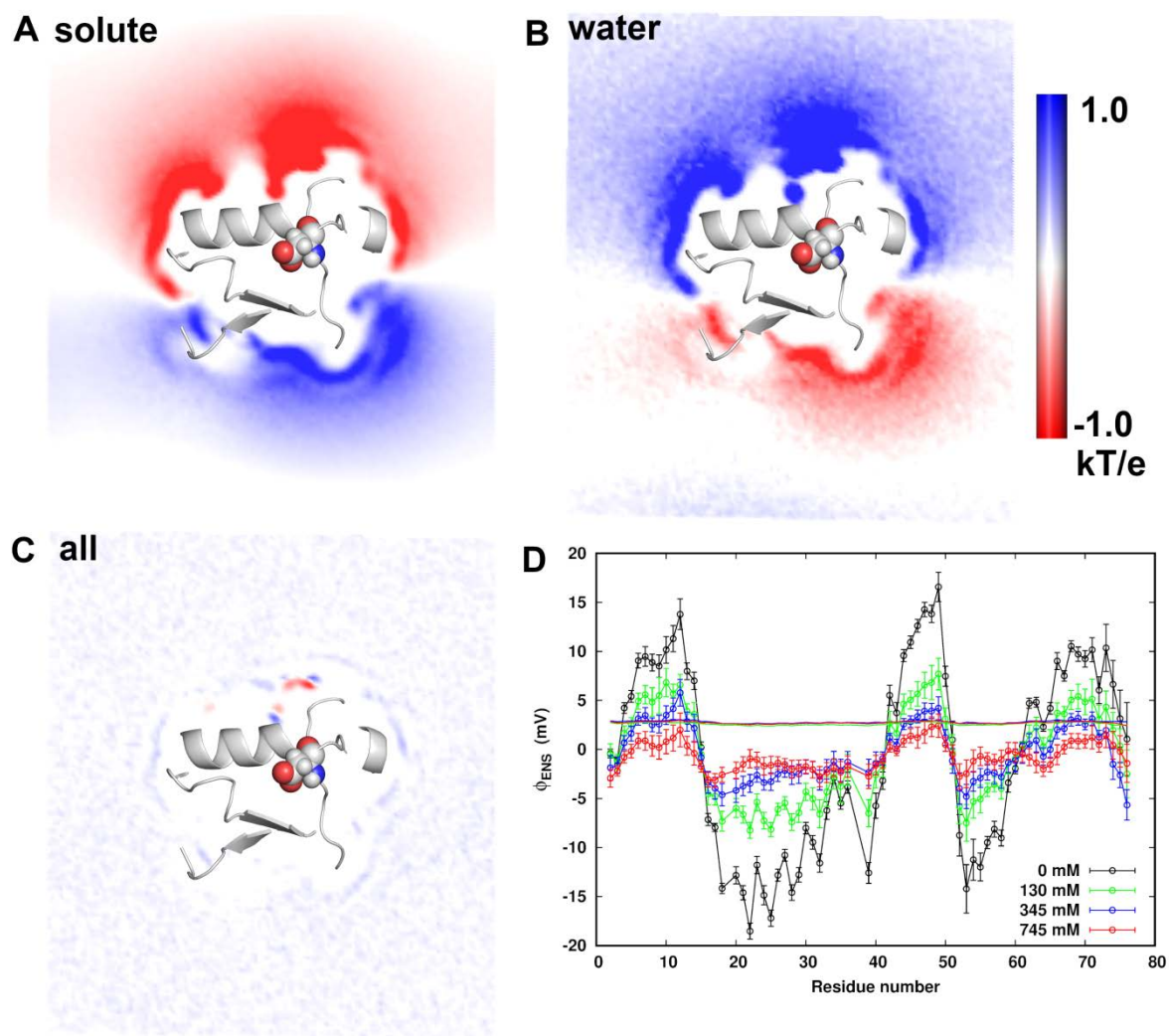


**Figure S2.** Overlay of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of ubiquitin at different KCl concentrations.

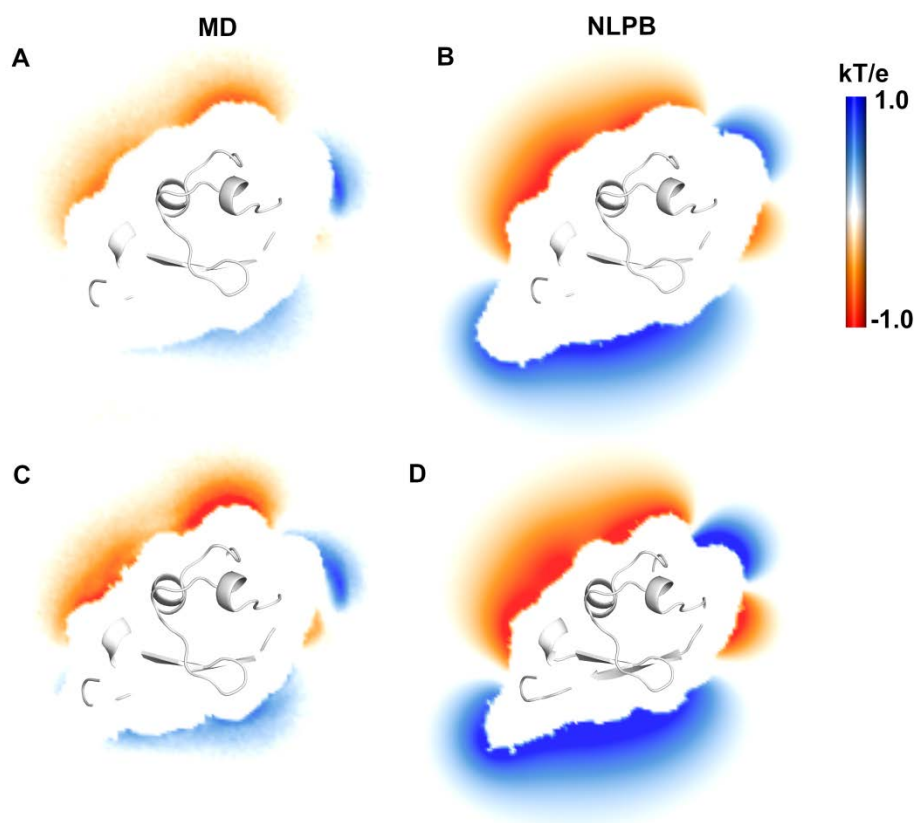




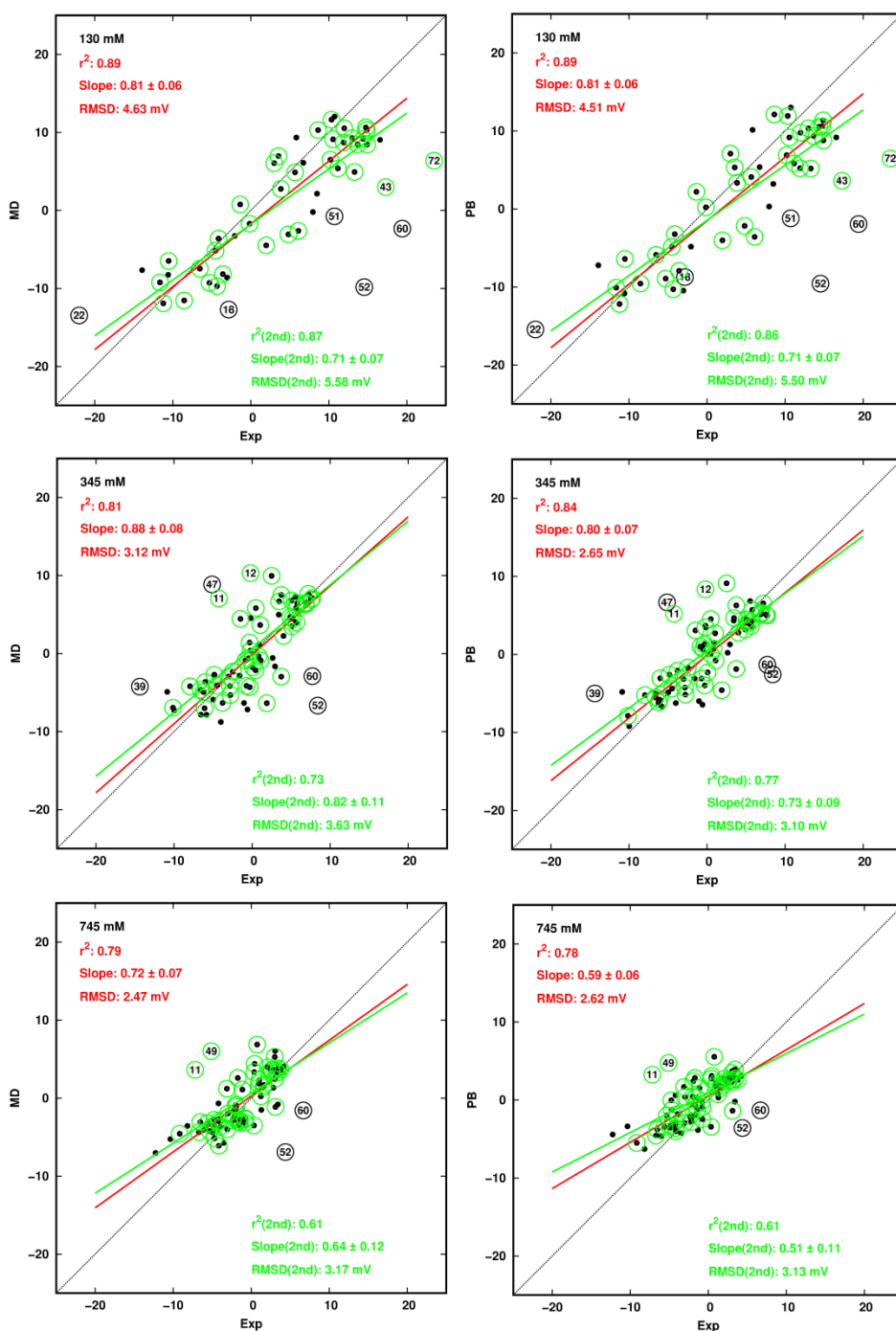
**Figure S3.** (A) RMSDs from the crystal structure of ubiquitin in different KCl salt solutions. (B) Superimposition of State 1 (in grey) and State 2 (in cyan) of ubiquitin excluding the C-terminal tail in the salt-free solution. Highlighted in blue is residue 31-41 involving major shifts from State 1.



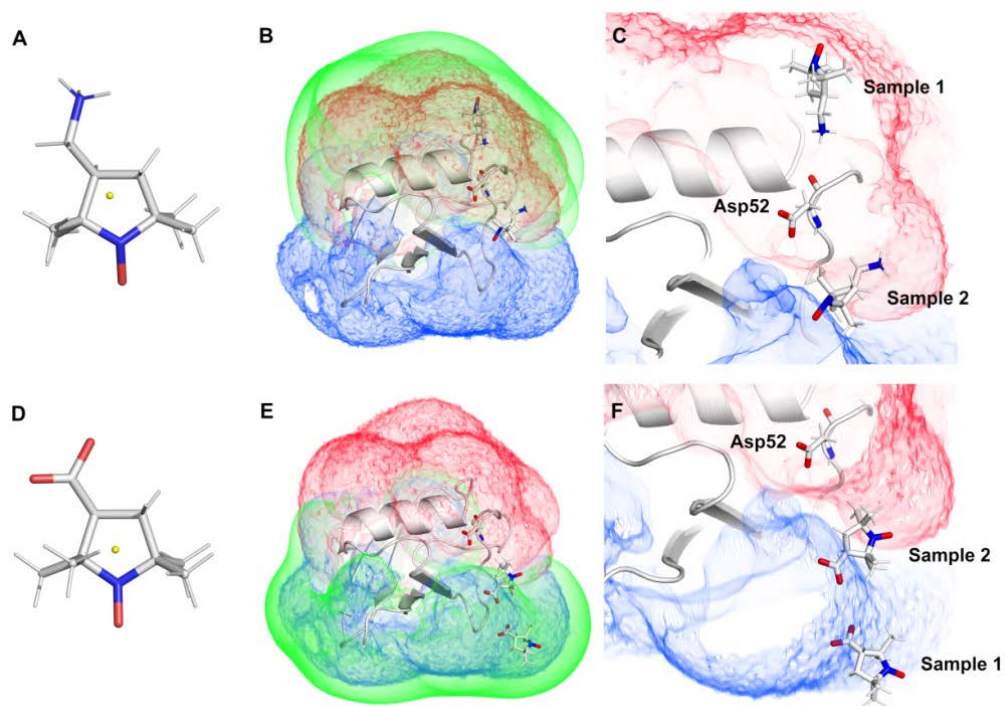
**Figure S4.** Slices of electrostatic potential component map around Asp52 (represented in vdW sphere) of ubiquitin in salt free solution. (A) Component from the solute atoms, (B) water molecules and (C) all atoms in the solution. (D) The MD-derived  $\phi_{ENS}$  potentials contributed from the solute and the mobile ions (with symbols) using a probe radius of 3.5 Å in different salt solutions. The lines at ~ 2.5 mV are for the total potentials contributed from the solute, the mobile ions and water.



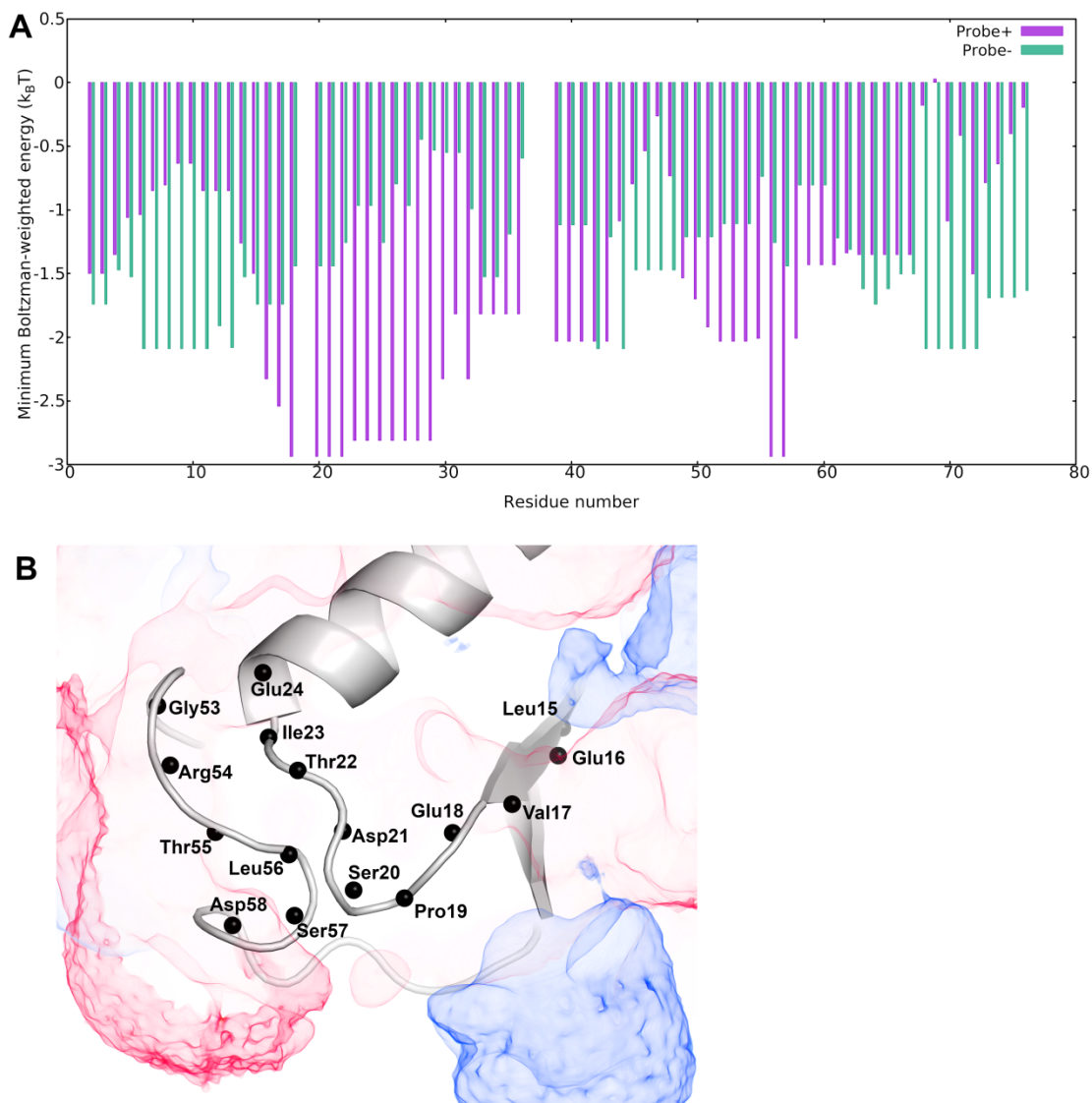
**Figure S5.** (A) and (C) Slices of grid potentials calculated from the MD simulations with the probe radius 3.5 Å and 2.0 Å, respectively. (B) and (D) Slices of grid potentials calculated from the PB-theory with the probe radius 3.5 Å and 2.0 Å, respectively. The potentials inside the probe-excluded-surface (sum of the probe radius 3.5 Å and average vdW radius 1.7 Å from all protein atoms) are set at  $0 k_B T/e$  for a convenience of display.



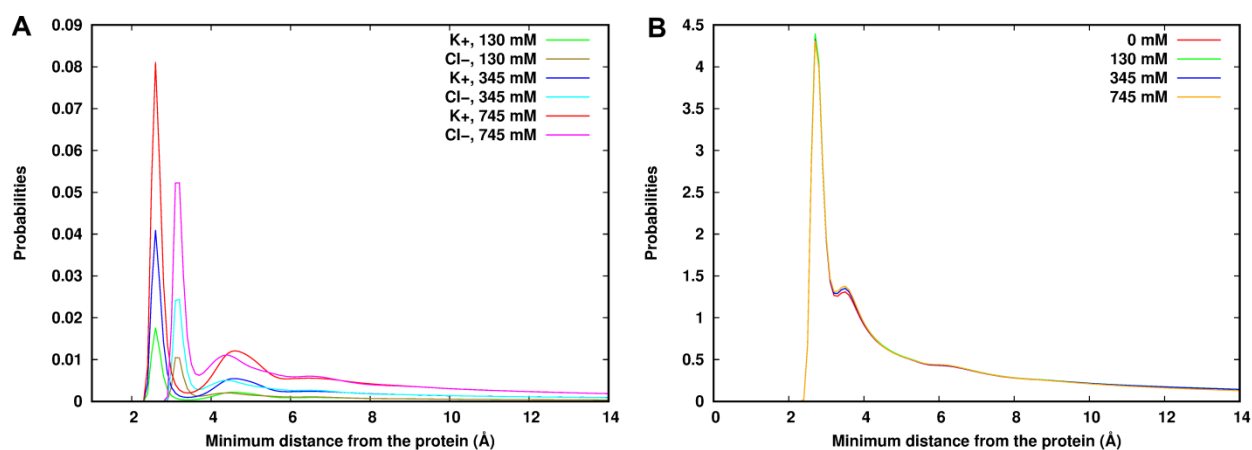
**Figure S6.** Correlation plots for the experimental  $\phi_{ENS}$  potentials and the predicted ones in (A) 130, (B) 345, and (C) 745 mM ionic strength, respectively. The outliers are labeled. The green circles represent the residues in the secondary structure region (2nd). The correlation coefficients of the original data points are  $r^2 = 0.77$  with the slope = 0.61 for 130 mM;  $r^2 = 0.61$  with the slope = 0.64 for 345 mM, and  $r^2 = 0.58$  with the slope = 0.52 for 745 mM, respectively. The black and red solid lines represent linear regression on data points without outliers and the residues in the secondary structure, respectively. The uncertainties in the slopes were estimated at a CI of 95%.



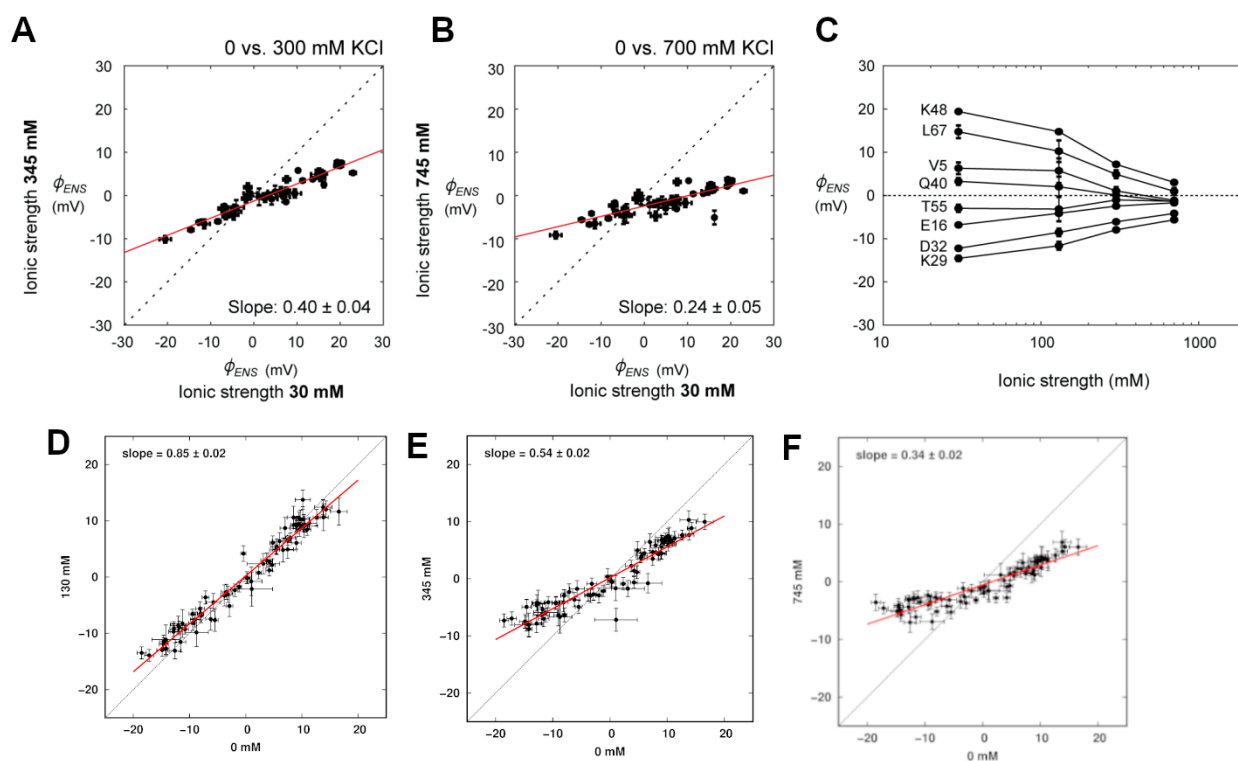
**Figure S7.** The model structures of (A) amino-methyl-PROXYL and (D) carboxy-PROXYL anionic probe, respectively. The yellow dot represents the center of the molecule. The oxygen with an unpaired electron is in a distance of  $\sim 2.9$  Å and  $2.4$  Å from the center of the cationic and anionic probe, respectively. Around ubiquitin in 130 mM KCl are negative potentials ( $-0.3 k_B T/e$  in red) and positive potentials ( $+0.3 k_B T/e$  in blue) rendered in the volume map. In addition, the grid points having favorable Boltzmann-weighted potential energies ( $-0.2 k_B T$ ) of ubiquitin with the cationic probe (B) and with the anionic probe (E) are in green color. Different orientations at the boundary of the positive and negative field around Asp52 are displayed as Sample1 and Sample 2 for cationic (C) and anionic (F) probe, respectively.



**Figure S8.** (A) The minimum effective potential energies in the near surface proximity zone within  $10 \text{ \AA}$  from the  $H_N$  atom of each individual residue of ubiquitin in 130 mM ionic strength. (B) Residue 15-24 and 53-58 of ubiquitin in the electrostatic potential field in 130 mM ionic strength. The backbone N atoms are represented as black spheres. Negative potential surface ( $-0.3 k_B T/e$  in red) and positive potential surface ( $+0.3 k_B T/e$  in blue) are rendered in the volume map.

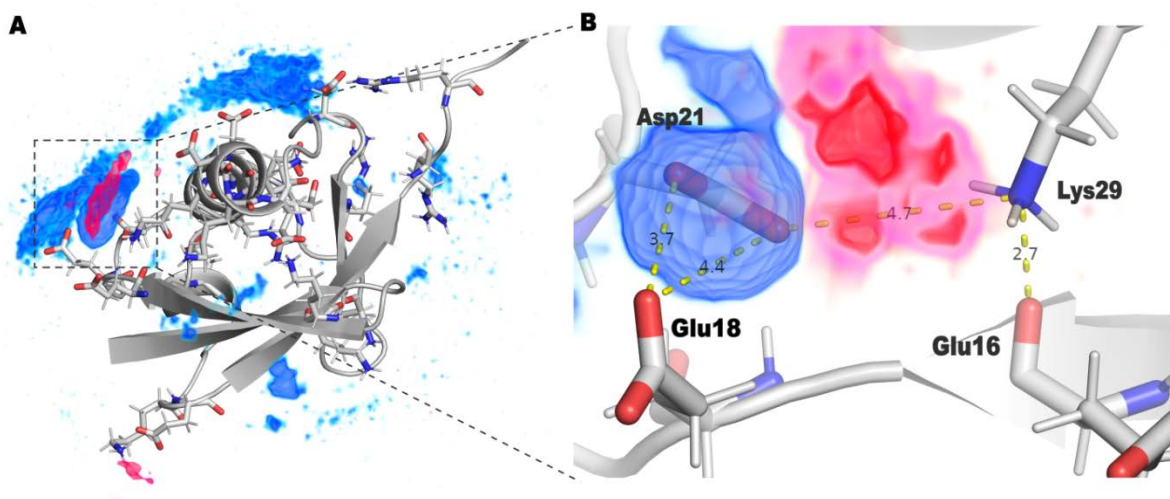


**Figure S9.** The probability of minimum distance of (A) the  $K^+$  ions and  $Cl^-$  ions and (B) water from ubiquitin surface. The highest peak is at 2.6 Å for  $K^+$  ions, 3.2 Å for  $Cl^-$  ions, and 2.7 Å for water molecules.



**Figure S10.** Comparison of  $\phi_{ENS}$  potentials at different ionic strengths. (A) and (B) NMR-derived data; (C) Changes in  $\phi_{ENS}$  potentials as a function of ionic strengths for  $H_N$  atoms of 8 residues. Data at 0 mM and 100 mM KCl (ionic strengths 30 and 130 mM) are from our previous work.<sup>13</sup> (D), (E) and (F) MD-derived data.





**Figure S11.** (A) Localized ion-ion electrostatic interaction energy in 745 mM KCl salt solution. This resulting energy corresponds to the electrostatic energy between an ion, which was located at a grid point, and all the other ions in the unit cell as well as all ions in the periodic image cells. The energies are in a range of  $(-1.0, 1.0) k_B T$  in color from red to blue. (B) Localized ion-ion electrostatic interaction energy around some residues.

## References

1. Ewald, P., The calculation of optical and electrostatic grid potential. *Ann. Phys* **1921**, *64*, 253.
2. de Leeuw, S.; Perram, J.; Smith, E., Simulation of Electrostatic Systems in Periodic Boundary Conditions. I. Lattice Sums and Dielectric Constants. *Proc. Roy. Soc. Lond.* **1980**, *A373*, 27-56.
3. Piana, S.; Lindorff-Larsen, K.; Shaw, D. E., Atomic-level description of ubiquitin folding. *Proc Natl Acad Sci U S A* **2013**, *110* (15), 5915-20.
4. Kitahara, R.; Yokoyama, S.; Akasaka, K., NMR snapshots of a fluctuating protein structure: ubiquitin at 30 bar-3 kbar. *J Mol Biol* **2005**, *347* (2), 277-85.
5. Majumdera, A.; Ghose, R., Probing slow backbone dynamics in proteins using TROSY-based experiments to detect cross-correlated time-modulation of isotropic chemical shifts. *Journal of Biomolecular NMR* **2004**, *28*, 213-227.
6. Massi, F.; Grey, M. J.; Palmer III, A. G., Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1 relaxation experiments. *Protein Science* **2005**, *14*, 735-742.
7. Schneider, D. M.; Dellwo, M. J.; Wand, A. J., Fast internal main-chain dynamics of human ubiquitin. *Biochemistry* **1992**, *31* (14), 3645-52.
8. Lange, O. F.; Lakomek, N. A.; Fares, C.; Schroder, G. F.; Walter, K. F.; Becker, S.; Meiler, J.; Grubmuller, H.; Griesinger, C.; de Groot, B. L., Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **2008**, *320* (5882), 1471-5.
9. Salmon, L.; Bouvignies, G.; Markwick, P.; Lakomek, N.; Showalter, S.; Li, D. W.; Walter, K.; Griesinger, C.; Bruschweiler, R.; Blackledge, M., Protein conformational flexibility from structure-free analysis of NMR dipolar couplings: quantitative and absolute determination of backbone motion in ubiquitin. *Angew Chem Int Ed Engl* **2009**, *48* (23), 4154-7.
10. Braatz, J. A.; Paulsen, M. D.; Ornstein, R. L., 3 Nsec Molecular Dynamics Simulation of the Protein Ubiquitin and Comparison with X-ray Crystal and Solution NMR Structures, *Journal of*



Biomolecular Structure and Dynamics. *Journal of Biomolecular Structure and Dynamics* **1992**, 9 (5), 935-949.

11. Nederveen, A. J.; Bonvin, A. M., NMR Relaxation and Internal Dynamics of Ubiquitin from a 0.2  $\mu$ s MD Simulation. *J. Chem. Theory Comput.* **2005**, 1 (3), 363-374.
12. Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Shaw, D. E., Picosecond to Millisecond Structural Dynamics in Human Ubiquitin. *J Phys Chem B* **2016**, 120 (33), 8313-20.
13. Yu, B.; Pletka, C. C.; Pettitt, B. M.; Iwahara, J., De novo determination of near-surface electrostatic potentials by NMR. *Proc Natl Acad Sci U S A* **2021**, 118 (25).