Electronic Supplementary Material (ESI) for Nanoscale.


**Supplementary information**

# Connecting conformational stiffness of the protein with energy landscape by a single experiment

Soham Chakraborty, Deep Chaudhuri, Dyuti Chaudhuri, Vihan Singh, Souradeep Banerjee, Debojyoti Chowdhury, Shubhasis Haldar*

Department of Biological Sciences, Ashoka University, Sonepat, Haryana 131029

*To whom correspondence may be addressed. Email- Shubhasis.haldar@ashoka.edu.in

**Plasmid construct and protein Expression:**

Polyprotein constructs of protein L and talin were engineered using *Bam*HI, *Bgl*II and *Kpn*I restriction sites in pFN18a expression vector[1,2]. The protein L construct contains eight Protein L B1 domains and talin construct is comprised of a single R3-IVVI domain, followed by eight tandem I91 repeats. Both the constructs are flanked by an N-terminal Halo-Tag enzyme and a C-terminal AviTag for biotinylation. For the purification, a $His_6$-tag is also present immediate before the AviTag sequence[3,4]. For protein expression, *Escherichia coli* BL21 (DE3) competent cell were grown in LB media with 50 µg/ml carbenicillin antibiotics at 37º C until optical density (OD) at 600 nm reached 0.6-0.8. Further, the protein overexpression was carried out by inducing with 1 mM Isopropyl β-D-1-thiogalactopyranoside and, followed by overnight incubation at 25º C.

Detailed sequence of the construct is listed below with colors labelled for corresponding component in the construct. There is a short dipeptide linker between two adjacent components of the construct are underlined.

HaloTag-$(PL)_8$-$His_6$Tag-AviTag

MAEIGTGFPFDPHYVEVLGERMHYVDVGPRDGTPVLFLHGNPTSSYVWRNIIPHVAPTH
RCIAPDLIGMGKSDKPDLGYFFDDHVRFMDAFIEALGLEEVVLVIHDWGSALGFHWAK
RNPERVKGIAFMEFIRPIPTWDEWPEFARETFQAFRTTDVGRKLIIDQNVFIEGTLPMGVV
RPLTEVEMDHYREPFLNPVDREPLWRFPNELPIAGEPANIVALVEEYMDWLHQSPVPKL
LFWGTPGVLIPPAEAARLAKSLPNCKAVDIGPGLNLLQEDNPDLIGSEIARWLSTLEISGA
SEEVTIKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKDNGEWTVDVADKGYTL
NIKFAGASEEVTIKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKDNGEWTVDVA
DKGYTLNIKFAGASEEVTIKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKDNGE
WTVDVADKGYTLNIKFAGASEEVTIKANLIFANGSTQTAEFKGTFEKATSEAYAYADTL
KKDNGEWTVDVADKGYTLNIKFAGASEEVTIKANLIFANGSTQTAEFKGTFEKATSEAY
AYADTLKKDNGEWTVDVADKGYTLNIKFAGASEEVTIKANLIFANGSTQTAEFKGTFEK
ATSEAYAYADTLKKDNGEWTVDVADKGYTLNIKFAGASEEVTIKANLIFANGSTQTAEF
KGTFEKATSEAYAYADTLKKDNGEWTVDVADKGYTLNIKFAGASEEVTIKANLIFANG
STQTAEFKGTFEKATSEAYAYADTLKKDNGEWTVDVADKGYTLNIKFAGHHHHHHGG
GLNDIFEAQKIEWHE

**Protein Purification:**

After inducing the protein overexpression for overnight, the cells were centrifuged at 8000 rpm for 10 minutes and subsequently harvested in 50 mM sodium phosphate buffer pH 7.4 with 300 mM NaCl and 10% Glycerol. Lysozyme and 10µl Phenylmethylsulfonyl fluoride were used as protease inhibitors. The dissolved pellet was incubated in ice for 20 minutes and shaken continuously on rocking platform at 4ºC for 10 minutes. Then Triton-X, DNase, RNase and $MgCl_2$ were added subsequently and again put in rocking platform for 10 minutes at 4º C and the cell were homogenized using French press at 19 psi. The lysate was centrifuged at 4º C and supernatant

is collected. Proteins were purified from lysate using $Ni^{2+}$-NTA affinity chromatography column of ÄKTA™ Pure (GE healthcare). The sodium- phosphate buffer with 20 mM imidazole is used as binding buffer to bind the protein in column which in turn was eluted by using the same buffer with 250 mM imidazole. After purification, the polyprotein was biotinylated by the biotinylation kit (GeneCopoeia). The reaction mixture containing Protein L was kept overnight at 4ºC followed by further purification by size exclusion chromatography with the Superdex-200 HR column of GE healthcare.

**Real time magnetic tweezers methodology:**

Real time magnetic tweezers (RT-MT) was built on an inverted microscope using an oil-immersion objective, attached with a nanofocusing piezo. The chamber was flashed using a collimated white LED light and the diffraction pattern images were collected with ximea camera (MQ013MG-ON). Force is applied using a pair of neodymium magnets to paramagnetic beads (Dynabeads M-270) attached with the protein. A linear voice coil, placed just above the sample chamber, was used to control the position of the magnets. We use multifunctional DAQ card for controlling the voice coil and piezo actuator and for getting accurate data. Detail information regarding the force calibration has been described previously by Chakraborty et al.[5]

The experiments are performed with glass chamber made by two sandwiched glass slides with the advantage of inlets and outlets. At first, glass surfaces were cleaned by Hellmanex and ethanol. Then they are silanized by incubating for 20 minutes in 1% solution of (3-Aminopropyl) trimethoxysilane (sigma Aldrich, 281778) in ethanol. The fluid chambers were then functionalized with glutaraldehyde (Sigma Aldrich, G7776) and HaloTag amine ($O_4$) ligand (Promega, P6741). The non-magnetic beads (2.5-2.9 µm, Spherotech, AP-25-10) were added as a reference bead. Finally, to block nonspecific interactions, the chambers were washed with freshly prepared blocking buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 1% BSA, 0.03 % $NaN_3$) and left them to passivate at least for 5 hours at room temperature. Single molecule experiments were carried out with 1-10 nM protein L solution in 20 mM Tris buffer with varying ammonium sulfate concentration[6]. Briefly, we separately prepared two buffers: 1) 20 mM Tris buffer, containing only 10 mM ascorbic acid of pH 7.2 and 2) 20 mM Tris buffer containing 10 mM ascorbic acid, supplemented with 1M ammonium sulfate at pH 7.2. Then we adjusted the ammonium sulfate concentration by mixing the buffer 1 and 2, accordingly. By applying different forces, we measured the step size or extension that described the unfolding and refolding of single protein domains. Gaussian fits were used to measure the absolute position and the required data corresponding to each step was obtained by length histograms. The differences in length between the centers of two peaks measured the step size.[3,7,8]

**Image Processing:**

One of the most critical functions of magnetic tweezers is the determination of the Z position of the two beads, non-magnetic and paramagnetic. Non-magnetic beads were used as reference bead,

stuck to the surface, to compensate the focal drift and instrumental vibration of the paramagnetic bead. To calculate the real time position of the bead, we used central processing unit (CPU) of the computer. We overclocked the CPU and enhance the image processing algorithm to get the satisfactory result.

There are several steps to determine the Z position of the bead:

1. Images are projected on a camera through objective that allows real-time screening of image connected to the computer and acquiring the region of interest (ROI).
2. Fourier transform (FFT) of the bead from the two ROI was acquired.
3. For determining the position of each bead, we calculate the radial profile and correlate to a Z-stack of the library and correlation profile was fitted.
4. The diffraction pattern around the bead can compute the change in end-to-end length of our protein of interest. We correlate measured radial vector and the stack at different position, for determining the z-position of the beads[3,9].

**Folding probability calculations:**

We represented the tables with residence time of each state at 10 pN force of figure 1 and calculated the folding probability using the eq. 1 mentioned in the result section.

| **Force = 10 pN** | | | | | |
|---|---|---|---|---|---|
| **150 mM $(NH4)_2SO_4$** | | | **500 mM $(NH4)_2SO_4$** | | |
| **State I** | **t** | **It/N** | **State I** | **t** | **It/N** |
| 8 | 0.000 | 0.000 | 8 | 0.000 | 0.000 |
| 7 | 0.000 | 0.000 | 7 | 0.000 | 0.000 |
| 6 | 0.000 | 0.000 | 6 | 0.290 | 0.217 |
| 5 | 0.000 | 0.000 | 5 | 0.136 | 0.085 |
| 4 | 0.000 | 0.000 | 4 | 0.297 | 0.148 |
| 3 | 0.000 | 0.000 | 3 | 0.101 | 0.037 |
| 2 | 0.248 | 0.062 | 2 | 0.089 | 0.022 |
| 1 | 0.287 | 0.035 | 1 | 0.017 | 0.002 |
| 0 | 0.464 | 0.000 | 0 | 0.066 | 0.000 |
| **FP = 0.097** | | | **FP = 0.51** | | |

**Supplementary Table 1:** State occupation at 10 pN force with 150 and 500 mM $(NH4)_2SO_4$ of figure 1

**Method for Molecular Dynamic simulation:**

The X-ray diffraction structure of PLB1 was obtained from PDB ID 1K53[10]. The final 62 residue PLB1 structure was produced in PyMOL 2.5.2 version by deleting the abovementioned 10 res sequence and reverting those two mutations. Molecular simulations of protein L B1 domain (PLB1) were run using Gromacs 2019.4 package using two different concentrations (150 mM and 650 mM) of ammonium sulphate. The protonation states of PLB1 residues were determined by H++ server[11]. For the simulation study, OPLS force field[12] was used to perform the simulations with SPC/E water model[13]. To optimize the topology, we have built GROMACS compatible force for ammonium sulphate. The protein molecule along with solvent water molecules and ammonium ions, sulphate ions were kept in 12 nm Box. The systems were equilibrated at 310 K and 1 bar for 1 ns and the overall simulation for both of the systems were run for 20 ns. Energy minimization were carried out using the steepest descent algorithm[14]. The systems were gradually heated up using verlet thermostat to reach a temperature of 310 K to perform equilibration in NVT ensemble. Later, the solvent density was maintained using a Parrinello-Rahman barostat (algorithm) at 310 K to perform equilibration in ensemble. Step size of the simulations were fixed to 2 femtoseconds. As we majorly focused on electrostatic interactions, only electrostatic interactions within the cut-off value of 1.0 nm were considered using Lenard-Jones potential. For calculating the protein-ammonium ion and protein-sulphate ion electrostatic interactions, a cut-off length was set at 4 Å. The hydrophobic core areas for PLB1, in presence of 150 mM and 650 mM ammonium sulphate, were generated using ProteinTools.

**Kuhn Length ($L_k$) Calculations:**

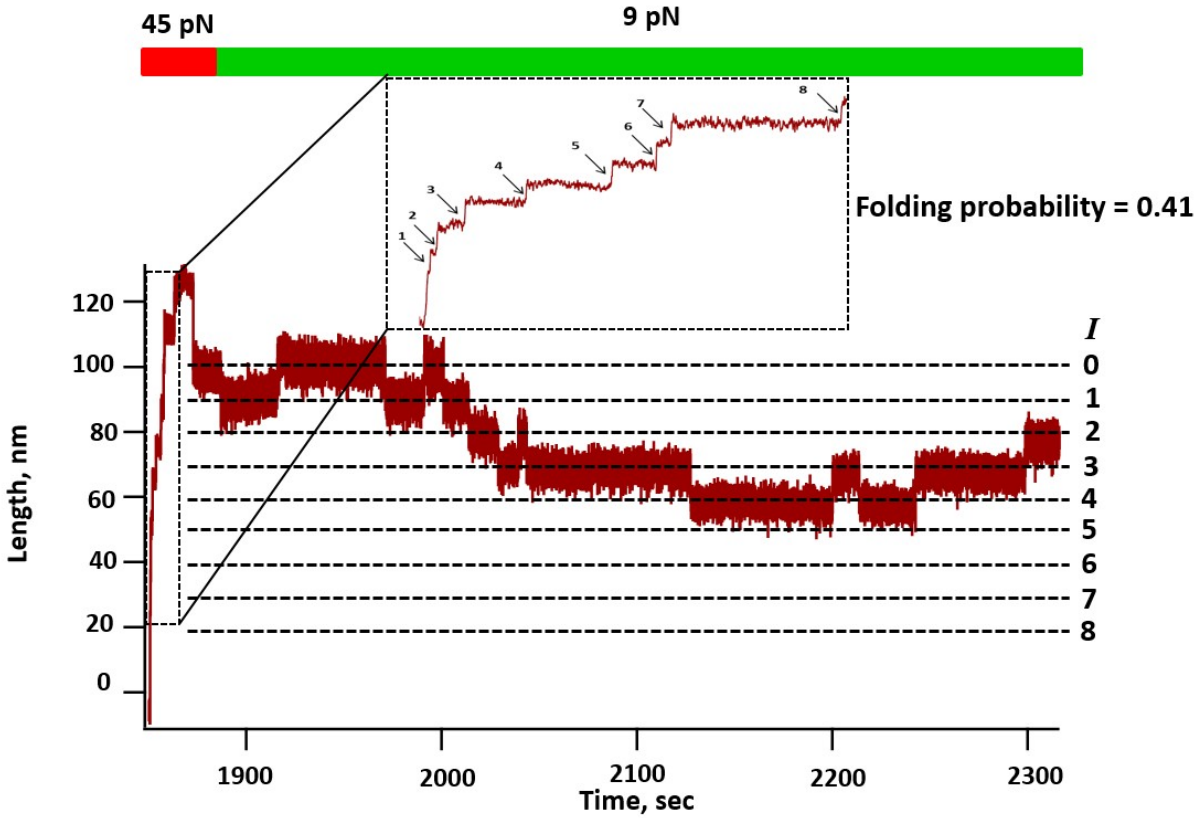
We determined the Kuhn length for protein L by fitting the step size (SS) data in electrolyte freely jointed chain (Ec-FJC) model (eq. 6) against the varying forces at a particular salt condition.

| Salt Concentration (mM) | Kuhn Length, $L_k$ (nm) |
|---|---|
| 150 | 1.18±0.2 |
| 325 | 1.05±0.1 |
| 500 | 0.82±0.1 |
| 650 | 0.58±0.1 |

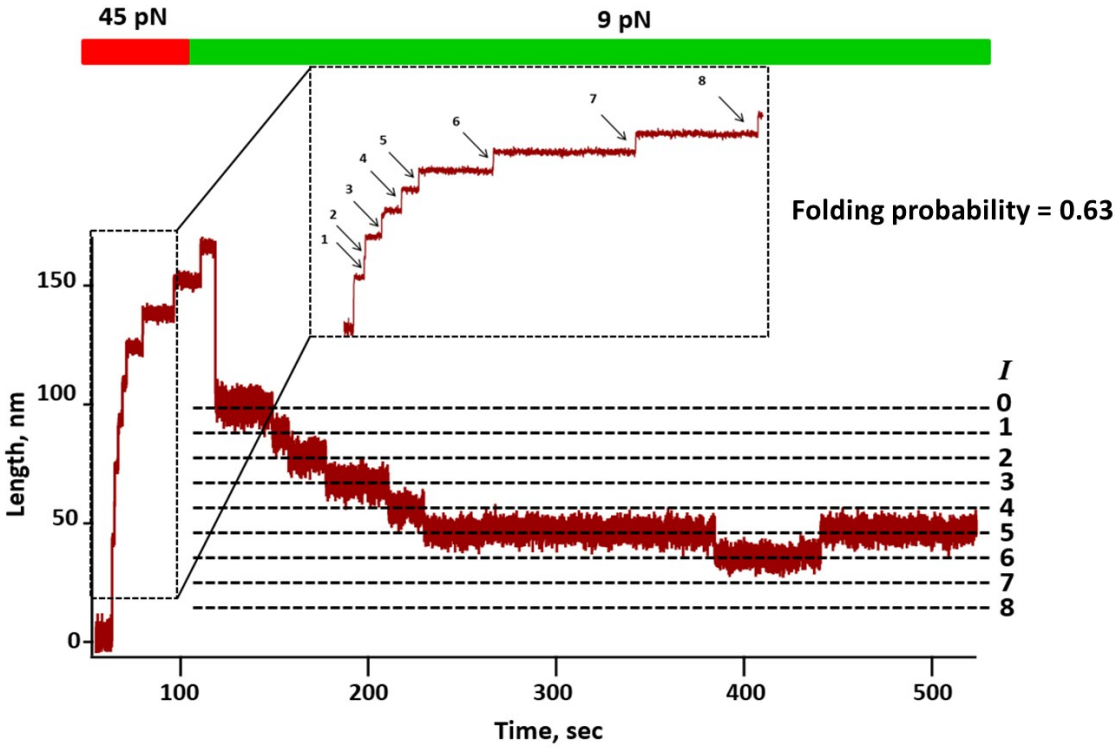
**Supplementary Table 2:** Variation in Kuhn length with increasing salt concentration

**Protein L folding dynamics under varying salt conditions:**

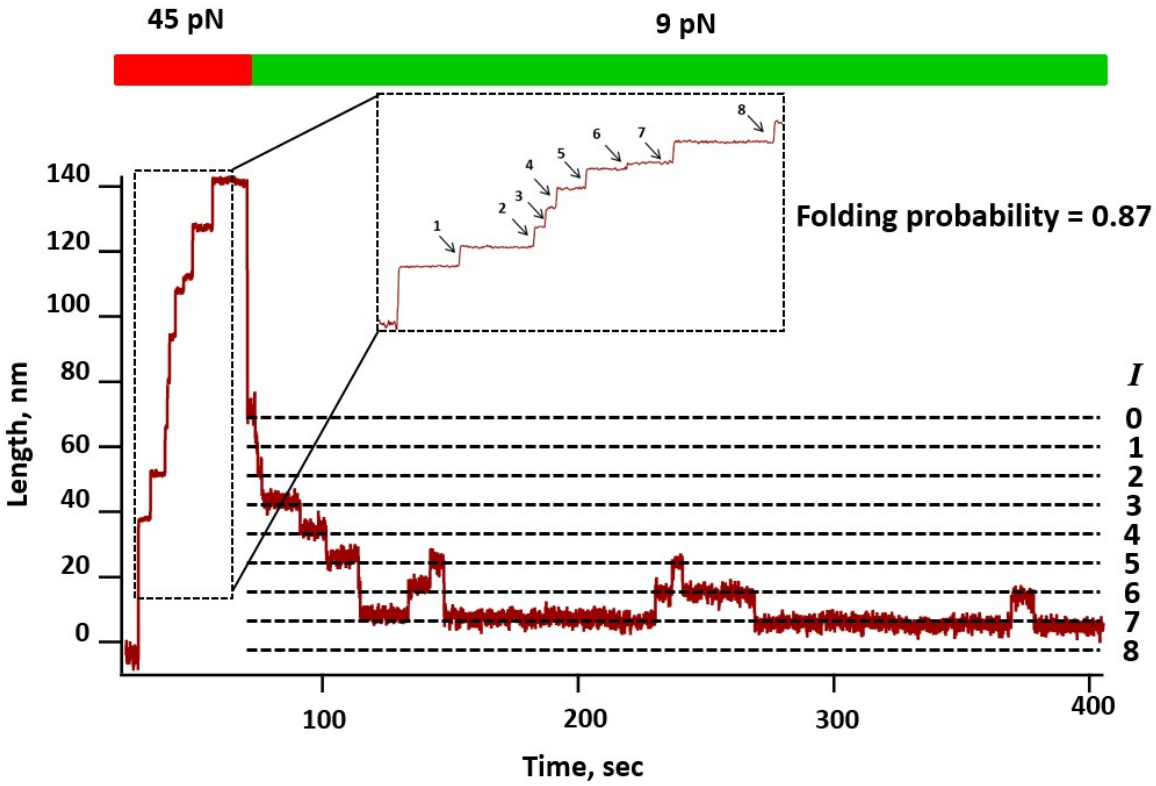
We calculated the folding dynamics of protein L within 150 to 650 mM salt concentrations at 9 pN and observed salt shifts the folding dynamics of protein L towards the folded state and increases the FP from 0.41 at 150 mM to 0.97 with 650 mM concentration (Supplementary figure 2 to 5). These experiments were performed on different protein L molecules and are the supportive data to Fig. 2 (main text).



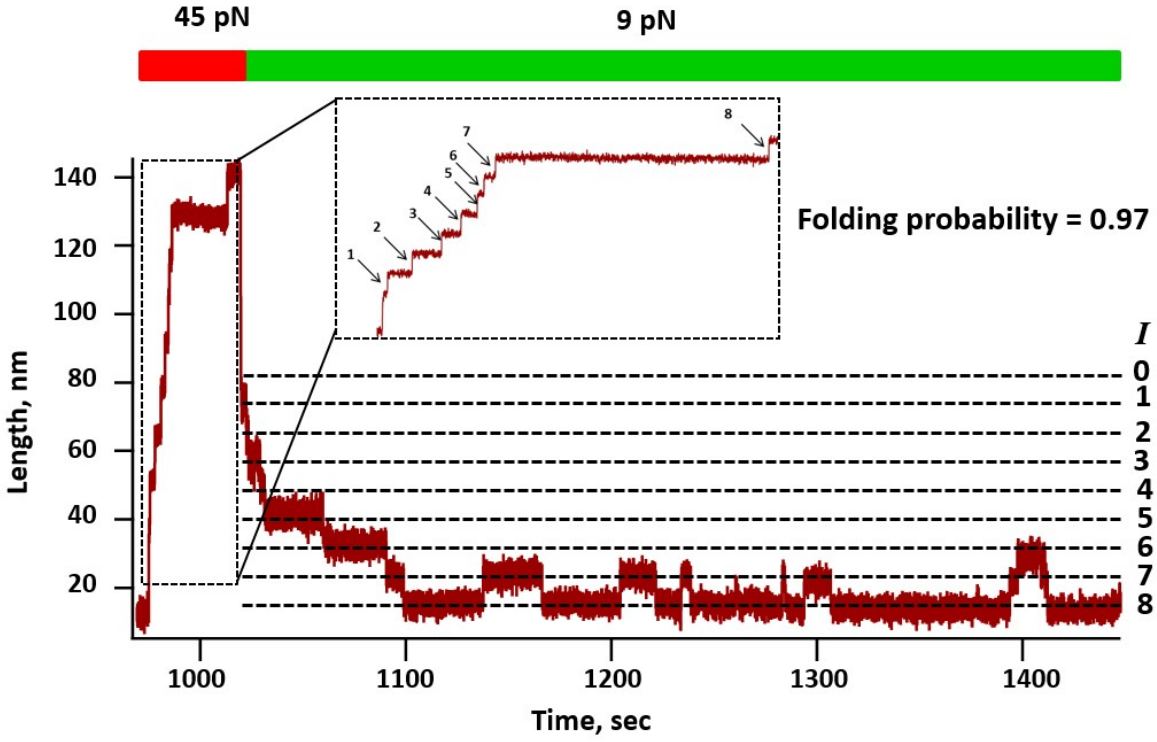

***Supplementary Figure 1: Effect of 150 mM $(NH_4)_2SO_4$ on protein L folding dynamics:*** *After full unfolding of protein L octamer at 45 pN, an equilibrium force of 9 pN was given in presence of 150 mM $(NH_4)_2SO_4$ where the polyprotein exhibits dynamic folding-unfolding transition. Here the observed folding probability at 9 pN is 0.41.*



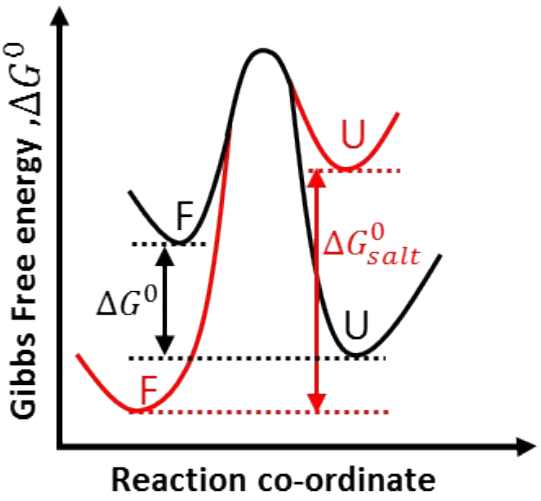

***Supplementary Figure 2: Effect of 325 mM $(NH_4)_2SO_4$ on protein L folding dynamics:*** *Unfolding force of 45 pN is applied on protein L, displaying eight discrete step sizes. Refolding pulse at 8 pN was provided to observe the folding dynamics at equilibrium, where we measured the folding probability is 0.63.*



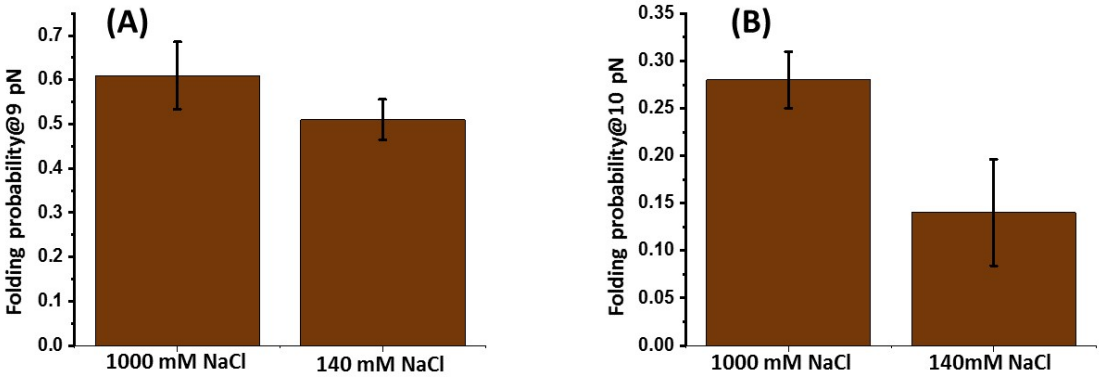

***Supplementary Figure 3: Effect of 500 mM $(NH_4)_2SO_4$ on protein L folding dynamics:*** *Initially, the polyprotein is fully unfolded at 45 pN force, getting eight unfolded domain that refolds back at 9 pN force where the folding probability has been observed to be 0.87.*



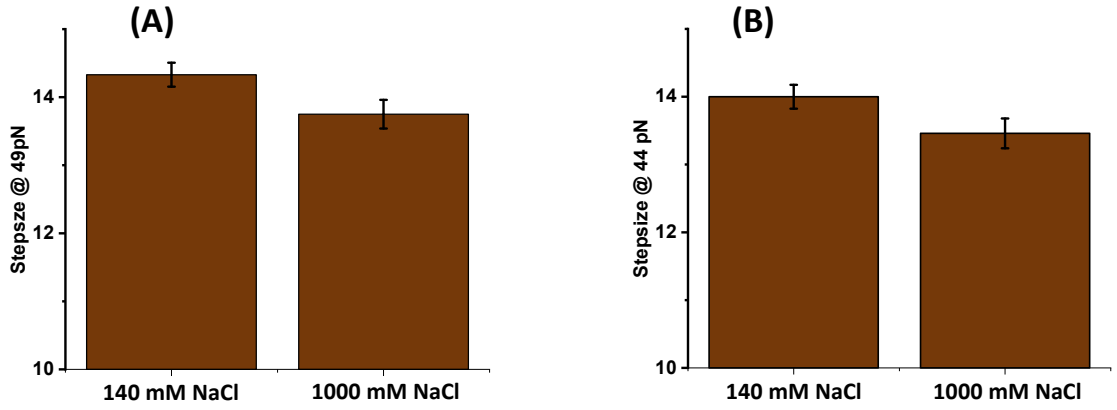

***Supplementary Figure 4: Effect of 650 mM $(NH_4)_2SO_4$ on protein L folding dynamics:*** *After applying an unfolding pulse at 45 pN, an equilibrium force of 9 pN was given in presence of 650 mM $(NH_4)_2SO_4$ where the polyprotein exhibits dynamic folding-unfolding transition. The observed folding probability at 9 pN is 0.97.*



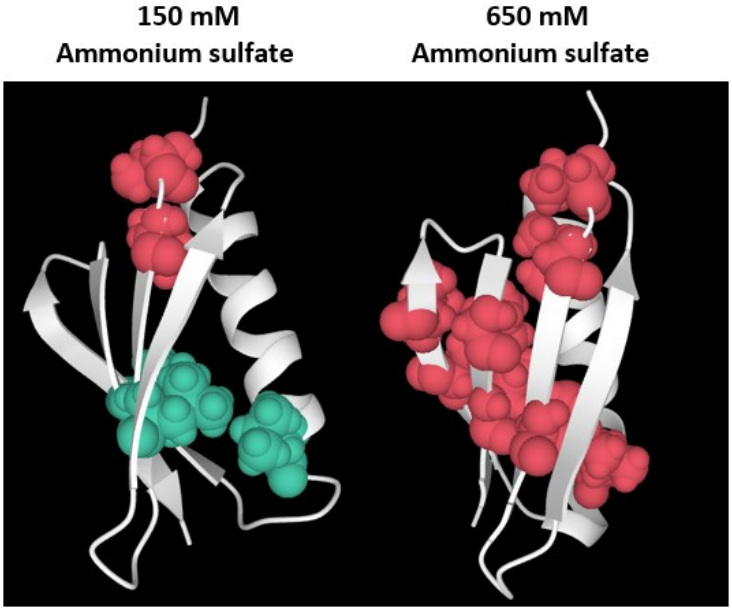

***Supplementary Figure 5: Salt reshapes the folding landscape of protein L:*** *Salt changes the free energy landscape of protein L by stabilizing folded state and destabilizing unfolded state and thus, increases the difference in $\Delta G^0$ between folded and unfolded state. This leads to an increase in the unfolding free energy barrier; however, salt does not change the distance to the transition state.*



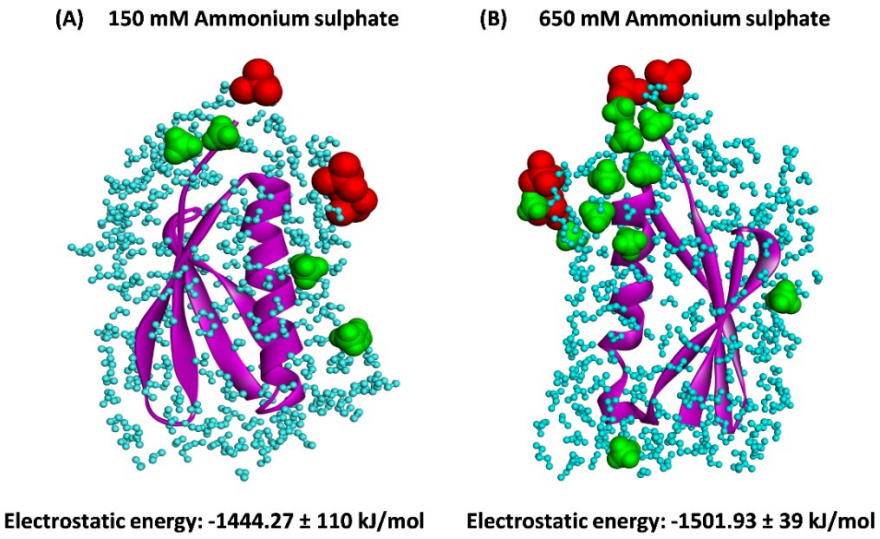

***Supplementary Figure 6: Variation of folding probability of protein L (A) at 9 pN:*** *With 140 mM salt, the folding probability is 0.61±0.07, while it decreases to only 0.51±0.05 in the presence of 1000 mM NaCl.* ***(B) at 10 pN:*** *Similarly, at 9 pN, with 140mM salt, the folding probability is 0.28±0.03 but reduces to only 0.14±0.06 at 1000 mM salt concentration. Both the Figure shows increase in folding probability with increase in salt concentration. Data points are calculated using >2500s and over six molecules per force. Error bars represent s.e.m.*




***Supplementary Figure 7: (A) Step size difference of protein L at 49 pN:*** *In the presence of 140mM NaCl, the step size is 14.33±0.18 nm, while in the presence of 1000 mM NaCl, it decreases to 13.75±0.21 nm.* ***(B) Step size difference of protein L at 44 pN:*** *Similarly, at 44 pN, with 140 mM NaCl, the step size is14±0.18 nm which decreases to 13.46±0.22 nm at 1000 mM salt concentration. More than 10 molecules are measured and averaged for step size at different salt concentration. Error bars are s.e.m.*




***Supplementary Figure 8: Hydrophobic cluster analysis:*** *With 150 mM salt, the hydrophobic cluster area is 256.9 Å$^2$ which increases to 639.2 Å$^2$ at 650 mM salt concentration. At 150 mM, there is two clusters (shown as green and red) in protein L locked conformation; however, in 650 mM protein L conformation, there is a single large hydrophobic cluster (shown as red) with 16 residue contacts in the protein L locked conformation.*




***Supplementary Figure 9: MD simulation revealed electrostatic contribution: (A)*** *At 150 mM ammonium sulfate concentration, 5 ammonium (green) and 3 sulfate (red) interacts with the protein L domain, whereas* ***(B)*** *at 650 mM ammonium sulfate, the 12 ammonium ions interact with the protein. However, the sulfate ion remains almost same (4 ions interact). This results in a minor change in electrostatic energy from -1444.2±110 kJ/mol to -1501.93 kJ/mol.*

**Reference:**


1 S. Garcia-Manyes, L. Dougan, C. L. Badilla, J. Brujić and J. M. Fernández, *PNAS*, 2009, **106**, 10534–10539.
2 O. Jw, K. De, B. D and Z. Ky, Structures of the B1 Domain of Protein L From Peptostreptococcus Magnus With a Tyrosine to Tryptophan Substitution, https://pubmed.ncbi.nlm.nih.gov/11264576/, (accessed May 20, 2020).
3 I. Popa, J. A. Rivas-Pardo, E. C. Eckels, D. J. Echelman, C. L. Badilla, J. Valle-Orero and J. M. Fernández, *J. Am. Chem. Soc.*, 2016, **138**, 10546–10553.
4 I. Popa, R. Berkovich, J. Alegre-Cebollada, C. L. Badilla, J. A. Rivas-Pardo, Y. Taniguchi, M. Kawakami and J. M. Fernandez, *J. Am. Chem. Soc.*, 2013, **135**, 12762–12771.
5 S. Chakraborty, D. Chaudhuri, S. Banerjee, M. Bhatt and S. Haldar, *Commun Biol*, 2022, **5**, 1–14.
6 K. D. Beyer and D. D. Ebeling, *Geophysical Research Letters*, 1998, **25**, 3147–3150.
7 S. Haldar, R. Tapia-Rojo, E. C. Eckels, J. Valle-Orero and J. M. Fernandez, *Nat Commun*, 2017, **8**, 668.
8 J. Valle-Orero, J. A. Rivas-Pardo, R. Tapia-Rojo, I. Popa, D. J. Echelman, S. Haldar and J. M. Fernández, *Angewandte Chemie International Edition*, 2017, **56**, 9741–9746.
9 C. Gosse and V. Croquette, *Biophys J*, 2002, **82**, 3314–3329.
10 J. W. O'Neill, D. E. Kim, K. Johnsen, D. Baker and K. Y. Zhang, *Structure*, 2001, **9**, 1017–1027.
11 R. Anandakrishnan, B. Aguilar and A. V. Onufriev, *Nucleic Acids Research*, 2012, **40**, W537–W541.

12 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.
13 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.
14 S. A. Adcock and J. A. McCammon, *Chem. Rev.*, 2006, **106**, 1589–1615.