Supplementary Information to

Ion mobility-mass spectrometry shows step-wise protein unfolding under alkaline conditions

Cagla Sahin, Nicklas Österlund, Axel Leppert, Jan Johansson, Erik G. Marklund, Justin L. P. Benesch, Leopold L. Ilag, Timothy M. Allison, and Michael Landreh

Experimental

Protein preparation. Bovine serum albumin (BSA), horse heart myoglobin, and lysozyme were purchased from Sigma. α-Synuclein and proSP-C BRICHOS were purified as described [1,2]. Proteins were dissolved to a final concentration of 10 μM (BSA, Myoglobin, α-Synuclein, lysozyme) or 30 μM (proSP-C BRICHOS) in dH_2O and stored at -20 °C. Buffer exchange was carried out using Micro Bio-Spin P6 gel columns (Bio-Rad). Proteins were exchanged into 7 M aqueous ammonia (ammonium hydroxide) solution, pH 12.0, 0.7 M ammonia for pH 11.6, or 0.07 M ammonia for pH 11.0, or into 1 M ammonium acetate for pH 6.9. Ammonia and ammonium acetate concentrations were chosen to minimize the risk of ESI-induced pH changes in the ESI droplets. For BSA, the ammonia solutions were titrated with 5, 50, or 500 mM final concentrations ammonium acetate. Water/formic acid was used for pH 2.0. It should be noted that the ionic strengths of the solutions differ, as pH 12 could only be reached at high ammonia concentrations (Figure S1).

*Mass spectrometry. n*ESI capillaries were purchased from Thermo. Mass spectra were acquired on a Micromass LCT ToF modified for analysis of intact protein complexes (MS Vision, The Netherlands) equipped with an offline nanospray source. The capillary voltage was 1.5 kV and the RF lens 1.5 kV. The cone voltage was 100 V and the pressure in the ion source was maintained at 9.0 mbar. IM-MS spectra were acquired on a Waters Synapt G2S travelling-wave ion mobility mass spectrometer (Waters, UK) equipped with an offline nanospray source. The capillary voltage was 1.5 kV, the source pressure was 3.4 mbar, and the source temperature was 80 °C unless otherwise noted. Collision-induced activation was performed by ramping the collision voltage in the ion trap from 5 to 100 V in steps of 5 V and a spectrum was recorded at each step. Wave height and wave velocity were, respectively, 35 V and 700 m/s in the IMS cell, and 10 V and 248 m/s in the transfer. IMS gas was nitrogen with a flow of 50 mL/min. CIU data were recorded at pH 2.0, 6.9, 11.0, and 12.0 for BSA to probe unfolding under native conditions, partially, or fully destabilizing conditions. For myoglobin, α-Synuclein, and lysozyme, CIU data were recorded at pH 6.9 and pH 12.0. CIU of proSP-C BRICHOS was performed at pH 6.9 and 11.0 to probe the effect of deprotonating the intermolecular salt bridge. IM-MS data for each protein system were recorded on the same day and with the same

instrument settings to enable direct comparisons of drift time values. Mass spectra were processed using MassLynx 4.1 (Waters). Average charges for BSA were calculated as weighted averages from the 10× smoothed spectra. IM-MS data were analyzed using PULSAR [3]. It is important to note that the drift time data in the CIU plots were not converted to CCS values, for two reasons: (1) CIU reduces the mobility of protein ions, leading to inaccurate CCS values when native-like proteins are used as calibrants [4]. (2)

CD Spectroscopy. Protein samples were prepared diluting protein stock solutions of BSA (50 mg/ml) and lysozyme (50 mg/ml) to 0.2 mg/ml, and proSP-C BRICHOS (18 mg/ml) to 0.18 mg/ml in 20 mM sodium phosphate (pH 7), 10 mM NaOH (pH 12.0) or 20 mM sodium dihydrogen phosphate adjusted with phosphoric acid (pH 2.0). The buffer conditions were chosen for optimal signal-to-noise ratios in far-UV spectra.[5] Samples were equilibrated in the respective buffer for 30 min prior to the experiment. Data acquisitions were performed using a JASCO J-1500 CD spectrophotometer. Spectra were recorded from 260 to 195 nm at 25 °C using a quartz cuvette with a 1 mm path length. The bandwidth was set to 1 nm and the scan steps to 0.5 nm at a scanning speed of 20 nm/min. Spectra show the average of five consecutive scans.

Size exclusion chromatography. Size exclusion chromatography (SEC) experiments were performed on an ÄKTA basic 10 FPLC system (GE Healthcare, UK) equipped with a Superdex 75 Increase 10/300 GL column (GE Healthcare, UK). 40 μM proSP-C BRICHOS, diluted from a 1 mM stock, were incubated for 30 min in either 20 mM Sodium dihydrogen phosphate (pH 5.0); 20 mM disodium hydrogen phosphate (pH 11.0) or a mixture of both (pH 8.0). Prior to the experiment the SEC column was equilibrated with the corresponding buffer.

Absorption spectroscopy. Myoglobin was diluted to 0.2 mg/mL in either ddH2O alone (pH 7), or with 12.5 % ammonia (pH 12.0) or 1 % formic acid (pH 2.0). Samples were prepared in triplicates in a black 96 well plate with clear bottom. Absorption spectra were recorded from 300 nm to 500 nm in 1 nm steps on a Multiskan Go, Thermo Scientific. Path length correction was performed at 900 and 975 nm using water as blank.

Isoelectric point calculations. The proteomes of nine different organisms were downloaded in FASTA format from NCBI Reference Sequence Database (https://ftp-ncbi.nlm.nih.gov /genomes/refseq/) (Table S1). The choice of these nine organisms was made to cover a variety of different organisms that need to cope with different environments. The organisms included acidophiles, alkalophiles, halophiles, thermophiles, animals, fungi, archaea, and various bacteria. The

isoelectric point (p*I*) was calculated for each protein in each genome with the iterative algorithm used by Bjellqvist[6] with a binary search strategy to find the pH corresponding to a net charge of zero. The search was stopped when the step size, which also determines the maximum error, reached 0.01 pH units. The Lehninger set of pK_a values were used for calculating protein net charge at a given pH (Table S2). The resulting p*I*-distributions were plotted as kernel-density-estinates with a bandwidth of 0.1 pH unit.

Supplementary References

1 H. Willander, G. Askarieh, M. Landreh, P. Westermark, K. Nordling, H. Keranen, E. Hermansson, A. Hamvas, L. M. Nogee, T. Bergman, A. Saenz, C. Casals, J. Aqvist, H. Jornvall, H. Berglund, J. Presto, S. D. Knight and J. Johansson, *Proc. Natl. Acad. Sci.*, 2012, **109**, 2325–2329.

2 C. Sahin, L. Kjær, M. S. Christensen, J. N Pedersen, G. Christiansen, A. M. W. Pérez, I. M. Møller, J. J. Enghild, J. S. Pedersen, K. Larsen and D. E. Otzen, *Biochemistry*, 2018, **57**, 5145–5158.

3 T. M. Allison, E. Reading, I. Liko, A. J. Baldwin, A. Laganowsky and C. V. Robinson, *Nat. Commun.*, 2015, **6**, 8551.

4 T. M. Allison, M. Landreh, J. L. P. Benesch, C. V. Robinson, *Anal. Chem.* 2016, **88**:5879-5884.

5 S. M. Kelly, T. J. Jess, N. C. Price, *Biochem. Biophys. Acta* 2005, **1751**:119-139.

6 Bjellqvist B, Hughes GJ, Pasquali Ch, Paquet N, Ravier F, Sanchez JCh, Frutiger S, Hochstrasser D. *Electrophoresis* 1993, **14**:1023-1031.

7 Nelson, D. L., *et al.* (2008). Lehninger Principles of Biochemistry, Macmillan.

Table S1. Organisms, the proteomes of which was subjected to pI calculations, listed with their RefSeq assembly names, types of extremophile, and taxonomic kingdom.

Organism	**Assembly**	**Extremophile**	**Kingdom**
Acetobacter aceti	GCF_000379545.1_ASM37954v1	Acidophile	*Bacteria*
Escherichia coli	GCF_000005845.2_ASM584v2		*Bacteria*
Halobaculum gomorrense	GCF_900129775.1_IMG-taxon_2695420988	Halophile	*Archaea*
Halorhodospira halochloris	GCF_002356555.2_HH1059_assembly_1.0	Halophile	*Bacteria*
Homo sapiens	GCF_000001405.39 (GRCh38)		*Animalia*
Mycobacterium tuberculosis	GCF_000195955.2_ASM19595v2		*Bacteria*
Mycoplasma pneumoniae	GCF_001272835.1_ASM127283v1		*Bacteria*
Saccharomyces cerevisiae	GCF_000146045.2_R64		*Fungi*
Thermus aquaticus	GCF_001399775.1_ASM139977v1	Thermophile	*Bacteria*
Yersinia pestis	GCF_000222975.1_ASM22297v1		*Bacteria*

Table S2. pK_a used for the pI calculations.[7]

Residue	pK_a
H	6.00
K	10.5
R	12.4
C	8.35
D	3.86
E	4.25
Y	10.0
N-terminus	9.69
C-terminus	2.34

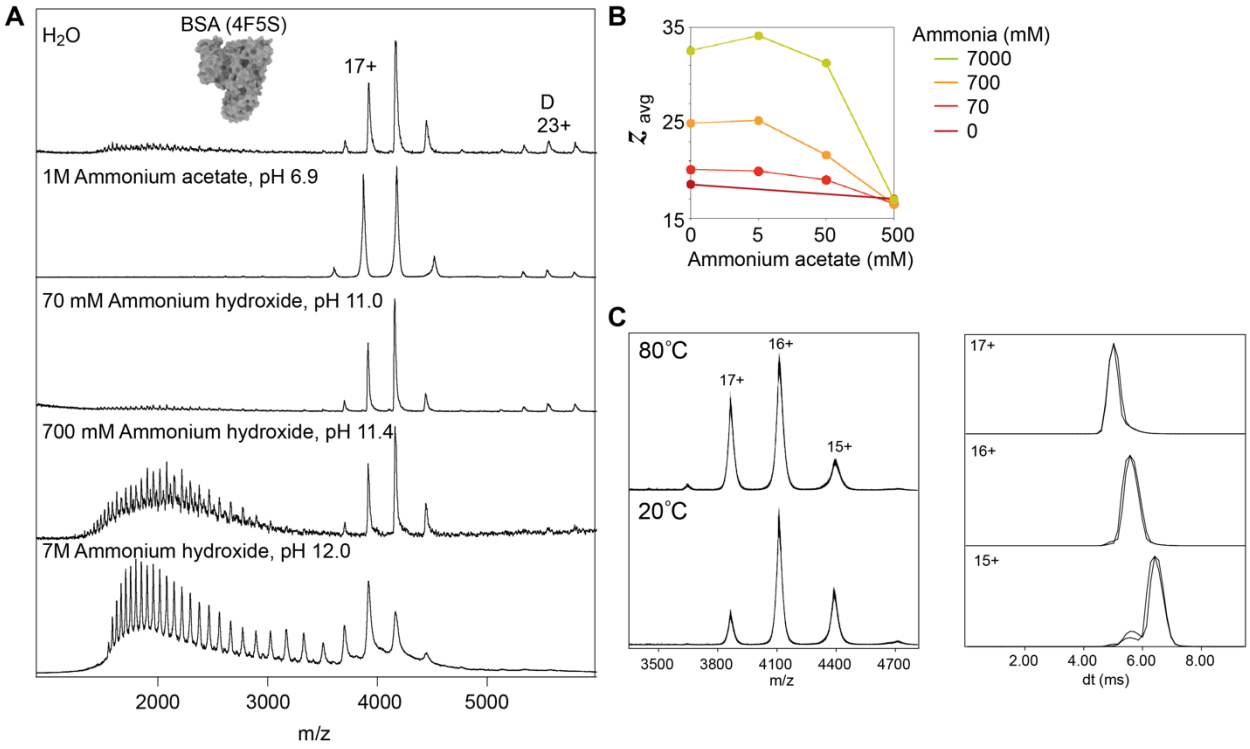

Figure S1. (A) MS spectra of BSA show a native-like CSD centered on the 16+ and 17+ ions at pH in pure water, 1M ammonium acetate, pH 6.9, or 70 mM ammonia, pH 11.0. Increasing the pH further to 11.6 and 12.0 by raising the ammonia concentration to 700 mM or 7 M results in appearance of a second CSD centered on the 37+ charge state. BSA dimers are indicated by a "D". (B) Effect of counterions on pH-induced unfolding. Ammonium acetate was added to BSA in 0, 70, 700 mM or 7 M ammonia solutions to final concentrations of 5, 500, or 500 mM, and the ion charge calculated as weighted average. Only 500 mM ammonium acetate resulted in a strong decrease in ion charge. It should be noted that the addition of 500 mM ammonium acetate lowered the pH of a 7 M ammonia solution from 12.0 to 11.0. (C) Effect of source temperature on arrival time distributions of BSA. Left, nESI spectra recorded at source temperatures of 80 and 20 °C. Right: Overlay of the drift time plots for the main charge states show no significant change in response to raising the temperature.

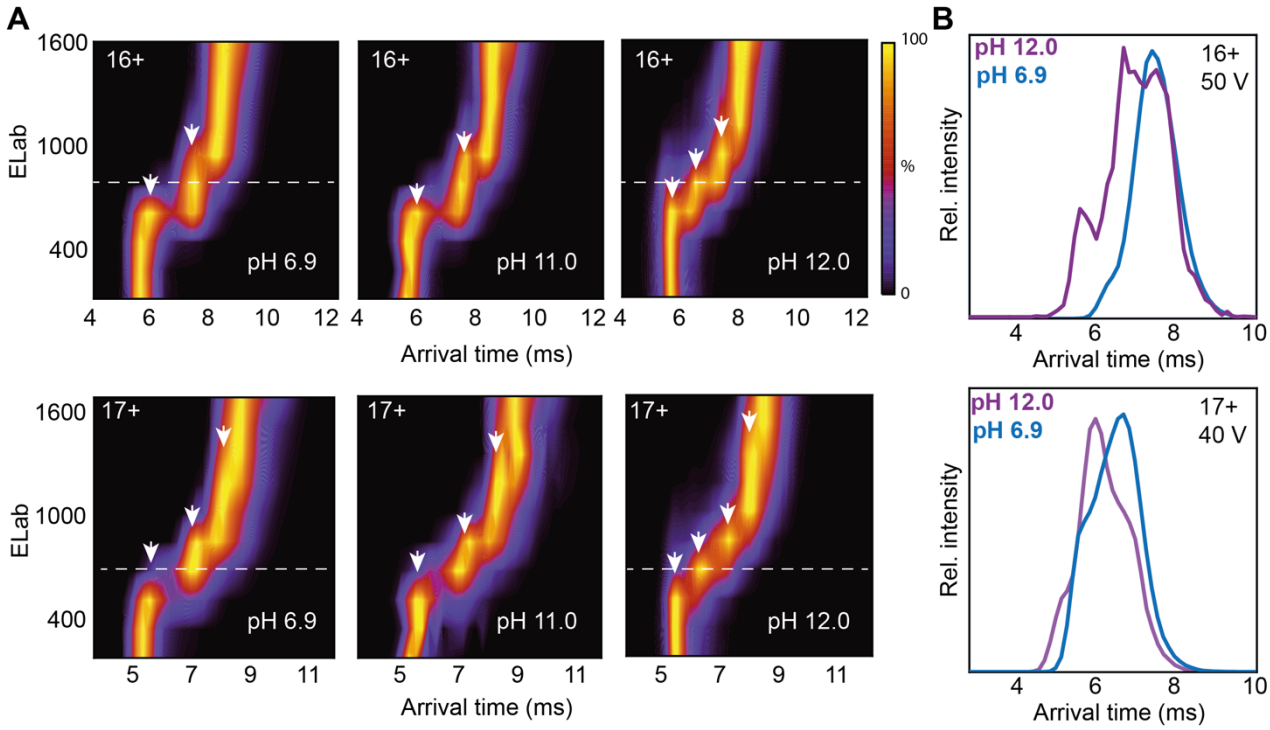

Figure S2. (A) CIU plots of the 16+ and 17+ charge states of BSA show additional unfolding steps at pH 12.0 compared to pH 11.0 and 6.9. Arrows indicate unfolding transitions. Dashed lines indicate the voltage steps of the drit time plots in panel B. (B) Drift time plots of the 16+ and 17+ charge states at 50 V and 40 V activation energy, respectively show the presence of additional populations at pH 12.0 that are not present at pH 6.9.

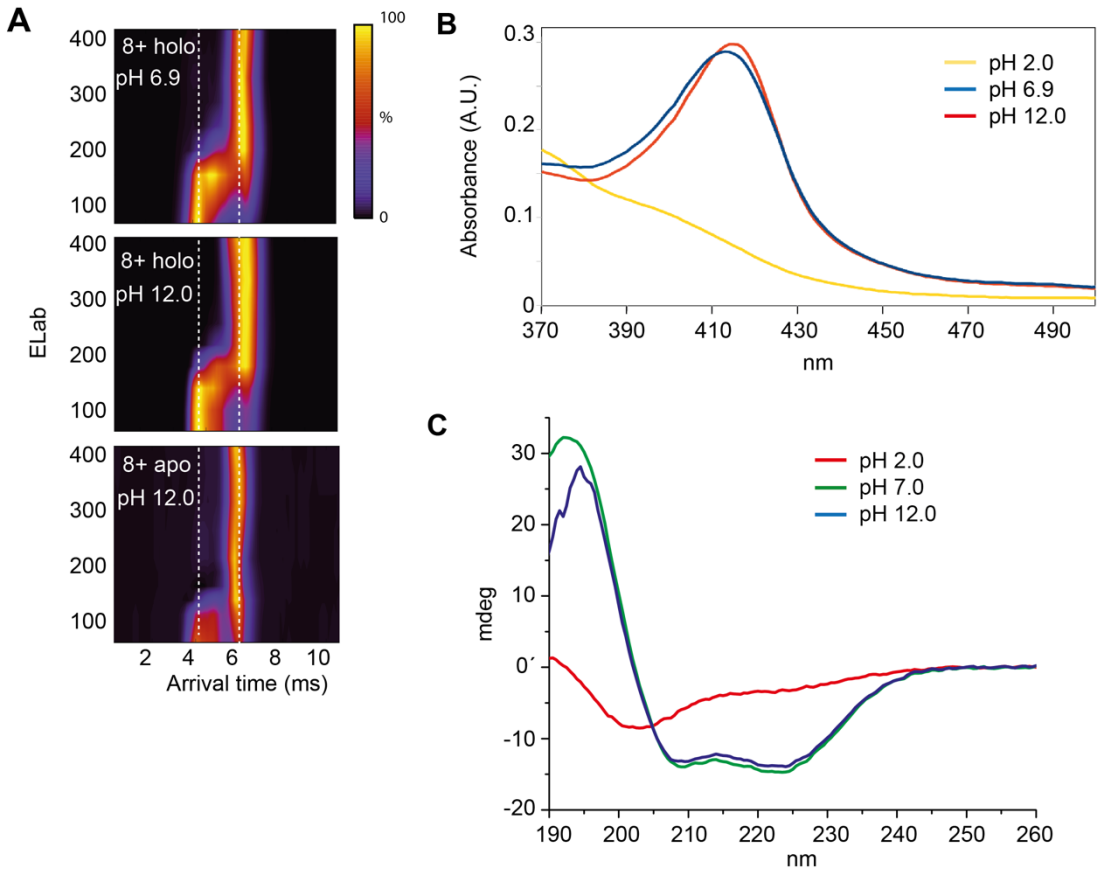

Figure S3. (A) CIU plots of the 8+ charge state of native holo-myoglobin, pH 6.9, and of holo- and apo-myoglobin, pH 12.0. (B) Absorbance spectra of the myoglobin heme group show that that the holo-protein in solution is completely dissociated at pH 2.0 but remains intact at pH 7.0 and 12.0. (C) CD spectroscopy of myoglobin shows a native-like secondary structure content at pH 12 and complete unfolding at pH 2.0.

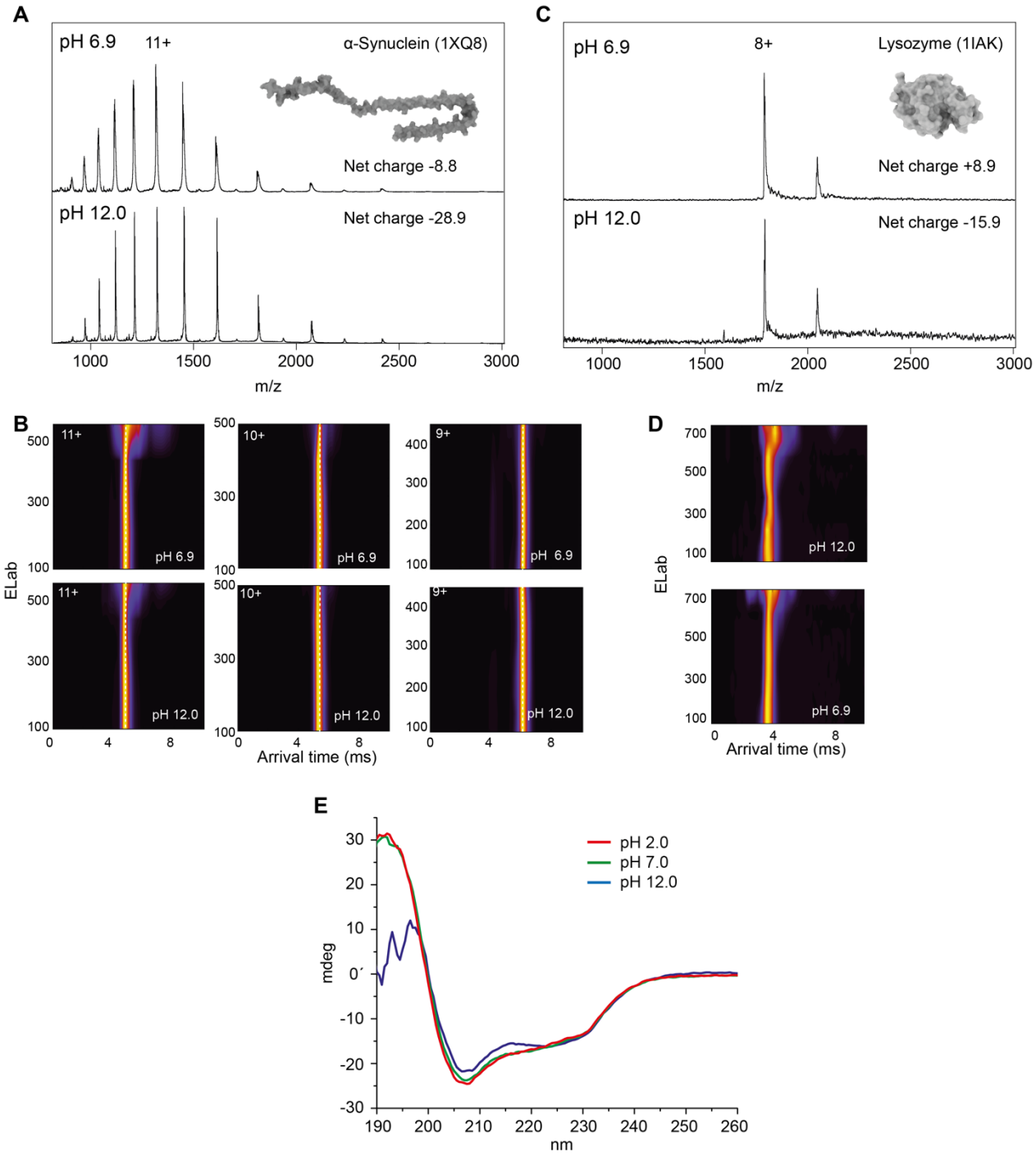

Figure S4. (A) The CSD of intrinsically disordered α-synuclein is not affected between pH 6.9 and 12.0. (B) CIU profiles of the three major low charge states show no changes in their arrival time distributions between pH 6.9 and 12.0. (C) Lysozyme retains a native-like CSD at pH 6.9 and pH 12.0 High-resolution structures for each protein are shown as inserts. (D) CIU plots of the 8+ charge state of lysozyme at pH 6.9 and 12.0. (E) CD spectra of lysozyme at pH 7.0 and 12.0 show no significant change in secondary structure.

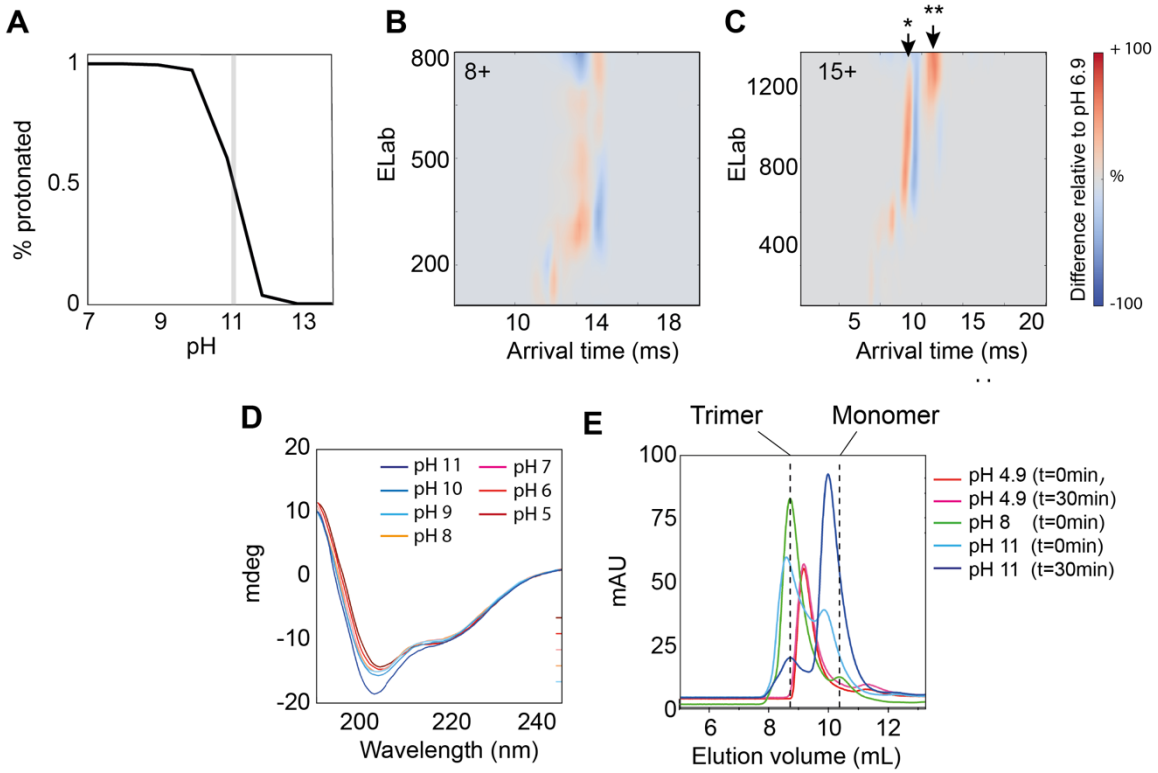

Figure S5. (A) The Arg139-Glu191 salt bridge has a predicted pKa of 10.9. The theoretical titration curve for the salt bridge was generated by analysing the PDB file using the DelPhiPKa server (compbio.clemson.edu/pka_webserver). CIU difference plots of the 8+ monomer (B) and the 15+ trimer (C) of proSP-C BRICHOS. Red shading indicates higher signal intensity at pH 6.9, and blue shading indicates higher signal intensity at pH 12.0. Arrows in (C) highlight the lower drift time of the unfolded BRICHOS trimer at pH 6.9 (*), as well as an additional unfolding step not observed at pH 12.0 (**). Separate plots are shown in Figure 3B and C. (D) CD spectroscopy of proSP-C BRICHOS shows a small increase in random coil content at pH 11.0, consistent with minor structural re-arrangements. (E) Size exclusion chromatography of proSP-C BRICHOS shows monomerization at pH 11.0.