Supplementary Information

Albumin Unfolding and Refolding Cycle by a Time-Controlled pH-Jump

Alessandra Del Giudice,*[a]+ Daniele Del Giudice,[a][b]+ Emanuele Spatola,[a][c] Valentina Alemanno,[a] Luciano Galantini,[a] Stefano Di Stefano[a]

[a] Dr. A. Del Giudice, Dr. D. Del Giudice, Dr. E. Spatola, V. Alemanno, Prof. L. Galantini, Prof. S. Di Stefano
Department of Chemistry
Sapienza University of Rome
P.le A. Moro 5 00185 Rome, Italy
E-mail: alessandra.delgiudice@uniroma1.it

[b] Present address: Institute for Complex Molecular Systems and the Laboratory of Macromolecular and Organic Chemistry, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

[c] Present address: Institute of Organic Chemistry, Ulm University Albert-Einstein-Allee 11, 89081 Ulm, Germany.

+ The authors both contributed equally to this work.

Experimental procedures

Synthesis of nitroacetic acid. Nitroacetic acid (abbreviated NAA or acid **1**) was obtained according to the protocol reported by Vanler et al,[1] with some modifications. In brief, nitromethane is initially carbonated by reacting with potassium hydroxide under air in an open reaction vessel, giving the dipotassium nitroacetate salt as a pale pink solid. The following protonation with HCl leads the desired product.

Sample preparation. For the fatty acid-free HSA (samples HdT5 and HdT10) HSA lyophilized powder, fatty acid- and globulin-free (≥ 99 %, type A3782, Sigma) was used, while for the "fatted" HSA (sample HfT5) and Bovine Serum Albumin (BSA, sample BfT5) the Sigma products A8763 (≥ 99 %) and 05470 (≥ 98 %) were employed, respectively. For buffer preparation, calculated amounts of Tris base (Sigma) stock solution 0.2 M, NaCl (Carlo Erba reagents RPE) 1 M stock solution and HCl (from 37% Carlo Erba RPE) 1 M solutions were mixed and brought to volume. pH measurements were performed with a Hamilton minitrode electrode connected to a Crison 2002 pH-meter and calibrated with standard buffer solutions at pH 7.00 and 4.01, at 25°C.

Initial protein solutions at concentrations close to 20 mg/mL were prepared by dissolving the powders in the buffer TrisHCl 5 mM (or 10 mM), NaCl 0.15 M at pH 8.1. Hence, the solutions were centrifuged at 12×10^3 g for 5 minutes. Diluted working solutions were prepared to reach a protein concentration of 3-4 mg/mL, and the protein concentration was estimated by measuring the absorbance at 280 nm and using the extinction coefficients[2] 35700 $M^{-1}cm^{-1}$ for HSA and 43324 for BSA using a Nanodrop spectrophotometer. For the time-dependent pH jump induced by acid **1** decarboxylation, the reaction was started by adding a known volume of protein solution to a weighted amount of solid acid **1** and by gently mixing for 30 seconds until full dissolution. The blanks needed for background subtraction were obtained by applying the same procedure using the buffer rather than the protein solution. Several combinations of acid **1** concentration and buffer concentration were used as detailed in the results. Nitroacetic acid concentration is expressed in mM considering the weighted amount, a molecular weight of 105.049 g/mol and the volume of buffer or protein solution in which it was dissolved.

Complementary samples of acidified protein were obtained for certain final pH values (Table S2), by mixing 450 µl of HSA solution in buffer and 50 µl conveniently diluted HCl in NaCl 0.15 M solution. The "static" samples were measured 1 hour after mixing and the pH checked after 24 hours.

Fluorescence measurements. The fluorescence measurements were carried out in a Horiba Jobin-Yvon FLUOROMAX 4spectrofluorometer (Kyoto, Japan), using a quartz cuvette with pathlengths 0.4 cm (excitation) x 1 cm (emission) and the

temperature was controlled at 25°C with the Peltier accessory. The intrinsic fluorescence was measured by exciting the protein solution at λ_{ex} =295 nm with a band-pass of both excitation and emission monochromators of 3 nm. Emission spectra were collected at 1 minute intervals between 305 and 550 nm with a step of 1 nm and an acquisition time of 0.05 s. The spectra were subtracted for the buffer background. A correction for the inner filter effect ($F_{corr}=F_{raw}10^{Abs(\lambda ex)+Abs(\lambda em)}$) was applied to the blank-subtracted spectra considering the experimental absorbance values at the excitation and emission wavelengths ($Abs(\lambda_{ex})$, $Abs(\lambda_{em})$) measured on an aliquot of the same initial sample. UV-vis absorption spectra were collected with a Cary 1E UV–VIS spectrophotometer between 240 and 550 nm with a step of 1 nm and an acquisition time of 0.1 s, at 1-minute intervals. The samples were measured in a quartz cuvette with 0.4 cm path length.

SAXS experiments and data analysis. SAXS measurements were performed at the BM29 beamline[3] of ESRF (Grenoble, France) (DOI10.15151/ESRF-ES-514136926, DOI10.15151/ESRF-ES-61665128) with X-ray wavelength λ = 0.992 Å and sample-to-detector distance 281 cm, allowing data collection in the scattering vector range $0.07 < q < 5$ nm^{-1}, where $q = \frac{4\pi \sin\theta}{\lambda}$ and 2θ is the scattering angle. The same sample was repeatedly measured at 25 °C by using the automatic sample changer mode of acquisition,[4] alternating sample and blank. Ten subsequent 1 s exposures were collected while flowing 50 µl of sample in the capillary. The frames were checked for overlap, automatically averaged, subtracted for the blank contribution, and converted to the molecular weight scale by the automatic pipeline. An approximate sampling rate of one averaged SAXS profile every three minutes was attained to follow the refolding process, and the time stamps of the SAXS acquisitions were used to calculate the elapsed time from the initial mixing between protein solution and acid **1** (NAA), representing the time "0" of the kinetics. Model-independent analysis of the subtracted scattering profiles to evaluate the radius of gyration R_g via the Guinier approximation ($\ln(I) \approx \ln(I(0)) - \frac{R_g^2}{3}q^2$) and the pair distance distribution function [P(R)] via the Indirect Fourier Transform was performed with the ATSAS package[5]. The SAXS data obtained for the fatted HSA preparation (HfT5) from refolding kinetics obtained with different acid **1** concentrations (40 and 80 mM) were merged in a global data set analysed by component analysis with the MCR-ALS toolbox.[6] The data matrix was built from the SAXS data restricted to the q range 0.165-3 nm^{-1} converted to Kratky plots $I(q)q^2$ vs. q, since this representation provides the largest relative variation in the unfolding process, making easier the identification of the minimum number of components needed to fully describe the process and their identity.[7] After extraction of a fundamental basis set of two components, the possibility to describe the data set as a linear combination of basis SAXS profiles including as second or third component a more expanded conformation was checked using the least-square optimization implemented in the Oligomer tool of the ATSAS package.[8] Such additional component identified with the putative "Expanded" form was represented by the SAXS profile collected for a sample of fatted HSA in sodium phosphate buffer 10 mM acidified with HCl 45 mM (pH 1.4), without further NaCl added.

The ensemble-based method implemented in EOM[9] (ATSAS 3.2.1 r14885) was applied to fit experimental SAXS profiles as a linear combination of selected conformers optimized from the same initial pool. With the RANCH tool a pool of 10,000 conformers were generated as nine folded loops (3 per each HSA domain) taken from the crystal structure[10] interconnected by eight flexible linkers having disordered statistics, according to a fragmentation scheme with high flexibility which allows to fit also fully extended conformations.[11–14] The theoretical SAXS profiles in solution and related size parameters were computed with the FFMAKER tool and eight repeated runs of the genetic algorithm optimization were performed.

pH-time calibration. A calibration of the pH variation over time allowed for reporting the variation of the experimental parameters as a function of pH rather than as a function of the time elapsed from the initial acid **1** dissolution. For samples identical to those analysed by SAXS and fluorescence and at the same temperature of the experiments (25°C), pH-vs.-time

profiles were measured with a Hamilton minitrode electrode connected to a Crison 2002 pH-meter and calibrated with standard buffer solutions at pH 7.00 and 4.01. The pH vs. time readings were interpolated at the time points of SAXS and fluorescence kinetics by using Matlab code. In the case of SAXS experiments with high nitroacetic acid content, the effervescence due to CO_2 release in the sample flowing into the capillary was detected as scattering from bubbles (e.g. Figure S4), and the data had to be discarded. For the samples with 80 mM nitroacetic acid the recovered trend of R_g and refolding index as a function of pH deviated from the runs with lower content (Figure 4d), possibly due to a faster pH trend compared to what measured off-line and therefore data collected after 10 minutes from the initial solubilization during the sharper pH jump were not considered to predict the trend. For the analysis in terms of structural components (HfT5), SAXS data measured with both 40 and 80 mM acid **1** were included, but corrected pH values were assigned to some data points of the kinetics with 80 mM **1**, to avoid outliers within the trends of R_g and Kratky plot subtended area.

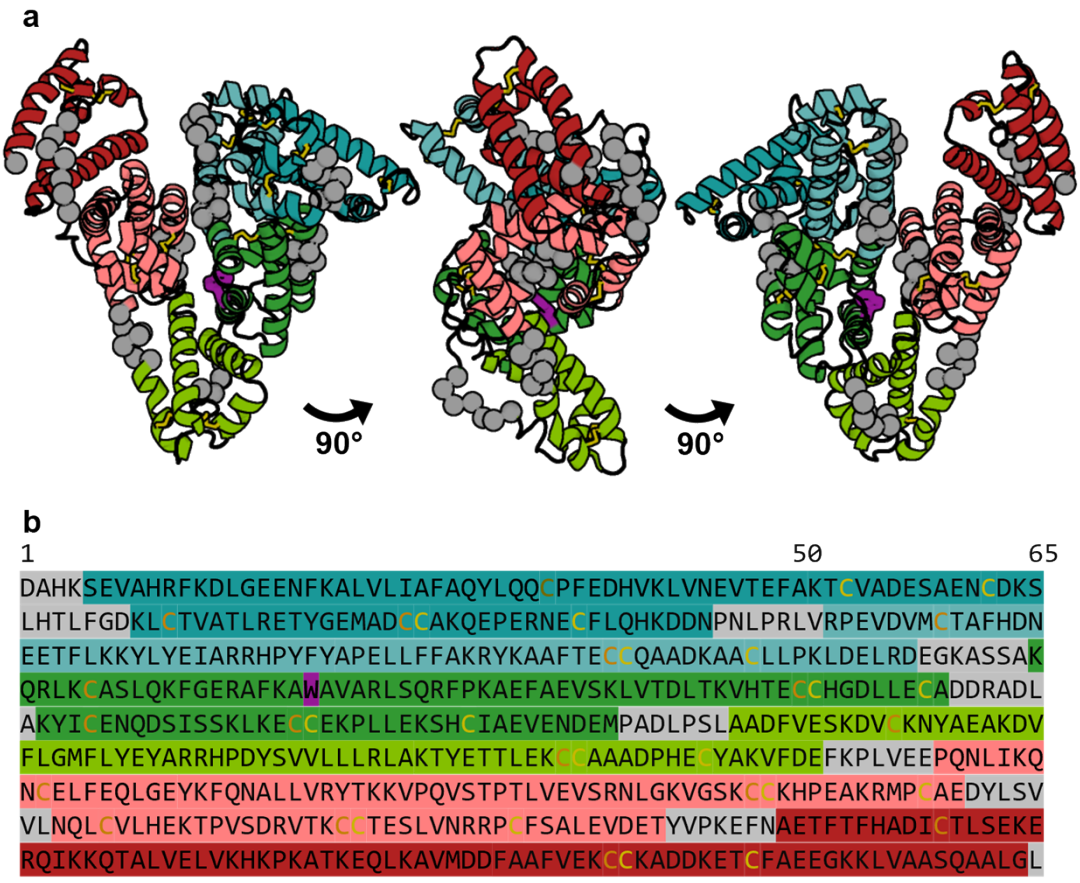

Figure S1. *Structure of HSA and location of the tryptophan residue.* a) The HSA protein is shown highlighting the fragmentation scheme adopted for ensemble modelling of partially unfolded conformations as rigid loops connected by flexible linkers (grey dummy atoms). The coordinates of the pdb entry 1AO6 were used. The protein chain is depicted in cartoon visualization with colours highlighting the three domains (I cyan, II green, III red) and the subdomains (A and B in different tints). Trp214 (thick, violet) and Cys residues (thin, yellow) are visualized as sticks. Representation made with PyMOL. b) Sequence of HSA, reproducing the colour code of a). Pairs of consecutive Cys residues involved in a disulfide bond are shown with the same yellow tint.

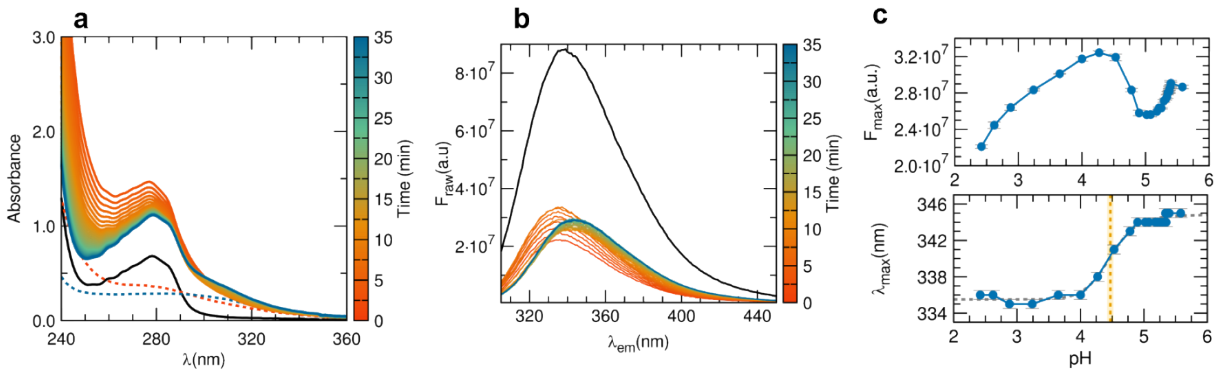

Figure S2. a) UV-vis absorbance spectra of HSA (0.055 mM) after the addition of 20 mM nitroacetic acid (**1**) as a function of time, as indicated by the colour bar (T = 25°C). The dotted lines are spectra of the blank without protein at the shortest and longest time after addition of acid **1**. b) Raw fluorescence emission spectra before applying the inner-filter effect correction. The spectrum of the initial stock solution is reported as a black line. c) Maximum fluorescence intensity (F_{max}, upper panel) and position (λ_{max}, lower panel) from raw spectra as a function of pH.

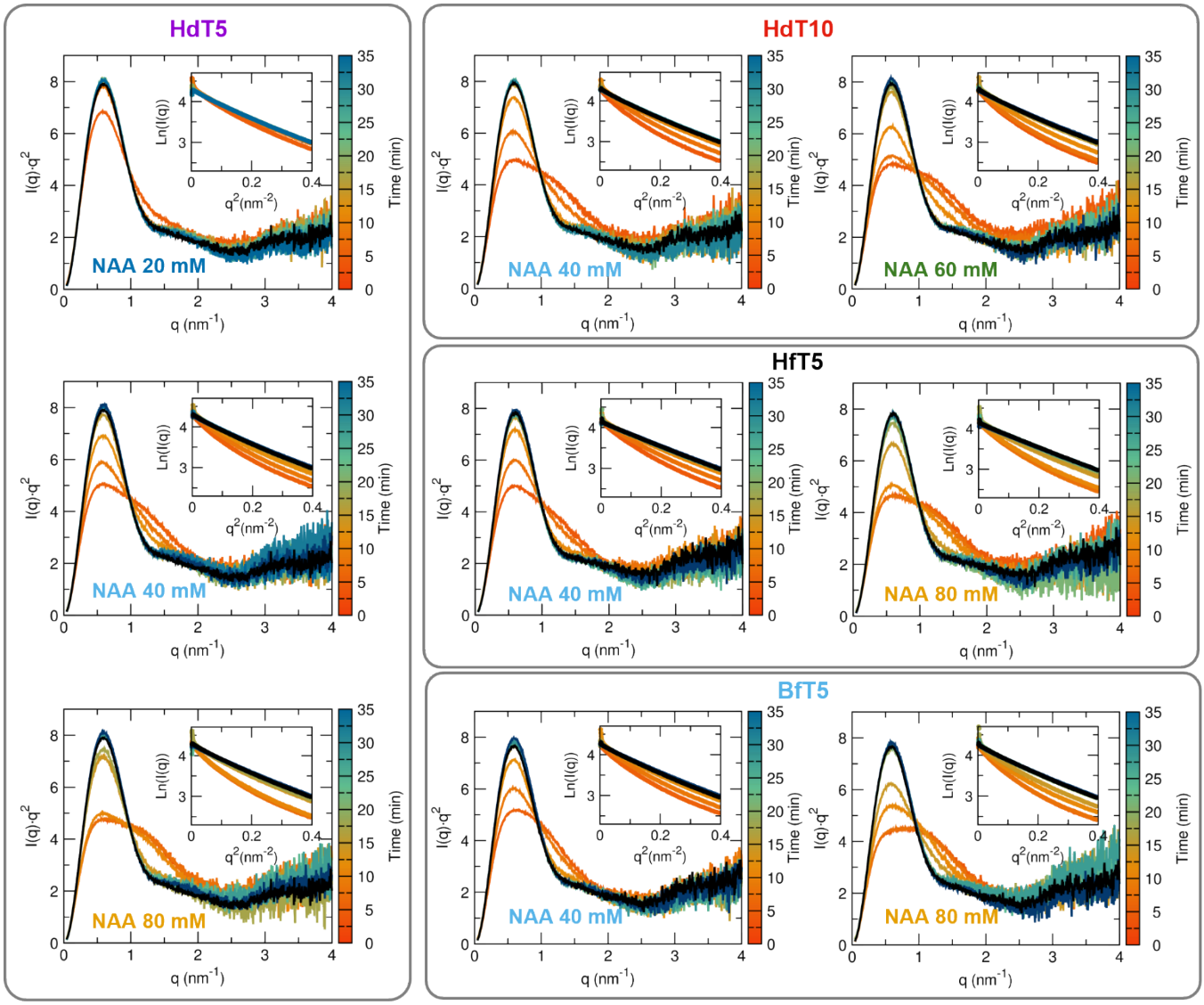

Figure S3. SAXS data collected during the refolding kinetics of different albumin preparations (defatted HSA in TrisHCl 5 mM (HdT5), defatted HSA with double concentration of TrisHCl buffer (HdT10), fatted BSA (BfT5), fatted HSA (HfT5) with addition of different concentrations of nitroacetic acid (NAA), in the form of Kratky plot and, in the inset, Guinier plot.

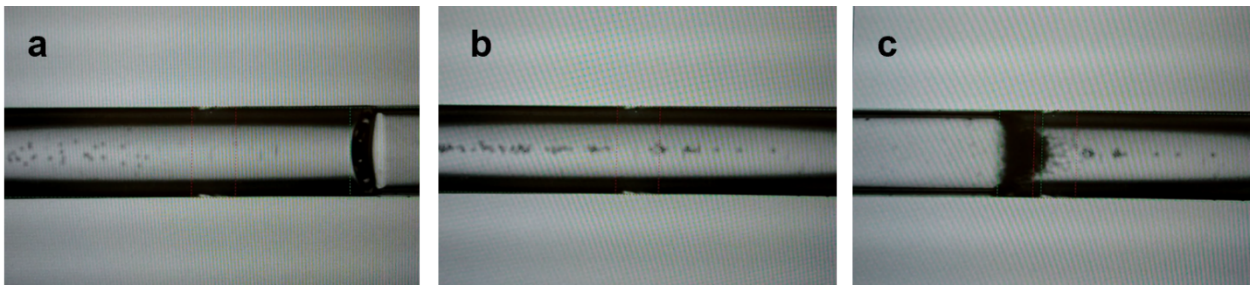

Figure S4. Pictures of the SAXS capillary during flow of a HdT10 sample with 40 mM nitoracetic acid initial concentration, as seen after 7'26''-7'40'' from initial mixing. The red dashed rectangle approximately signals the position where the x-ray beam hits the capillary, while the green dashed rectangle identifies the meniscus position which moves from left to right during flow, according to the automatic image analysis of the acquisition software. a) initial meniscus is visible b) no meniscus c) final meniscus. Pictures taken from the control desktop of the automatic sample changer in operation at BM29.

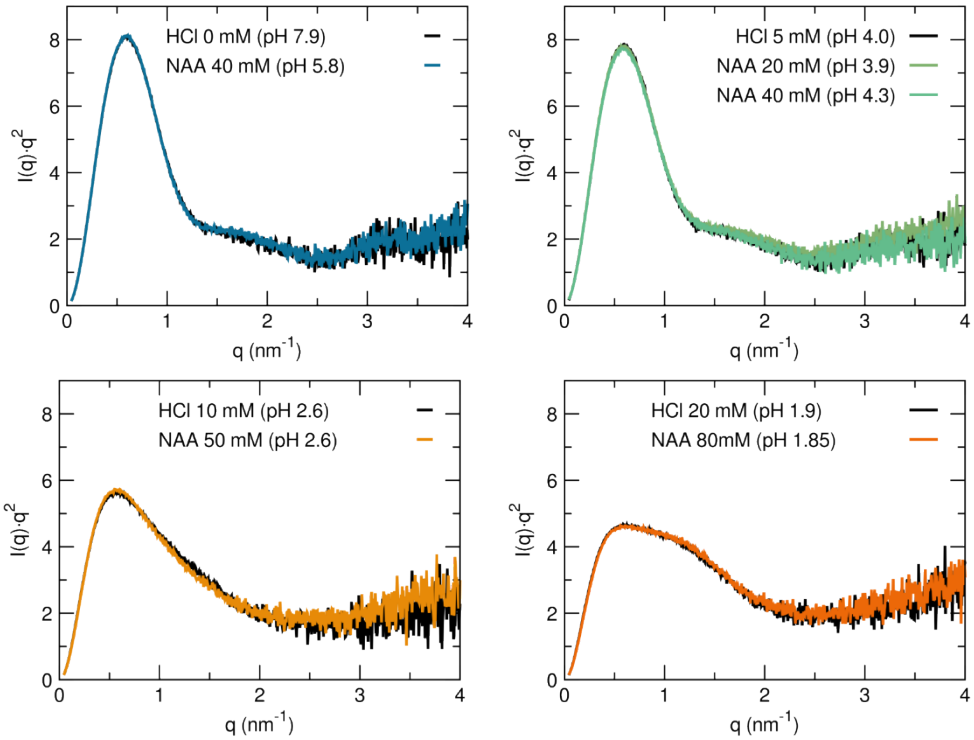

Figure S5. Comparison between the SAXS data collected for HdT5 during the nitroacetic acid (NAA)-induced refolding kinetics at some selected pH values deduced from off-line pH measurements and for HdT5 acidified with static additions of HCl and 1h equilibration.

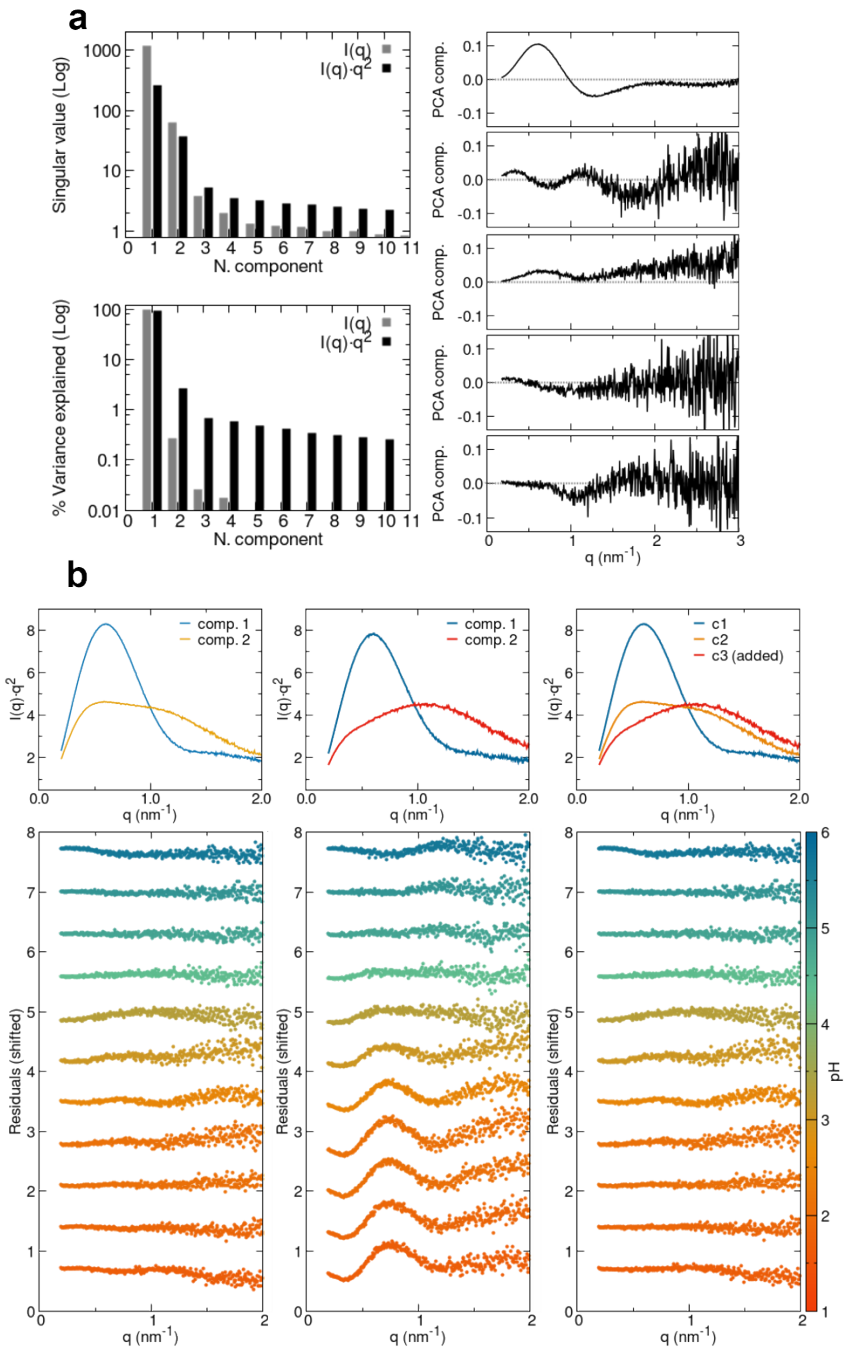

Figure S6. Best-fit of SAXS data as a linear combination of components. a) Results of the PCA(SVD) analysis of the HfT5 refolding dataset: eigenvalues, % of variance explained including each additional component (both on Log scale) and the first 5 eigenvectors (from top to bottom). b) The components used in the linear combination are shown on top, and the residuals between experimental and optimized intensity are shown below: left panel - modelling considering the freely optimized components with MCR-ALS, middle panel - considering the native optimized component plus a very extended conformation as second component, right panel - the three-component assumption as reported in Fig. 5b of the main text.

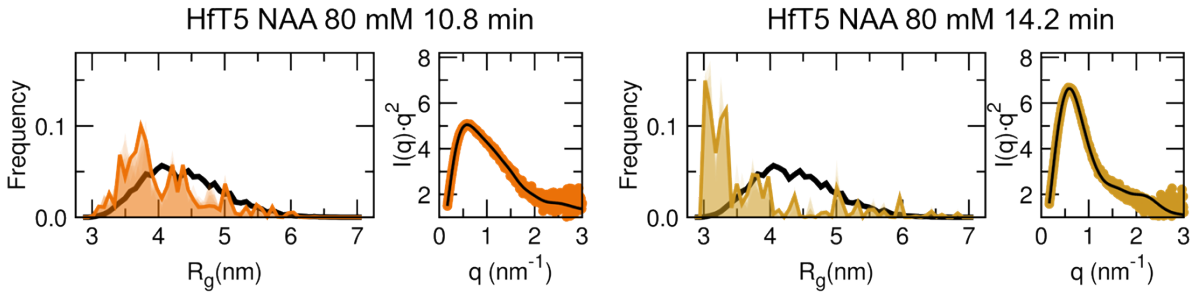

Figure S7. Additional EOM fits for selected SAXS profiles obtained during the time-dependent refolding of the sample HfT5 after acidification with 80 mM nitroacetic acid.

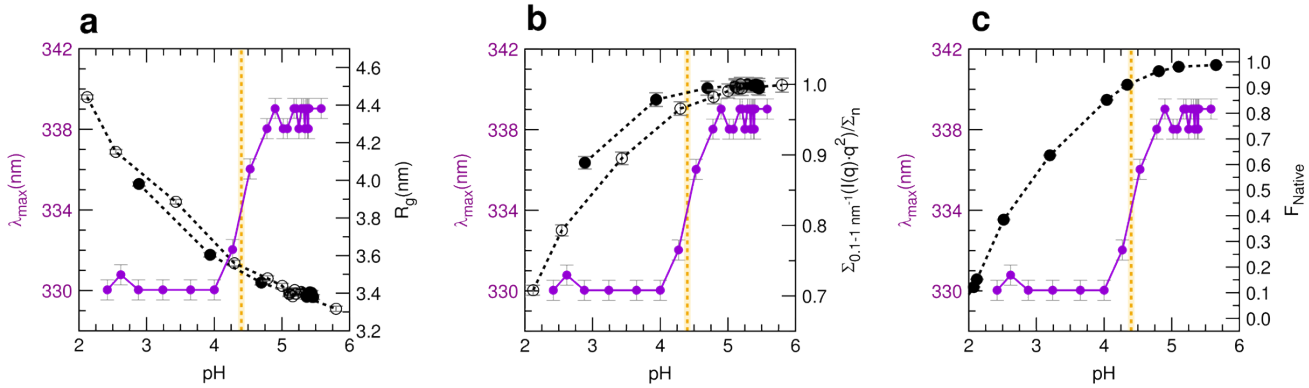

Figure S8. Comparison between the experimental parameters derived from fluorescence and SAXS to monitor albumin refolding as a function of pH: the position of the tryptophan emission maximum (purple points, sample HdT5 NAA 20 mM) is compared to a) the SAXS radius of gyration (samples HdT5 NAA 20 mM and 40 mM); b) the degree of refolding as measured by normalized subtended area of the SAXS data in the form of Kratky plot (samples HdT5 NAA 20 mM and 40 mM); c) the fraction of the native form as optimized by least square fitting of SAXS data as linear combination of 3 scattering components (samples HfT5 40 and 80 mM).

Table S1. SAXS data collection parameters.

Source, instrument	**BM29 (ESRF)**
Detector	PILATUS 2M
Beam geometry (mm²)	0.2 x 0.2
Wavelength (Å)	0.9919
sample-to-detector distance (m)	2.813
q-measurement range (nm^{-1})	0.044-5.2
Absolute scaling method	water scattering I(0)= 0.01632 cm^{-1} at 20°C
Capillary path length (mm)	1
Sample volume (µl)	50/70
Exposure time (s)	1
Number of exposures	10
Extra flow time (s)	10
Sample temperature (ºC)	25°C

Table S2. Details on the albumin samples analysed by SAXS. Concentrations were measured by UV absorbance at 280 nm only on the stock solutions and were considered constant upon addition of solid nitroacetic acid (NAA). A dilution factor of 0.9 was considered for the HCl-acidified samples considering that 50 µl of acid were added to 450 µl of albumin solution.

Sample ID	Protein	Conc. (mM)	Conc. (mg/ml)	Buffer	Nitroacetic acid (mM)
HdT5	Sigma A3782	0.055	3.64	Tris/TrisHCl 5 mM NaCl 0.15 M	-
HdT5_NAA_20mM	Sigma A3782	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	20
HdT5_NAA_40mM	Sigma A3782	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	40
HdT5_NAA_50mM	Sigma A3782	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	50
HdT5_NAA_80mM	Sigma A3782	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	80
HdT10	Sigma A3782	0.057	3.77	Tris/TrisHCl 10mM NaCl 0.15 M	-
HdT10_NAA_40mM	Sigma A3782	-	-	Tris/TrisHCl 10mM NaCl 0.15 M	40
HdT10_NAA_60mM	Sigma A3782	-	-	Tris/TrisHCl 10mM NaCl 0.15 M	60
HfT5	Sigma A8763	0.057	3.77	Tris/TrisHCl 5 mM NaCl 0.15 M	-
HfT5_NAA_40mM	Sigma A8763	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	40
HfT5_NAA_80mM	Sigma A8763	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	80
BfT5	Sigma 05470	0.057	3.79	Tris/TrisHCl 5 mM NaCl 0.15 M	-
BfT5_NAA_40mM	Sigma 05470	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	40
BfT5_NAA_80mM	Sigma 05470	-	-	Tris/TrisHCl 5 mM NaCl 0.15 M	80
Sample ID	**Protein**	**Conc. (mM)**	**Conc. (mg/ml)**	**Buffer**	**HCl (mM)**
HdT5_HCl0	Sigma A3782	0.049	3.28	Tris/TrisHCl 5 mM NaCl 0.15 M	0
HdT5_HCl0.005	Sigma A3782	0.049	3.28	Tris/TrisHCl 5 mM NaCl 0.15 M	5
HdT5_HCl0.008	Sigma A3782	0.049	3.28	Tris/TrisHCl 5 mM NaCl 0.15 M	8
HdT5_HCl0.01	Sigma A3782	0.049	3.28	Tris/TrisHCl 5 mM NaCl 0.15 M	10
HdT5_HCl0.02	Sigma A3782	0.049	3.28	Tris/TrisHCl 5 mM NaCl 0.15 M	20
Hf_E*	Sigma A8763	0.050	3.32	Na_2HPO_4/NaH_2PO_4 5 mM	45

*Collected before ESRF-EBS upgrade (DOI10.15151/ESRF-ES-61665128). Mixing protocol: 200 µl albumin solution in 10 mM buffer + 200 µl acid in water.

References cited in the Supporting Information:

1 S. F. Vanler, G. Larouche, R. P. Wurz and A. B. Charette, *Org. Lett.*, 2010, **12**, 672–675.

2 C. N. Pace, F. Vajdos, L. Fee, G. Grimsley and T. Gray, *Protein Sci.*, 1995, **4**, 2411–2423.

3 M. D. Tully, J. Kieffer, M. E. Brennich, R. Cohen Aberdam, J. B. Florial, S. Hutin, M. Oscarsson, A. Beteva, A. Popov, D. Moussaoui, P. Theveneau, G. Papp, J. Gigmes, F. Cipriani, A. McCarthy, C. Zubieta, C. Mueller-Dieckmann, G. Leonard and P. Pernot, *J. Synchrotron Radiat.*, 2023, **30**, 258–266.

4 A. Round, F. Felisaz, L. Fodinger, A. Gobbo, J. Huet, C. Villard, C. E. Blanchet, P. Pernot, S. McSweeney, M. Roessle, D. I. Svergun and F. Cipriani, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2015, **71**, 67–75.

5 K. Manalastas-Cantos, P. V. Konarev, N. R. Hajizadeh, A. G. Kikhney, M. V. Petoukhov, D. S. Molodenskiy, A. Panjkovich, H. D. T. Mertens, A. Gruzinov, C. Borges, C. M. Jeffries, D. I. Svergun and D. Franke, *J. Appl. Crystallogr.*, 2021, **54**, 343–355.

6 J. Jaumot, A. de Juan and R. Tauler, *Chemom. Intell. Lab. Syst.*, 2015, **140**, 1–12.

7 F. Herranz-Trillo, M. Groenning, A. van Maarschalkerweerd, R. Tauler, B. Vestergaard and P. Bernadó, *Structure*, 2017, **25**, 5–15.

8 P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch and D. I. Svergun, *J. Appl. Crystallogr.*, 2003, **36**, 1277–1282.

9 G. Tria, H. D. T. Mertens, M. Kachala and D. I. Svergun, *IUCrJ*, 2015, **2**, 207–17.

10 S. Sugio, A. Kashima, S. Mochizuki, M. Noda and K. Kobayashi, *Protein Eng.*, 1999, **12**, 439–46.

11 C. Leggio, L. Galantini and N. V. Pavel, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6741–6750.

12 L. Galantini, C. Leggio, P. V. Konarev and N. V. Pavel, *Biophys. Chem.*, 2010, **147**, 111–122.

13 A. Del Giudice, C. Leggio, N. Balasco, L. Galantini and N. V Pavel, *J. Phys. Chem. B*, 2014, **118**, 10043–10051.

14 A. Del Giudice, C. Dicko, L. Galantini and N. V. Pavel, *J. Phys. Chem. B*, 2017, **121**, 4388–4399.