Supporting Information

It takes a dimer to tango: Oligomeric small heat shock proteins dissociate to capture substrate

Indu Santhanagopalan[1,§], Matteo T. Degiacomi[2,3], Dale A. Shepherd[2,‡], Georg K.A. Hochberg[2,†] Justin L.P. Benesch[2] and Elizabeth Vierling[1]

From the [1]Department of Biochemistry and Molecular Biology, University of Massachusetts Amherst, Amherst MA 01003 USA; [2] Department of Chemistry, Physical & Theoretical Chemistry, Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford, OX1 3TA, U.K; [3] Department of Chemistry, Durham University, South Road, Durham, DH1 3LE, U.K.

Running title: sHsp oligomer dynamics in substrate interaction

Present addresses: [§] Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, U.K; [‡] Waters Corporation, Stamford Avenue, Wilmslow, SK9 4AX, U.K; [†] Department of Ecology and Evolution, University of Chicago, Illinois 60637 U.S.

To whom correspondence must be addressed: Elizabeth Vierling, Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, MA 01003. Tel. (413) 577-2890 Fax. (413) 545-3291 vierling@biochem.umass.edu

Contains:
Supporting materials and methods
Supporting references
Table S1-S3
Figures S1-S5

SUPPORTING MATERIALS AND METHODS

Modelling of Ta16.9 and Ps18.1 dodecamers

To build models of Ta16.9 dodecamers, we employed an approach, based on that described previously (1), by combining a protomeric building-block with a polyhedral scaffold . As building block we used the ACD dimer with a C-terminal peptide bound to each monomer,extracted from crystal structure 1GME (2) (residues 43 to 137, and 144-151). The tetrahedral assembly space of this building block was subsequently scanned systematically by exploiting a unit-sized tetrahedral scaffold (a list of four vertices, along with their connectivity). A multimeric tetrahedral arrangement can be obtained by aligning on the scaffold edge a building block so that its first axis of inertia is superimposed on the edge. Two degrees of freedom can then be controlled: a radial expansion of the whole tetrahedron (radius), and a rotation of each individual subunit around their inertia tensor (angle). A total of 990 initial models were produced, with radii from 40 and 50 Å (in steps of 1 Å), and angles between 0 and 180° (with 2° steps). We only selected the 608 assemblies with a negative van der Waals energy (9-6 Lennard-Jones), i.e. featuring protein contacts, but no atomic clashes, for subsequent analysis.

For all energetically favorable models, we calculated the CCS using IMPACT (3), and rescaled the resulting value by 1.14 to obtain a relevant number to scale against experimental data (4). Only models having a CCS within an error bound of 4.8% (i.e. 8629 ± 414 $Å^{2}$) derived from a combination of experimental (standard deviation of multiple charge states and three repeats) and computational uncertainties were considered acceptable. The models that passed this test were assessed for connectivity between dimers by measuring the solvent accessible path distance between the C-termini of the ACDs and the N-termini of bound peptides using DynamXL (5). We only accepted those models for which the N-terminus of each C-terminal peptide was smaller than a specified cut-off distance from the closest ACD C-terminus. The cut-off distance was set at 22.8 Å, the longest distance the missing amino acids between ACD and peptide could span (based on a maximum length for a single amino acid of 3.8 Å and six missing amino acids, 138-143). Furthermore, only cases where every ACD is associated to exactly one peptide would be accepted. Between the selected models, representatives were identified by means of hierarchical clustering using a 2-Å RMSD clustering threshold. This resulted in 4 models that respected our combined selection criterion. In all of these, we added the missing amino acids connecting each ACD to its cognate C-terminus using MODELLER (6).

The volume of the central cavity of each assembly was calculated in two steps. First, we determined the volume of the largest sphere fitting inside the cavity without clashing with any atom (distance less than 2.3 Å). To estimate the volume of the remainder of the cavity, we built a collection of larger, concentric spheres, the largest having a radius equivalent to the tetrahedron's circumradius. All of these spheres were sampled with equally spaced beads (i.e. distributed according to a golden sphere distribution), each having a 1.9 Å radius. The number of beads sampling each sphere was determined so that beads would touch each other. Concentric spheres were spaced by 3.8 Å i.e. twice their sampling beads' radii. We calculated the volume occupied by the beads not clashing with any protein atom, and summed it to the volume of the largest sphere that we could accommodate inside the cavity.

SUPPORTING REFERENCES

1. Baldwin, A. J., Lioe, H., Hilton, G. R., Baker, L. A., Rubinstein, J. L., Kay, L. E., and Benesch, J. L. (2011) The polydispersity of alphaB-crystallin is rationalized by an interconverting polyhedral architecture. *Structure* **19**, 1855-1863

2. van Montfort, R. L., Basha, E., Friedrich, K. L., Slingsby, C., and Vierling, E. (2001) Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat Struct Biol* **8**, 1025-1030

3. Marklund, E. G., Degiacomi, M. T., Robinson, C. V., Baldwin, A. J., and Benesch, J. L. (2015) Collision cross sections for structural proteomics. *Structure* **23**, 791-799

4. Benesch, J. L., and Ruotolo, B. T. (2011) Mass spectrometry: come of age for structural and dynamical biology. *Curr Opin Struct Biol* **21**, 641-649

5. Degiacomi, M. T., and Dal Peraro, M. (2013) Macromolecular symmetric assembly prediction using swarm intelligence dynamic modeling. *Structure* **21**, 1097-1106

6. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713

7. Hochberg, G. K. A., Shepherd, D. A., Marklund, E. G., Santhanagoplan, I., Degiacomi, M. T., Laganowsky, A., Allison, T. M., Basha, E., Marty, M. T., Galpin, M. R., Struwe, W. B., Baldwin, A. J., Vierling, E., and Benesch, J. L. P. (2018) Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions. *Science* **359**, 930-935

Table S1: Residue positions chosen for mutagenesis in Ta 16.9

Residue pairs	% Accessiblities[#]		Distance (Å)[$]	
N_i/Ch X --- N_j/Ch Y	N_i/X	N_j/Y	Cα- Cα	Cβ-Cβ
Ta16.9 mutants				
W48/Ch A - H97/Ch B	0	9.6	5.4	4.3
K49/Ch A - W96/Ch B	13.5	21.5	4.6	5.8
V73/Ch A - K145/Ch L	2.8	25.6	5.3	4.8
E74 /Ch A - V144/Ch L	4.6	46.5	5.2	5.6

[#] % Accessibilities were calculated from the coordinates of PDB file 1GME as described in the methods section of main text.

[$] Distances were calculated using the distance measurement tool of Pymol.

Table S2: DTNB assay measurements for calculation of free thiols in sHsp molecules

Protein	Absorbance at 412 nm	[-SH] in μM	Number of thiols per sHsp monomer[$]
Ta16.9	0.005 ± .002	0.38	0.02
Ta$_{CT-ACD}$	0.006 ± .002	0.44	0.03
Ta$_{dimer}$	0.020 ± .003	1.57	0.10
Ta$_{V144C}$	0.189 ± .004	13.8	0.92
Ps18.1	0.007 + .002	0.52	0.04
Ps$_{CT-ACD}$	0.010 ± .004	0.71	0.05
Ps$_{V151C}$	0.176 ± .003	12.8	0.85
Ta$_{CT-ACD}$ + 20 mM DTT*	0.435 ± .010	31.8	2.12
Ta$_{dimer}$+ 20 mM DTT*	0.427 ± .005	31.2	2.08
Ps$_{CT-ACD}$ + 20 mM DTT*	0.467 ± .010	34.1	2.27
Ta$_{CT-ACD}$ + 0.5 mM TCEP*	0.011 ± .004	0.80	0.05
Ta$_{dimer}$ + 0.5 mM TCEP*	0.398 ±.010	29.0	1.93
Ps$_{CT-ACD}$ + 0.5 mM TCEP*	0.021 ±.007	1.53	0.10

* Samples treated with reducing agent (10 mM DTT/0.5 mM TCEP) were desalted into 10 mM sodium acetate buffer of pH 4.0, prior to treatment with DTNB in the presence 100 mM Tris pH 8.0, and 4M urea.

[$] sHsp monomer concentration was 15 μM.

Table S3: Normalized total scattering intensity determined by DLS at different temperatures

Protein[$]	Temperature °C					
	25	35	45	55	65	25 *
Ta 16.9	1.00 ± 0.03	0.76 ± 0.02	0.52 ± 0.04	3.25 ± 0.63	7.51 ± 0.76	1.18 ± 0.04
$Ta_{CT\text{-}ACD}$	1.00 ± 0.02	0.86 ± 0.02	0.69 ± 0.03	0.82 ± 0.03	2.08 ± 0.05	0.68 ± 0.01
$rTa_{CT\text{-}ACD}$[#]	1.00 ± 0.01	0.76 ± 0.05	0.18 ± 0.01	1.68 ± 0.12	3.60 ± 0.10	5.69 ± 0.16
Ta_{dimer}	1.00 ± 0.01	1.24 ± 0.09	1.05 ± 0.11	3.16 ± 0.61	3.60 ± 0.07	1.05 ± 0.18
rTa_{dimer}[#]	1.00 ± 0.04	0.89 ± 0.34	0.64 ± 0.03	3.72 ± 0.51	9.32 ± 0.23	3.36 ± 0.42
Ps 18.1	1.00 ± 0.10	1.04 ± 0.04	0.39 ± 0.05	1.31 ± 0.10	3.22 ± 0.32	0.85 ± 0.02
$Ps_{CT\text{-}ACD}$	1.00 ± 0.07	0.66 ± 0.02	0.49 ± 0.01	0.36 ± 0.02	0.24 ± 0.01	0.30 ± 0.03
$rPs_{CT\text{-}ACD}$[#]	1.00 ± 0.08	0.99 ± 0.02	1.03 ± 0.15	0.66 ± 0.01	5.72 ± 0.35	5.50 ± 0.56

[$] Protein concentration used is 10 μM. Values were normalized with respect to measurement at 25 °C, which was assigned a value of 1. All measurements are averages of 3 experiments.

* The samples were cooled to 25 °C, after measurements at 65 °C.

[#] The prefix 'r' indicates that measurements were made in the presence of 20 mM DTT.
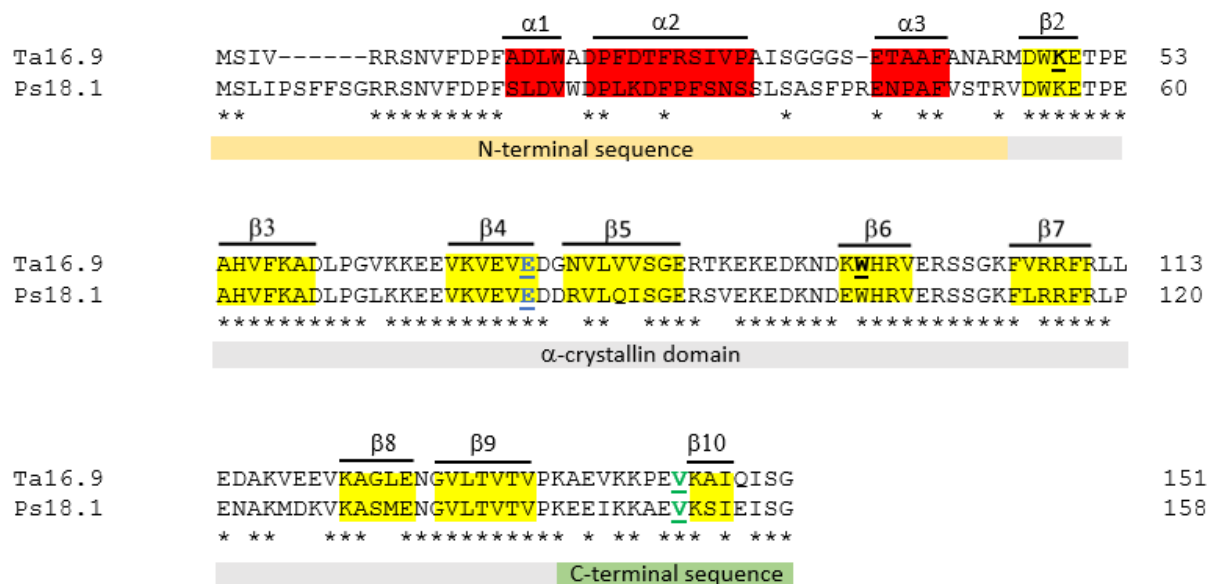
Figure S1: Homology of Ta16.9 and Ps18.1. Alignment of Ta16.9 and Ps18.1 amino acid sequences with helices and strands highlighted in red and yellow respectively based on PDB 1GME of Ta16.9 and homology of Ps18.1. Residues mutated to cysteines for construction of Ta$_{CT-ACD}$ and Ps$_{CT-ACD}$ (green), and Ta$_{dimer}$ (black) are underlined. Identical residues (*) in the two sequences, and stretches of NT, ACD and CT are indicated.
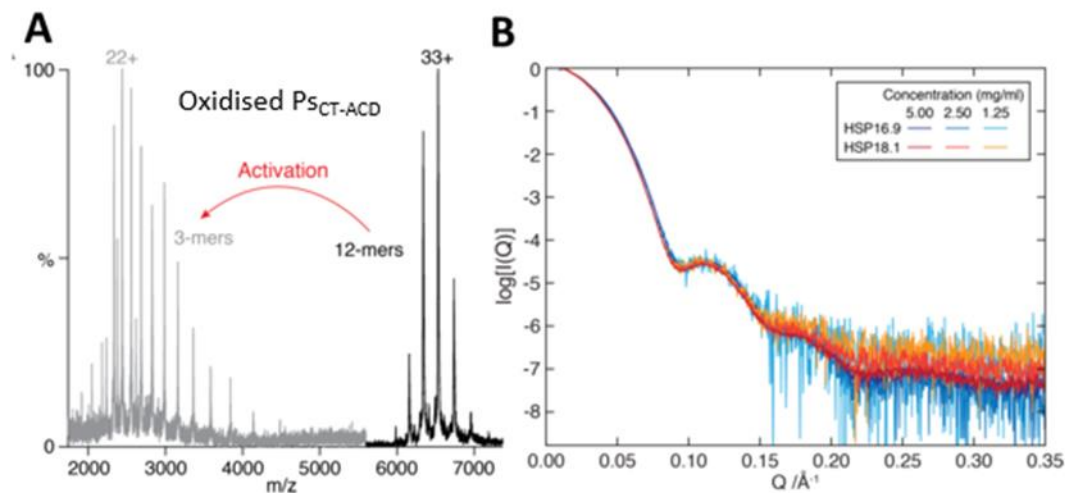
Figure S2: **Ps18.1 and Ta16.9 have comparable dodecameric structures. (A) Mass spectrum of oxidized Ps$_{CT-ACD}$ confirms presence of covalent trimers, consistent with a tetrahedral dodecamer only (7).** Mass spectrum for Ps$_{CT-ACD}$ (Ps18.1) under native (black) and non-native (grey) MS analysis conditions. **(B) Wild type Ta16.9 and Ps18.1 have overlapping SAXS profiles.** Wild type Ta16.9 and Ps18.1 were examined by means of SAXS as the indicated concentrations. Note that we did not use these SAXS data in our modeling of the structures presented in the main text. This is because the models do not include the N-terminal regions of Ta16.9 (almost 30% of the sequence), whereas the SAXS data reflects the pairwise distances between all atoms in the wild-type protein (i.e. 100% of the sequence). Therefore there cannot be a good correspondence between theoretical curves calculated from the models and the experimental measurements, precluding meaningful comparisons.
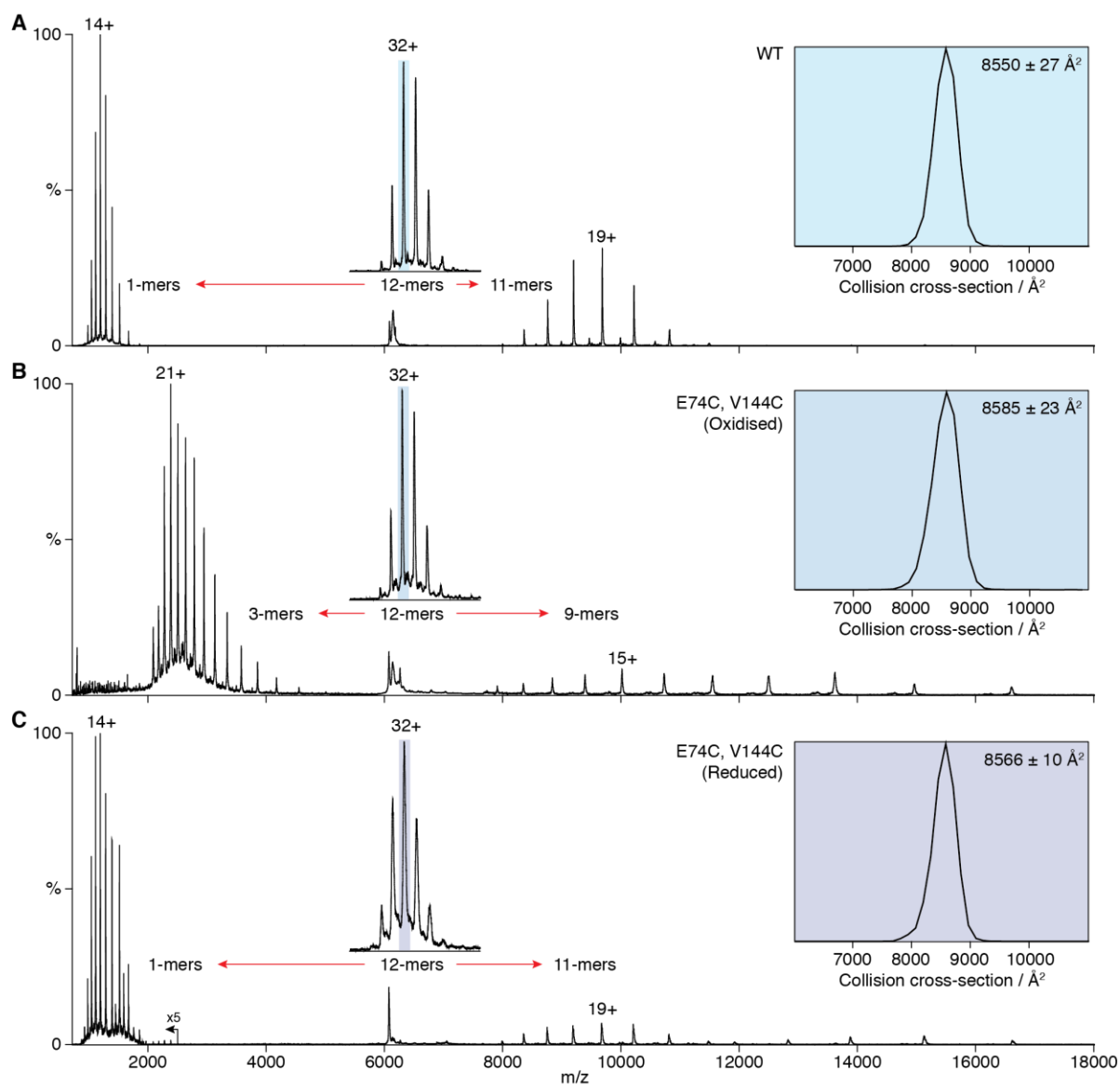
Figure S3: Wild type Ta16.9 (A), and oxidized (B) and reduced (C) Ta_{CT-ACD} compared by mass spectrometry and ion mobility mass spectrometry. Left side of panels shows MS1 of the dodecamers with the 32+ charge state highlighted, which was then dissociated in the MS2 to the constituent covalent substructure. Wild type (A) and reduced Ta_{CT-ACD} dissociate to monomers, while as shown in Figure 1C in the text, oxidized Ta_{CT-ACD} dissociates into trimers. Insets on right side show the collisional cross-section of the corresponding dodecamers.
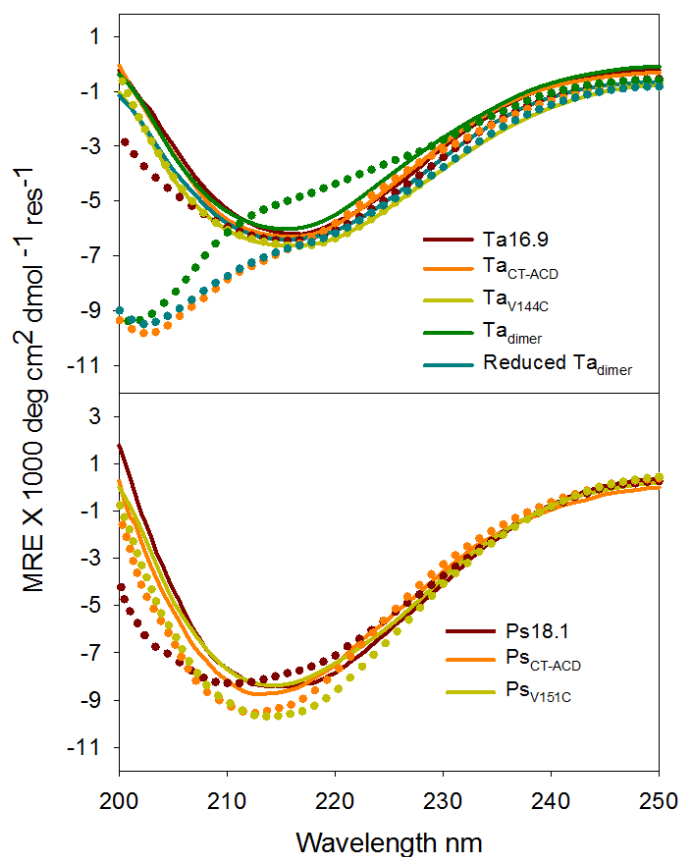
Figure S4: Reversibility of sHsp unfolding. Following heating of 10 μM proteins either to 65 °C (Ta 16.9, Ta$_{V144C}$ and reduced Ta$_{dimer}$), to 75 °C (Ps18.1 and Ps$_{V151C}$) or to 85 °C (Ta$_{CT-ACD}$, Ta$_{dimer}$ and Ps$_{CT-ACD}$), proteins were cooled to 25 °C. Far UV CD spectra collected for the cooled proteins (refolded-dotted lines) are plotted along with those for native proteins (continuous lines) to ascertain the reversibility of thermal unfolding. Legend indicates the colors used for different proteins.
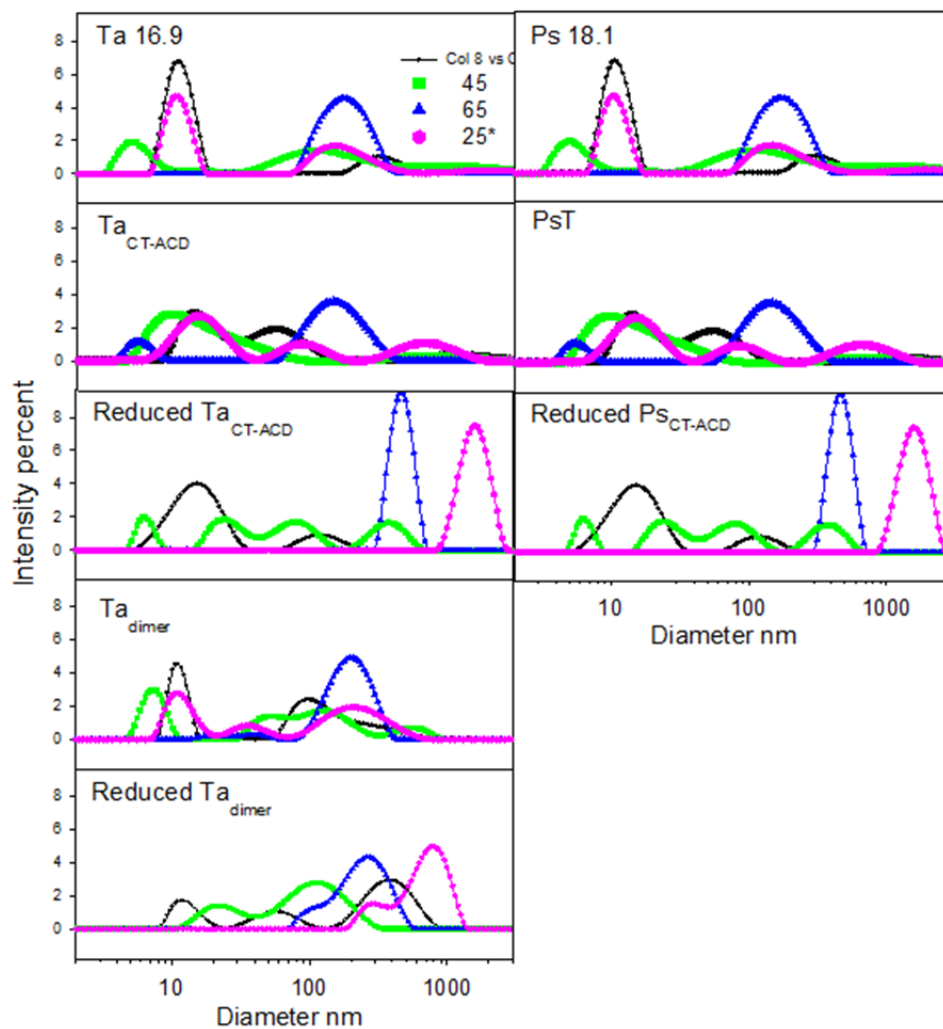
Figure S5: Particle intensity distribution (PID) plots for 10 µM protein were recorded at 25, 45 and 65 °C, and again at 25 °C after cooling from 65 °C (25*). Plots are representative of three experiments.