Figures and figure supplements

Evolution towards simplicity in bacterial small heat shock protein system

Piotr Karaś *et al*.

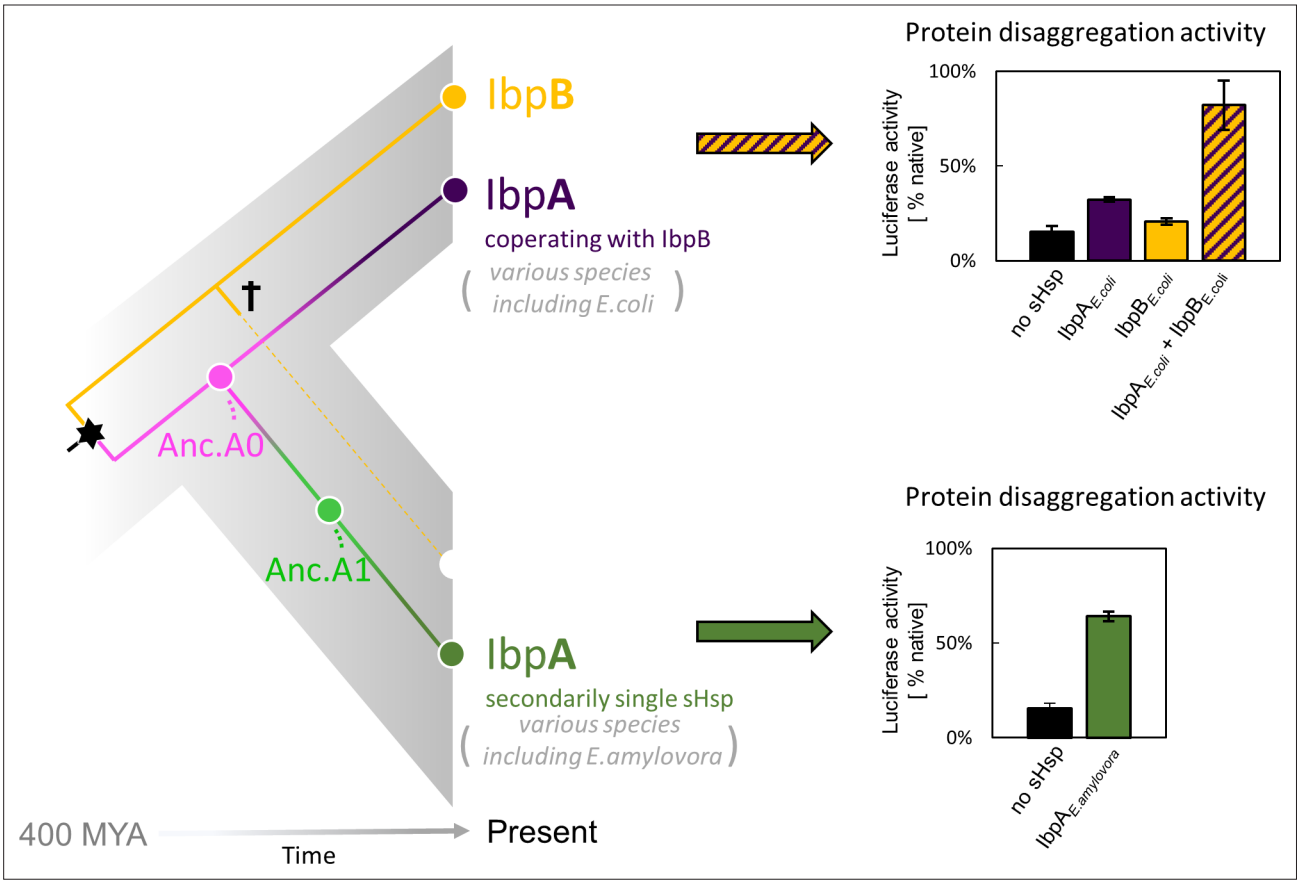

Figure 1. small heat shock protein (sHsp) systems in *Enterobacteriacrae* and *Erwiniaceae*. Left - schematic phylogeny of sHsps in *Enterobacterales*. Gene duplication resulting in IbpA + IbpB two-protein system is marked with a star, while the loss of *ibpB* gene in *Erwiniaceae* clade is marked with a cross. $AncA_0$ - reconstructed last common ancestor of IbpA from *Erwiniaceae* and *Enterobacteriaceae*, expressed as a part of two-protein system. $AncA_1$ - reconstructed last common ancestor of secondarily single IbpA from *Erwiniaceae*. Right - representative extant sHsps' ability to stimulate luciferase refolding. sHsps were present during the luciferase thermal denaturation step. Refolding of denatured luciferase was performed by the Hsp70-Hsp100 chaperone system. Activity of luciferase was measured after 1 hr refolding at 25 °C and shown as an average of 3 repeats. Error bars represent standard deviation.

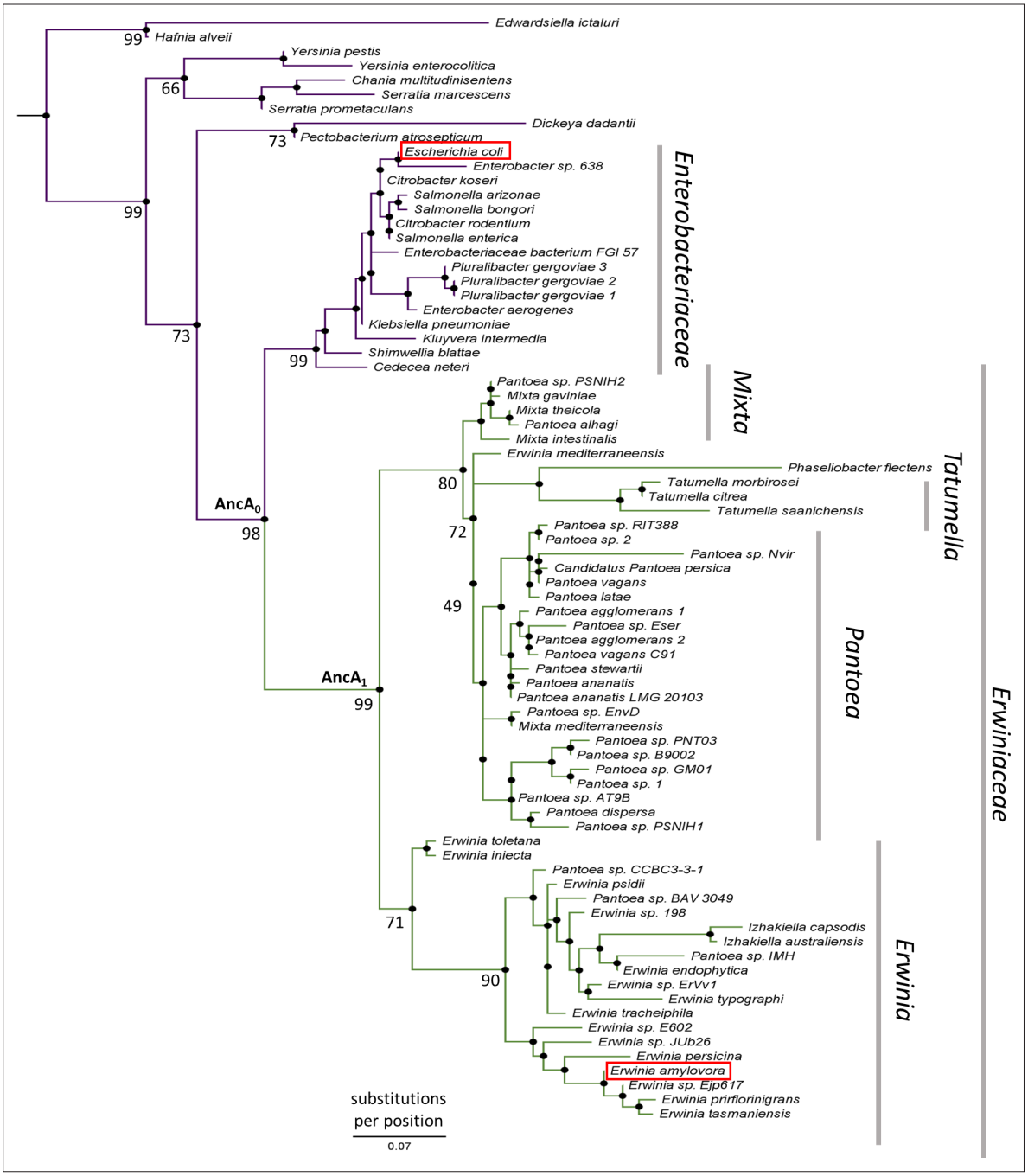

Figure 2. IbpA phylogeny in *Enterobacterales*. Phylogeny was reconstructed from 77 IbpA orthologs from *Enterobacterales* using the Maximum Likelihood algorithm with JTT + R3 substitution model. $AncA_0$ - node representing the last common ancestor of IbpA from *Erwiniaceae* and *Enterobacteriaceae*. $AncA_1$ - node representing the last common ancestor of IbpA from *Erwiniaceae*. Bootstrap support is noted for the major nodes. Extant IbpAs from *E. coli* and *E. amylovora* are marked with a red frame. Scale bar - substitutions per position.

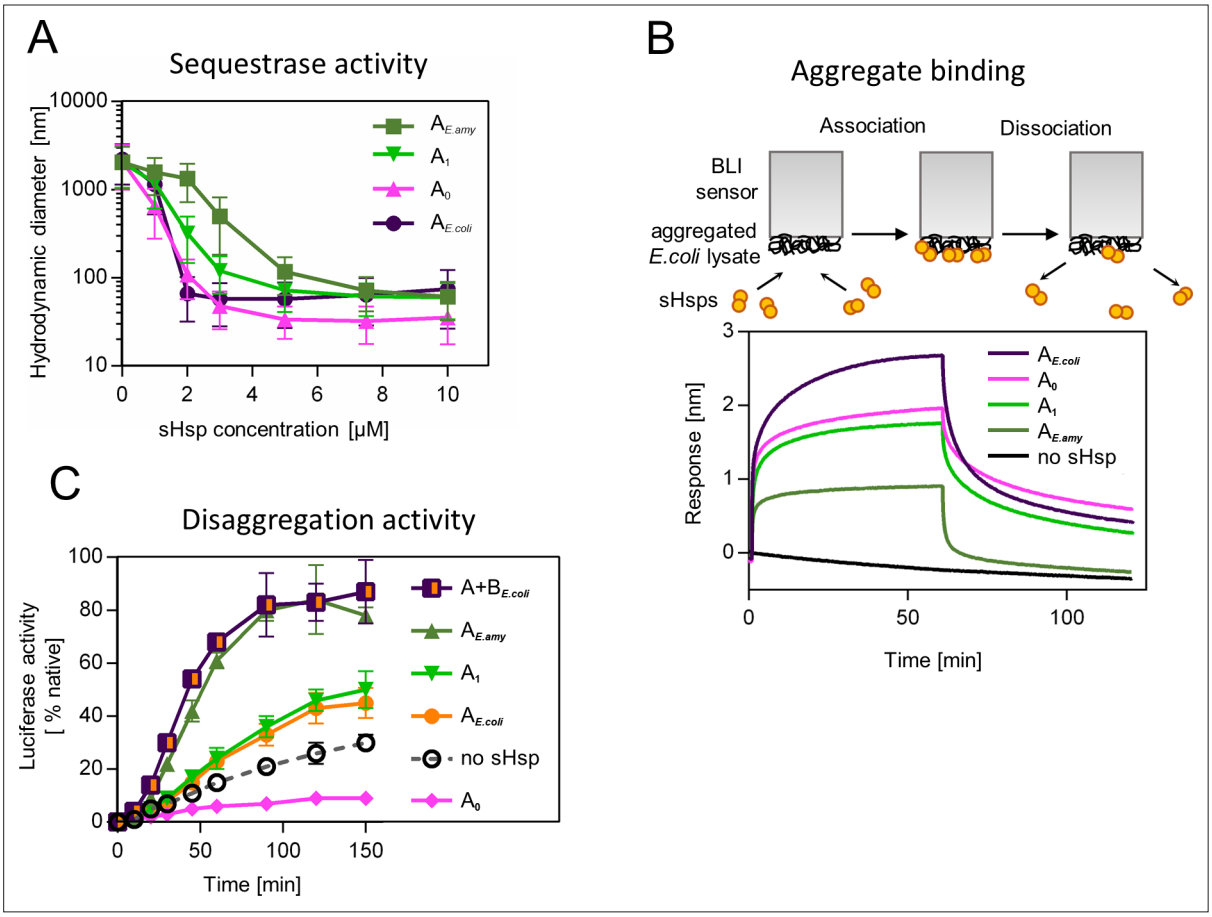

Figure 3. Functional changes during the evolution of secondarily single small heat shock protein (sHsp) in *Erwiniaceae*. (**A**) Sequestrase activity of extant and ancestral sHsps. Luciferase was heat denatured in the presence of different concentrations of sHsps and size of formed sHsps-substrate assemblies was measured by DLS. For every measurement, ten subsequent series of ten 10-s runs were averaged and particle size distribution was calculated by fitting to 70 size bins between 0.4 and 10,000 nm. Results are presented as an average hydrodynamic diameter of measured particles weighted by intensity (Z-average), calculated with Malvern Zetasizer Software 7.13. Error bars represent standard deviation obtained from the distribution. (**B**) Binding of extant and ancestral sHsps to heat-aggregated *E. coli* proteins. *E. coli* proteins were heat aggregated and immobilized on a Biolayer interferometry (BLI) sensor. sHsps were heat-activated before the binding step. (**C**) Extant and ancestral sHsps' ability to stimulate luciferase refolding. Experiment was performed at 25 °C. Luciferase activity at each timepoint was shown as an average of 3 repeats. Error bars represent standard deviation.

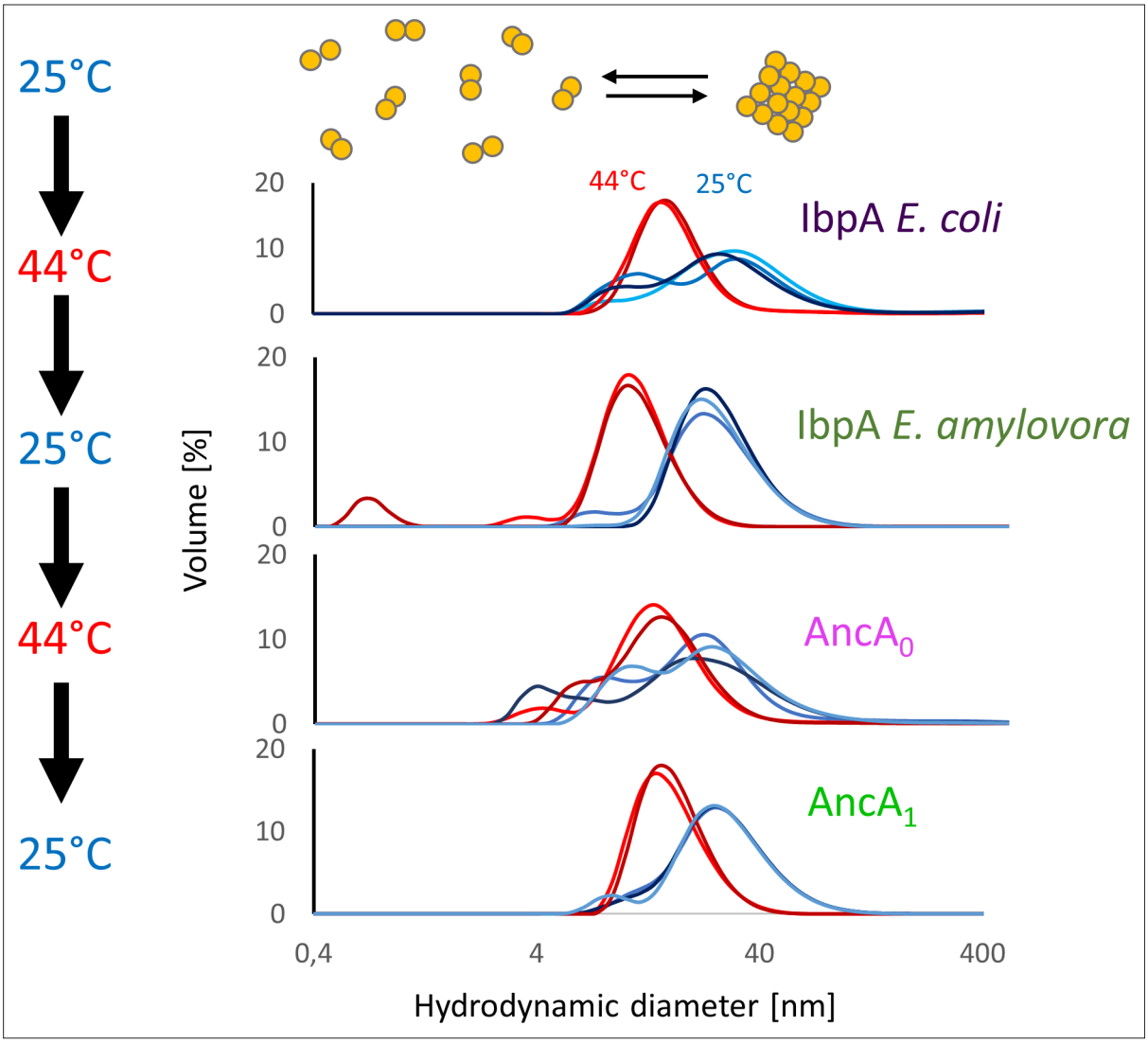

Figure 3—figure supplement 1. Both reconstructed proteins reversibly deoligomerize at heat shock temperature, similarly to extant proteins. Hydrodynamic diameter distributions of small heat shock protein (sHsp) oligomers incubated subsequently at 25 °C (blue graphs) and 44 °C (red graphs). Size distributions were measured by DLS.

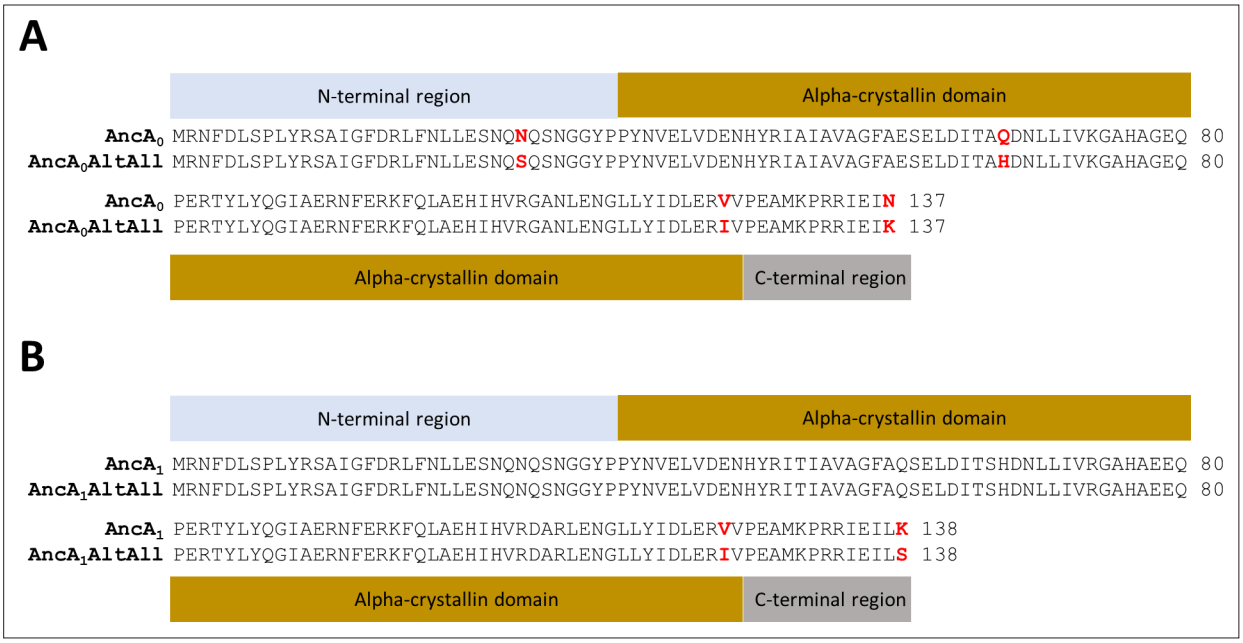

Figure 3—figure supplement 2. Amino acid sequences of most likely (ML) and AltAll variants of reconstructed ancestral proteins. $AncA_0$, $AncA_1$ - ML variants; $AncA_0$AltAll, $AncA_1$AltAll - AltAll variants. Positions differing between most likely (ML) and AltAll variants are marked in red. (**A**) Alignment of $AncA_0$ and $AncA_0$AltAll sequences. (**B**) Alignment of $AncA_1$ and $AncA_1$AltAll sequences.

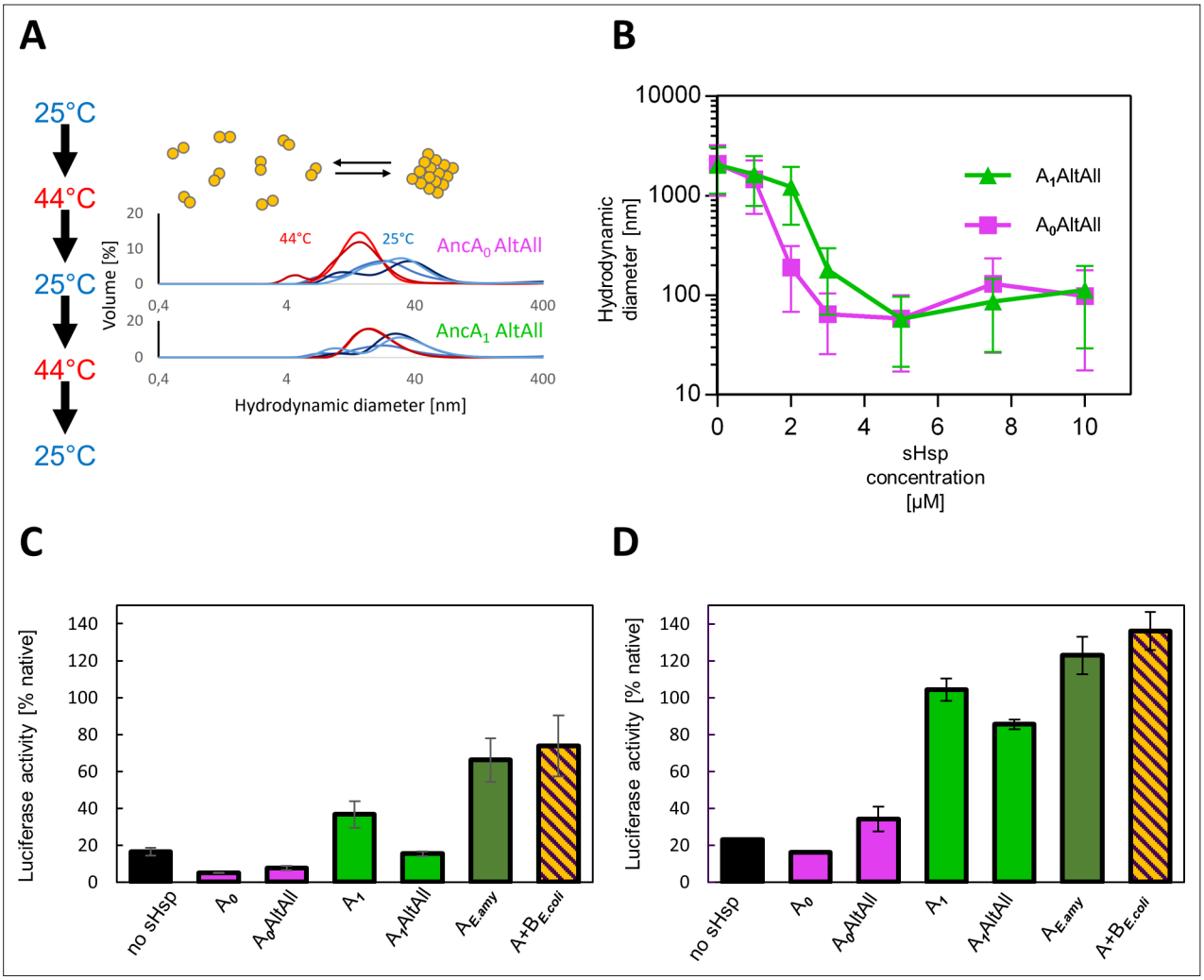

Figure 3—figure supplement 3. AltAll variants of $AncA_0$ and $AncA_1$ have similar properties to most likely (ML) variants. (**A**) Reversible deoligomerization at heat shock temperature; measurement were performed as in ***Figure 3—figure supplement 1***. (**B**) Sequestrase activity. Luciferase sequestration assay was performed as in ***Figure 3B***. For every measurement, ten subsequent series of ten 10-s runs were averaged and particle size distribution was calculated by fitting to 70 size bins between 0.4 and 10,000 nm. Results are presented as an average hydrodynamic diameter of measured particles weighted by intensity (Z-average), calculated with Malvern Zetasizer Software 7.13. Error bars represent standard deviation obtained from the distribution. (**C–D**) Ability to stimulate Hsp70-Hsp100-dependent luciferase disaggregation; Luciferase refolding assay was performed as in ***Figure 1***. Activity of luciferase was measured after 1 hr refolding at 25 °C and shown as an average of 3 repeats. Error bars represent standard deviation; assays were performed at two concentrations of Hsp70 system: (**C**) standard (1 μM DnaK, 0.4 μM DnaJ 0.3 μM GrpE) or (**D**) increased (2 μM DnaK, 0.8 μM DnaJ 0.6 μM GrpE).

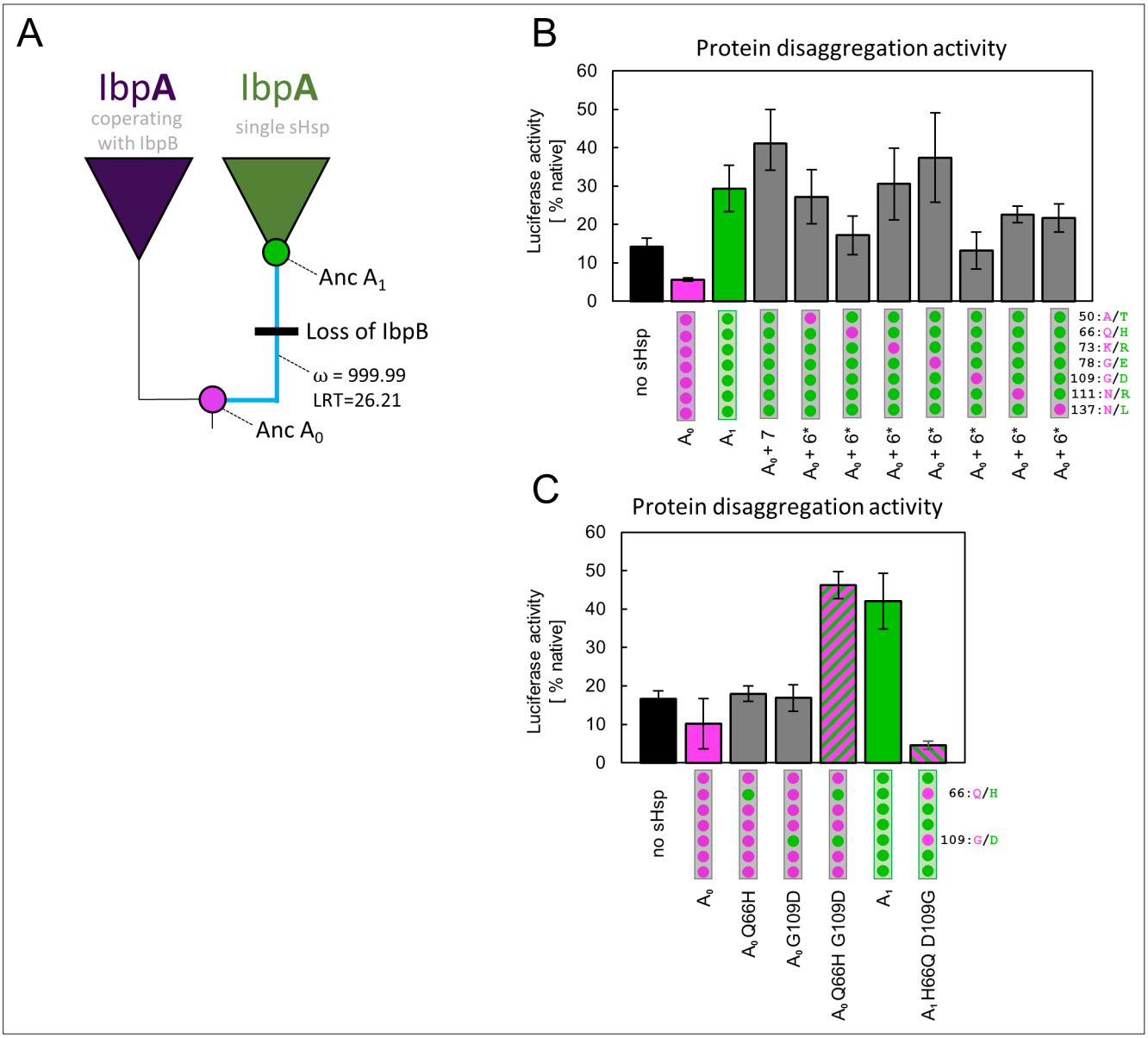

Figure 4. Substitutions at positions 66 and 109 that occurred between nodes A_0 and A_1 are crucial for ancestral small heat shock proteins (sHsps) to work as a single protein. Luciferase refolding assay was performed as in ***Figure 1***. Activity of luciferase was measured after 1 hr refolding at 25 °C and shown as an average of 3 repeats (6 in the case of no sHsp and proteins $AncA_0$ and $AncA_1$ in panel C). Error bars represent standard deviation. (**A**) Schematic phylogeny of *Enterobacterales* IbpA showing increased ratio of nonsynonymous to synonymous substitutions (ω) on the branch between nodes $AncA_0$ and $AncA_1$. Loss of cooperating IbpB is marked on a tree. Value of the Likelihood Ratio Test (LRT) is given for the selection model. (**B**) Identification of substitutions necessary for $AncA_0$ to obtain $AncA_1$-like activity in luciferase disaggregation; seven candidate mutations were introduced into $AncA_0$ ($AncA_0$ + 7); subsequently, in series of six mutants, each of the candidate positions was reversed to a more ancestral state ($AncA_0$ + 6* variants) (**C**) Effect of substitutions at positions 66 and 109 on the ability of $AncA_0$ and $AncA_1$ to stimulate luciferase refolding.

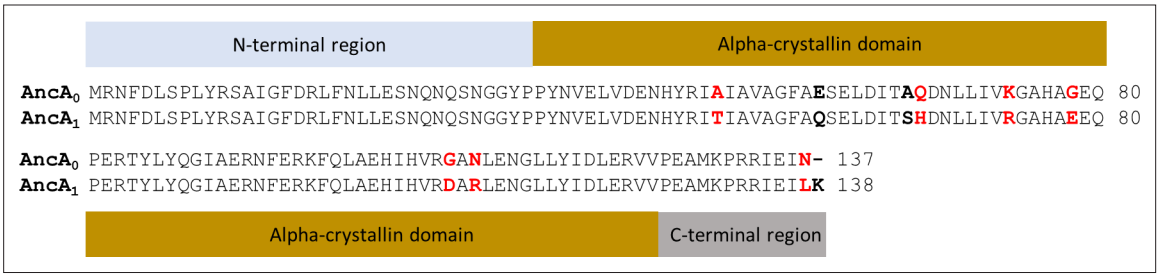

Figure 4—figure supplement 1. Amino acid sequence differences between $AncA_0$ and $AncA_1$: Differing positions used in further analysis were marked in red, an differing positions omitted from further analysis due to low conservation in extant *Erwiniaceae* were marked in bold.

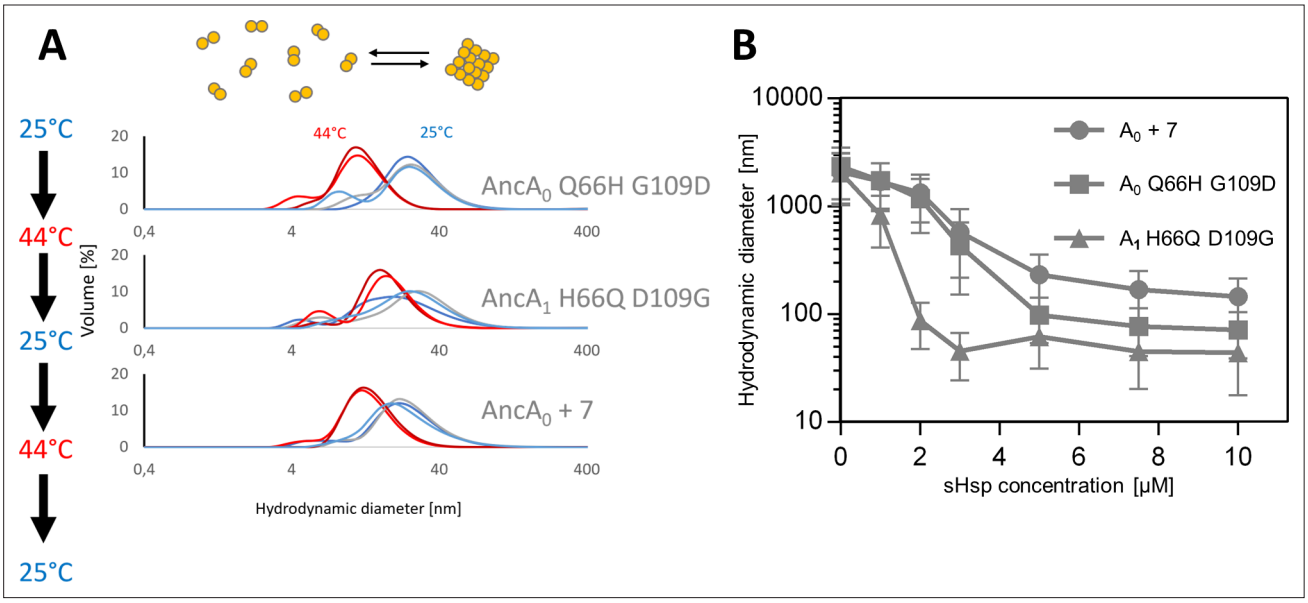

Figure 4—figure supplement 2. $AncA_0$ Q66H G109D, $AncA_0$ + 7 and $AncA_1$ H66Q D109G exhibit sequestrase activity and reversibly deoligomerize at heat shock temperature. (**A**) Reversible deoligomerization at heat shock temperature; measurements were performed as in ***Figure 3—figure supplement 2***. (**B**) Sequestrase activity; luciferase sequestration assay was performed as in ***Figure 3B***. For every measurement, ten subsequent series of ten 10-s runs were averaged and particle size distribution was calculated by fitting to 70 size bins between 0.4 and 10,000 nm. Results are presented as an average hydrodynamic diameter of measured particles weighted by intensity (Z-average), calculated with Malvern Zetasizer Software 7.13. Error bars represent standard deviation obtained from the distribution.

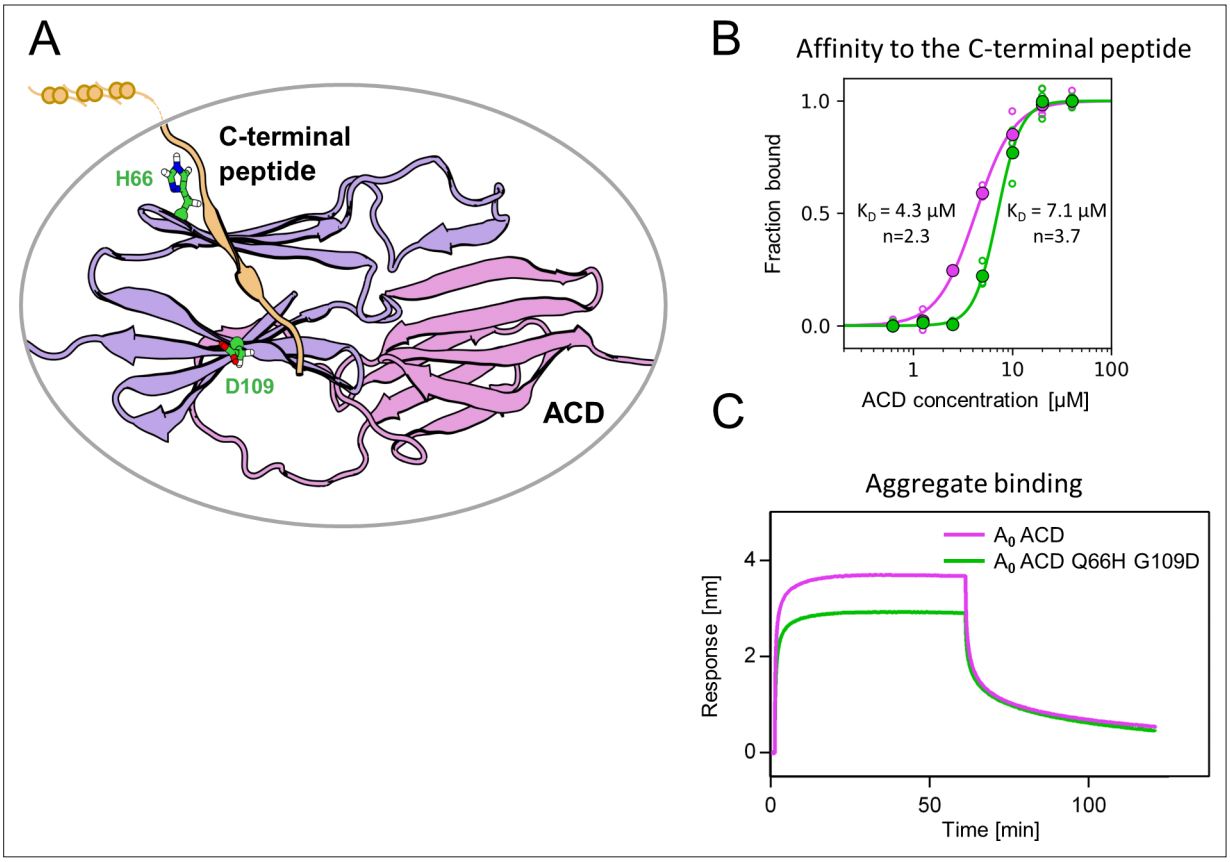

Figure 5. Substitutions at positions 66 and 109 decreased the affinity of $AncA_0$ ACD to C-terminal peptide and aggregated substrate. (**A**) Structural model of complex formed by $AncA_0$ Q66H G109D α-crystallin domain dimer (purple and lilac) and $AncA_0$ C-terminal peptide (orange). (**B**) Effect of Q66H G109D substitutions (green) on $AncA_0$ (purple) ACD's affinity to the C-terminal peptide assayed by Biolayer interferometry (BLI). Biolayer thickness at the end of the association step was used to calculate the fraction of bound peptide. Filled circles represent means of triplicate measurements, individual data points are shown as hollow circles and were fitted to the cooperative binding model (Hill equation). Values of fitted binding affinities [$K_{0.5}$] ($AncA_0$ 4.3 ± 0.2 µM, $AncA_0$ Q66H G109D 7.1 ± 0.2 µM) and Hill coefficients [n] ($AncA_0$ 2.3 ± 0.17, $AncA_0$ Q66H G109D 3.7 ± 0.34) are indicated on the plot. (**C**) Effect of Q66H G109D substitutions on $AncA_0$ ACD's affinity to aggregated *E. coli* proteins bound to BLI sensor. Analysis was performed as in ***Figure 3A***.

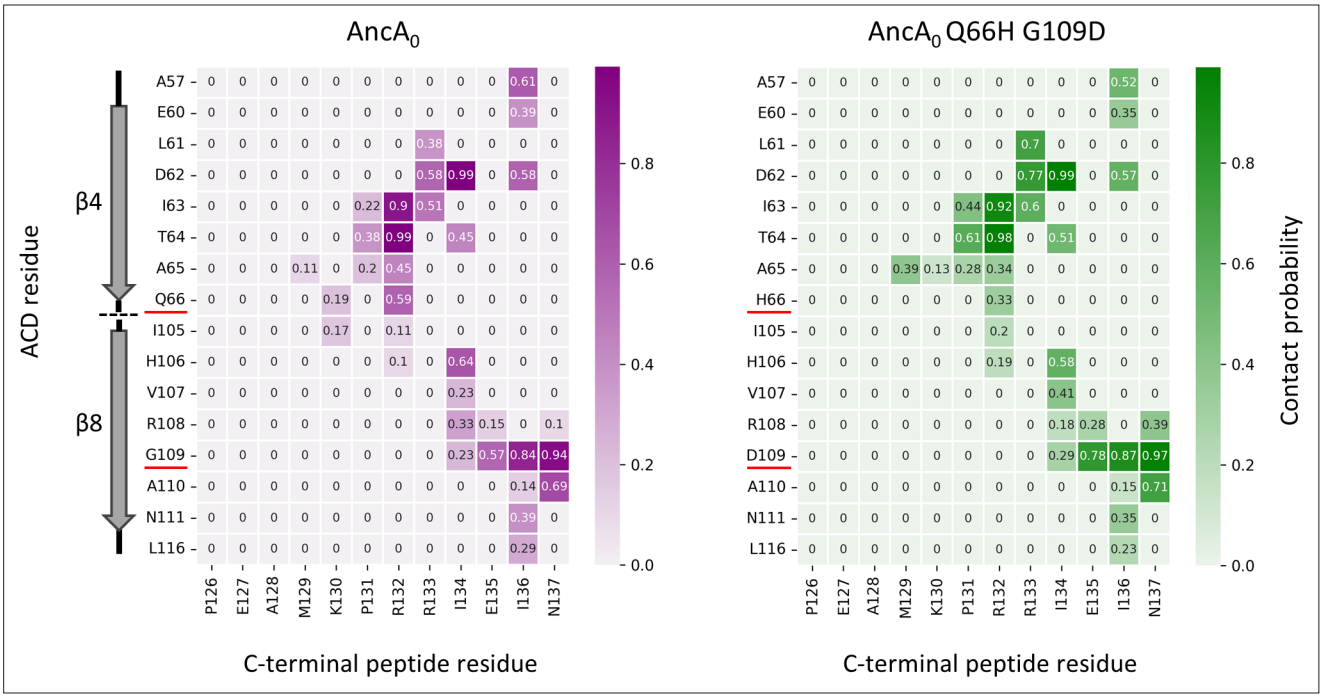

Figure 5—figure supplement 1. $AncA_0$ residues 66 and 109 contact the C-terminal peptide, although overall contact pattern remains similar upon introduction of Q66H G109D substitutions. Heatmaps illustrating contact probabilities, derived from last 400 ns of equilibrium molecular dynamics (MD) simulations, between residues of the α–crystallin domain (ACD) (vertical-axis) and the C-terminal peptide (horizontal axis) for AncA0 (left) and AncA0 Q66H G109D (right). Number within each cell of the heatmap and cell shading represent contact probability. Schematic next to the vertical axis of AncA0 heatmap represents positioning of the ACD residues within secondary structure elements. Substituted ACD residues are underlined.

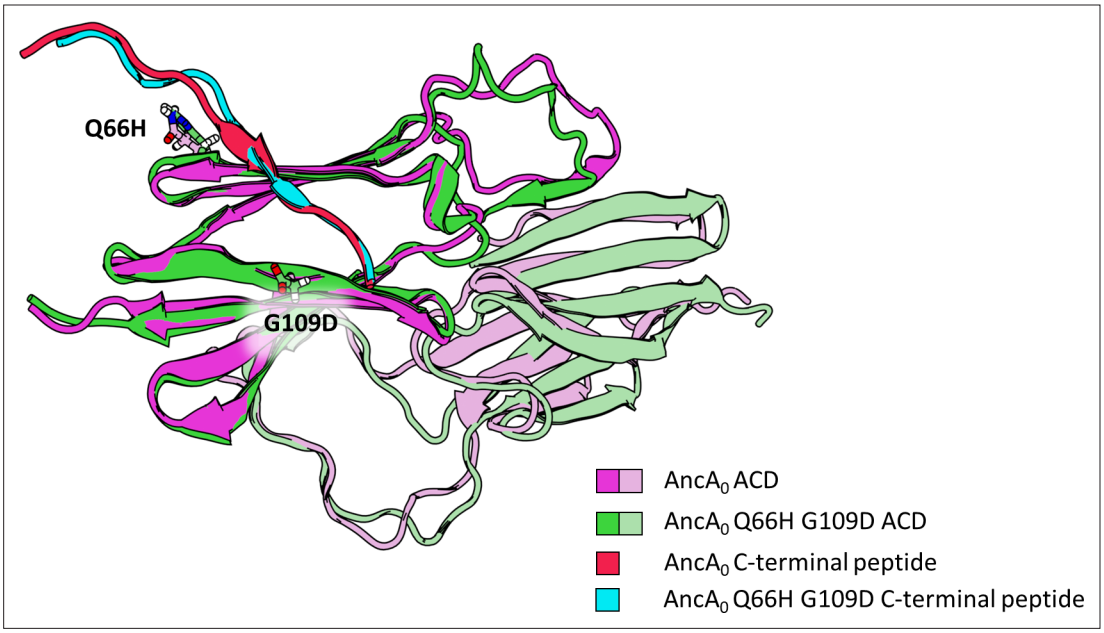

Figure 5—figure supplement 2. Comparison of the structural models of AncA0 (purple) and AncA0 Q66H G109D (green) α–crystallin domain (ACD) dimers complexed with the C-terminal peptides. Models represent middle structures of largest clusters obtained from equilibrium molecular dynamics simulation trajectories based on the AlphaFold2 prediction. Models were superimposed on the backbone heavy atoms of the C-terminal-peptide interacting ACD monomer without dimerization loop (residues 40–74 and 95–126) and the stably interacting C-terminal peptide fragment (residues 132–137), RMSD of the superimposed regions backbone atoms – 0.81 Å, RMSD of the whole complex backbone atoms – 3.42 Å. Substituted residues in positions 66 and 109 are shown in licorice representation.

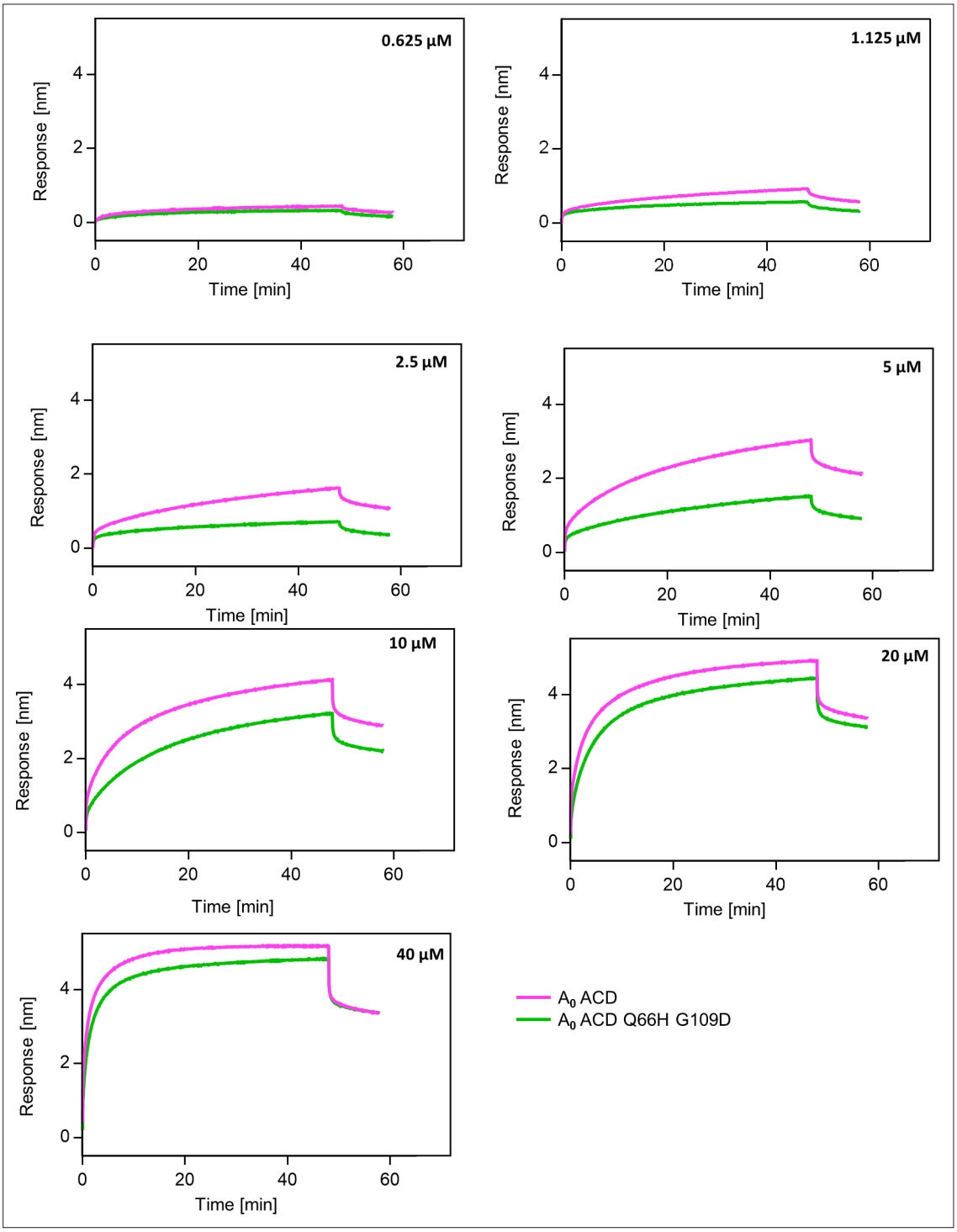

Figure 5—figure supplement 3. Substitutions Q66H G109D decrease affinity of $AncA_0$ ACD to $AncA_0$ C-terminal peptide. Representative binding curves for different concentrations of $AncA_0$ ACD and $AncA_0$ Q66H G109D ACD interacting with $AncA_0$ C-terminal peptide; background was not subtracted. Peptide was immobilized on BLI NTA sensor via His_6-Sumo tag.

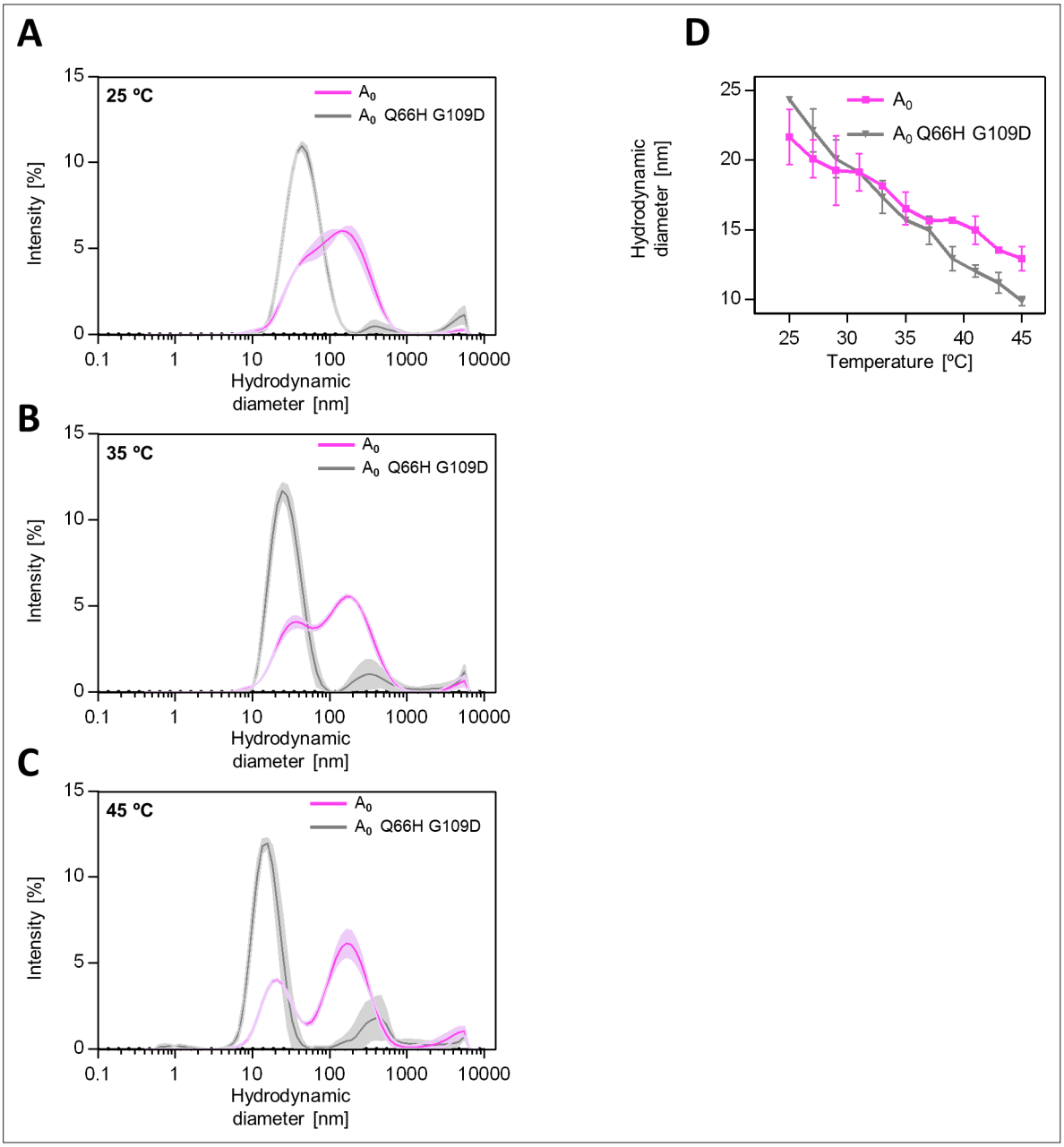

Figure 5—figure supplement 4. Substitutions Q66H G109D influence $AncA_0$ oligomerization. (**A**) Hydrodynamic diameter distributions shown as DLS signal intensity for better visualization purposes of $AncA_0$ and $AncA_0$ Q66H G109D measured at 25, 35, and 45 °C. Size distributions were measured by DLS. Results are shown as an average of 3 repeats. Error bars (shades) represent standard deviation. (**B**) Changes of hydrodynamic diameter representing maximum of the dominant peak of hydrodynamic diameter distribution by volume with temperature. Results are shown as an average of 3 repeats. Error bars represent standard deviation.

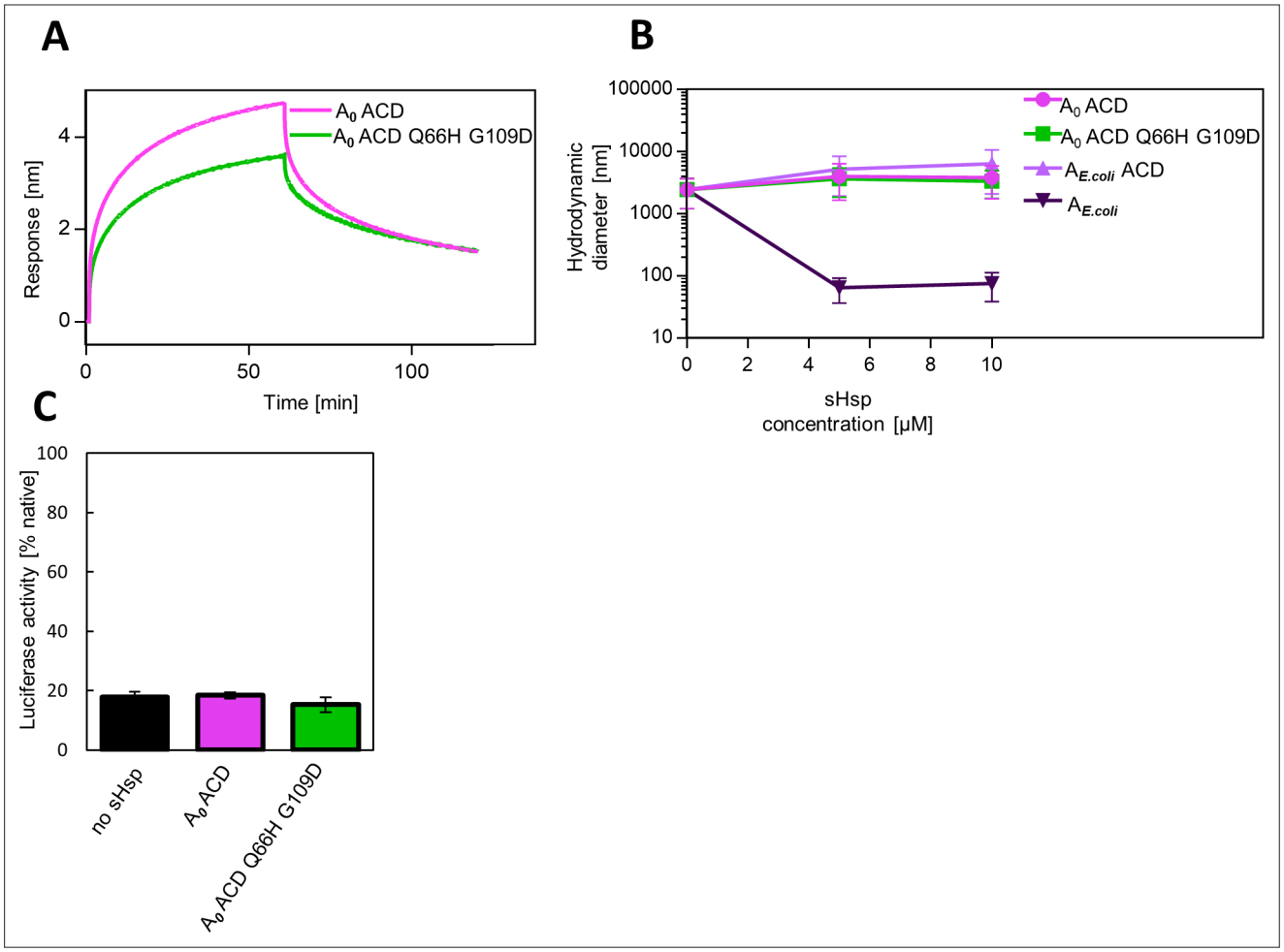

Figure 5—figure supplement 5. $AncA_0$ ACD and $AncA_0$ ACD Q66H G109D can bind to aggregated luciferase, but do not exhibit sequestrase activity or ability to stimulate luciferase refolding. (**A**) Effect of Q66H G109D substitutions on $AncA_0$ ACD affinity to aggregated luciferase; analysis was performed as in Fig.***Figure 3A***. (**B**) Ability of $AncA_0$ ACD and $AncA_0$ Q66H G109D ACD to sequester aggregating luciferase. Results were compared to α–crystallin domain (ACD) of extant *E. coli* IbpA, as well as full-length extant *E. coli* IbpA. Analysis was performed as in ***Figure 3B***. For every measurement, ten subsequent series of ten 10-s runs were averaged and particle size distribution was calculated by fitting to 70 size bins between 0.4 and 10,000 nm. Results are presented as an average hydrodynamic diameter of measured particles weighted by intensity (Z-average), calculated with Malvern Zetasizer Software 7.13. Error bars represent standard deviation obtained from the distribution. (**C**) Ability of $AncA_0$ ACD and $AncA_0$ Q66H G109D ACD to stimulate Hsp70-Hsp100-dependent luciferase disaggregation. Luciferase refolding assay was performed as in ***Figure 1***. Activity of luciferase was measured after 1 hr refolding at 25 °C and shown as an average of 3 repeats. Error bars represent standard deviation.

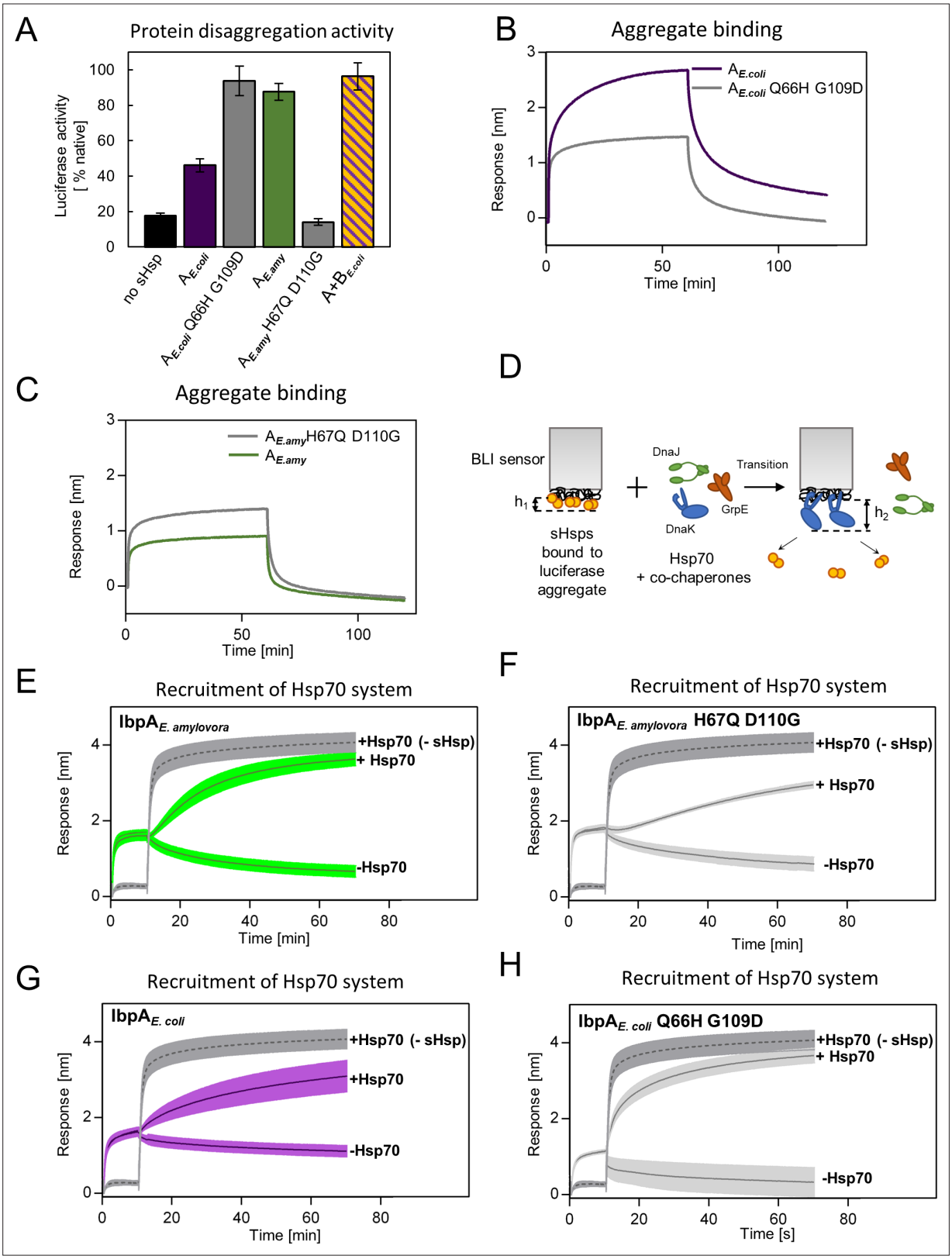

Figure 6. Differences at positions 66 and 109 determine functional differences between extant IbpA proteins from *E. coli* and *E. amylovora*. (**A**) Effect of substitutions at position 66 and 109 (and homologous) on the ability of IbpA from *E. amylovora* and *E. coli* to stimulate luciferase refolding. Assay was performed as in ***Figure 1B***. Activity of luciferase was measured after 1 h refolding at 25 °C and shown as an average of 3 repeats. Error bars represent standard deviation. (**B, C**) Effect of substitutions at analyzed positions on binding of IbpA from *E. coli* (**B**) and *E. amylovora* (**C**) to heat - aggregated *E.*

Figure 6 continued on next page

Figure 6 continued

coli proteins. Assay was performed as in 3 A. (**D–H**) Effect of substitutions at analyzed positions on inhibition of Hsp70 system binding to aggregates by extant small heat shock proteins (sHsps) (**D**) Experimental scheme. (**E–H**) Aggregate-bound sHsps differently inhibit Hsp70 binding. Biolayer interferometry (BLI) sensor with immobilized aggregated luciferase and aggregate bound sHsps was incubated with Hsp70 or buffer (spontaneous dissociation curve).Dark gray traces (dashed line) represent Hsp70 binding to immobilized aggregates in the absence of sHsps. Results are presented as an average of 5 (in the case of IbpA $_{E.\ coli}$ in panel G), 4 (in the case of IbpA $_{E.\ coli}$ Q66H G109d with Hsp70 in panel H), 2 (in the case of IbpA $_{E.\ coli}$ Q66H G109d without Hsp70 in panel H) or 3 repeats (in the case of the remaining curves). Error bars (shades) represent standard deviation.

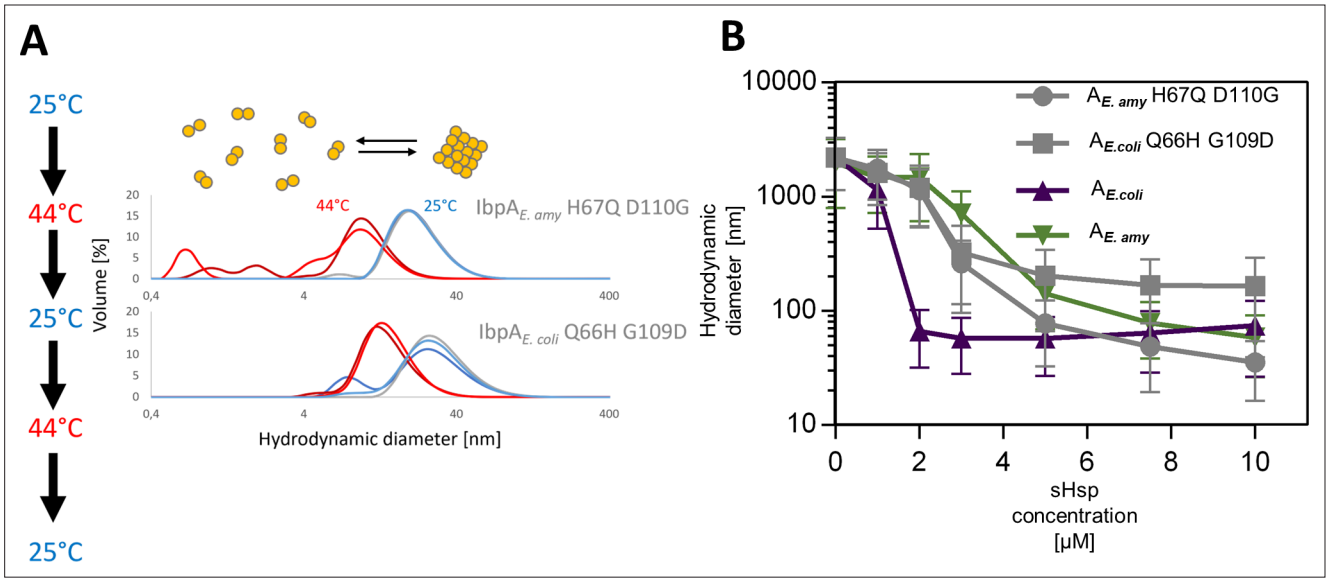

Figure 6—figure supplement 1. IbpA$_{E.coli}$ H66Q G109D and IbpA$_{E.amyl}$ H67Q D110G exhibit sequestrase activity and reversibly deoligomerize at heat shock temperature: (**A**) Reversible deoligomerization of IbpA$_{E. coli}$ H66Q G109D and IbpA$_{E.amyl}$ H67Q D110 at heat shock temperature. Measurements were performed as in ***Figure 3—figure supplement 2***. (**B**) Sequestrase activity of IbpA$_{E. coli}$ H66Q G109D and IbpA$_{E.amyl}$ H67Q D110 at heat shock temperature. Measurements were performed as in ***Figure 3B***. For every measurement, ten subsequent series of ten 10-s runs were averaged and particle size distribution was calculated by fitting to 70 size bins between 0.4 and 10,000 nm. Results are presented as an average hydrodynamic diameter of measured particles weighted by intensity (Z-average), calculated with Malvern Zetasizer Software 7.13. Error bars represent standard deviation obtained from the distribution.