Published in final edited form as:
Proteins. 2012 December ; 80(12): 2758–2768. doi:10.1002/prot.24159.

Flexible connection of the N-terminal domain in ClpB modulates substrate binding and the aggregate reactivation efficiency

Ting Zhang[1], **Elizabeth A. Ploetz**[2], **Maria Nagy**[1,4], **Shannon M. Doyle**[3], **Sue Wickner**[3], **Paul E. Smith**[2], and **Michal Zolkiewski**[1,*]

[1]Department of Biochemistry, Kansas State University, Manhattan, KS 66506, U.S.A

[2]Department of Chemistry, Kansas State University, Manhattan, KS 66506, U.S.A

[3]Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, U.S.A

Abstract

ClpB reactivates aggregated proteins in cooperation with DnaK/J. The ClpB monomer contains two nucleotide-binding domains (D1, D2), a coiled-coil domain, and an N-terminal domain attached to D1 with a 17-residue-long unstructured linker containing a Gly-Gly motif. The ClpB-mediated protein disaggregation is linked to translocation of substrates through the central channel in the hexameric ClpB, but the events preceding the translocation are poorly understood. The N-terminal domains form a ring surrounding the entrance to the channel and contribute to the aggregate binding. It was suggested that the N-terminal domain's mobility that is maintained by the unstructured linker might control the efficiency of aggregate reactivation. We produced seven variants of ClpB with modified sequence of the N-terminal linker. To increase the linker's conformational flexibility, we inserted up to four Gly next to the GG motif. To decrease the linker's flexibility, we deleted the GG motif and converted it into GP and PP. We found that none of the linker modifications inhibited the basal ClpB ATPase activity or its capability to form oligomers. However, the modified linker ClpB variants showed lower reactivation rates for aggregated glucose-6-phosphate dehydrogenase and firefly luciferase and a lower aggregate-binding efficiency than wt ClpB. We conclude that the linker does not merely connect the N-terminal domain, but it supports the chaperone activity of ClpB by contributing to the efficiency of aggregate binding and disaggregation. Moreover, our results suggest that selective pressure on the linker sequence may be crucial for maintaining the optimal efficiency of aggregate reactivation by ClpB.

Keywords

molecular chaperone; AAA+ ATPase; protein aggregation; aggregate reactivation; conformational fluctuations

Introduction

Bacteria, protozoa, fungi, and plants contain aggregate-reactivating systems of molecular chaperones from the Hsp100 and Hsp70 families.[1–3] Hsp100 chaperones (bacterial ClpB, yeast Hsp104) belong to the AAA+ superfamily of ATPases associated with various cellular

*Correspondence to: Michal Zolkiewski, Department of Biochemistry, Kansas State University, 141 Chalmers Hall, Manhattan, KS 66506, U.S.A. Tel: +1-785-532-3083; Fax: +1-785-532-7278; : michalz@ksu.edu.
[4]Present address: Yale University School of Medicine and Howard Hughes Medical Institute, New Haven, CT 06510, U.S.A.
Work performed at: Kansas State University and the National Cancer Institute

activities.[4] Within the AAA+ family, ClpB shows a unique activity of resolubilization of aggregated proteins. ClpB contains two AAA+ ATP-binding sequence modules (D1, D2) with an inserted coiled-coil middle domain and a distinct N-terminal domain connected to D1 with an unstructured linker (see Fig. 1A, B).[5] Like many other AAA+ ATPases, ClpB forms nucleotide-stabilized cylinder-shaped hexamers with a narrow axial channel (see Fig. 1C, D).[6] The channel is surrounded by D1 and D2 domains, while the middle domain is exposed on the outside of the cylinder.[5] In the hexameric ClpB, the N-terminal domains form a "crown" on top of the channel entrance in D1.

The mechanism of protein disaggregation mediated by ClpB involves the ATP-dependent extraction of polypeptides from aggregated particles and their forced separation from the aggregates by translocation through the channel in the hexameric ClpB.[7] Selected substrates can be disaggregated by ClpB alone,[8] but reactivation of most aggregated proteins requires cooperation between ClpB and DnaK/DnaJ/GrpE.[1,2,7] The mechanism of aggregate recognition by ClpB, the events preceding substrate translocation, and the role of the co-chaperones remain poorly understood.

The N-terminal domain of ClpB provides contacts for binding of large aggregates[9] and supports heat-shock survival of the DnaK-deficient *E. coli*.[10] The location of the N-terminal domain on the top surface of the hexameric cylinder[11] suggests that its interactions with aggregated substrates may precede their insertion into the translocation channel. Interestingly, the N-terminal domain does not maintain direct physical contacts with the rest of ClpB.[12] Thus, the position of the N-terminal domain is solely maintained by the unstructured 17-residue linker that attaches that domain to D1.[5] Among the three *T. thermophilus* ClpB molecules that formed the crystallographic asymmetric unit, each showed a different orientation of the N-terminal domain with a rotation by as much as 120°.[5] The high mobility of the N-terminal domain within the ClpB oligomer has been also demonstrated by the apparent lack of visible structures on top of the D1 ring in the averaged cryo-EM images of ClpB.[5]

Due to a dual translation-initiation site, *E. coli* produces two isoforms of ClpB: the 95-kDa full-length ClpB and the 80-kDa truncated form which does not contain the N-terminal domain.[13,14] We have previously shown that the full aggregate-reactivating potential of ClpB is achieved through a synergistic cooperation of the two isoforms and formation of hetero-oligomers.[15] Normal-mode structural analysis of the ClpB95/80 hetero-oligomers suggested that the mobility of the N-terminal domain is enhanced in the hetero-oligomers due to the insertion of subunits lacking that domain.[15] A recent study has shown that a disulfide-mediated immobilization of the N-terminal domain decreased the aggregate reactivation efficiency.[16] Altogether, the structural and biochemical characterization of ClpB suggested that the mobility of the N-terminal domain supports the aggregate-reactivation activity. In this work, we explored the hypothesis that while the unstructured linker does not apparently participate in aggregate binding, it may support the ClpB function by controlling the position and motions of the aggregate-binding N-terminal domain.

The sequence of the ClpB linker is poorly conserved (see Fig. 2A), but it contains a high number of charged residues, which maintain an unstructured conformation and multiple glycines, which could provide conformational flexibility. We hypothesized that amino acid substitutions within the linker might modulate the N-terminal domain mobility and, as a consequence, the aggregate-reactivation activity of ClpB. In this study, we compared the biochemical properties and activity of wt *E. coli* ClpB and seven variants with different linker sequences (see Fig. 2B). To increase the number of the conformational degrees of freedom within the linker, we inserted up to four Gly in the vicinity of Gly149/150. To decrease the number of the degrees of freedom, we deleted Gly149 and Gly150 and replaced

one or both Gly with Pro. As described below, we observed a modulation of the ClpB functionality by the linker sequence modifications. Our results show that the linker does not merely maintain a physical connection between the ClpB domains, but it also plays a role in the mechanism of protein disaggregation.

Materials and Methods

Proteins and Aggregates

Protein concentration was determined spectrophotometrically and reported in monomer units. Chaperones (ClpB, DnaK, DnaJ, GrpE) were produced or obtained as previously described.[9] G6PDH from *Leuconostoc mesenteroides* was produced in *E. coli* BL21(DE3) harboring the plasmid pET25-G6PDH (gift from Dr. Michael S. Cosgrove, Syracuse University). Briefly, the cells were grown at 37 °C with shaking to OD(600nm)~0.5 and then induced with 0.4 mM IPTG. 4 hours post-induction, cells were collected by centrifugation and disrupted by sonication. 0.004 g/g cell PEI was added to the soluble fraction to precipitate DNA. The soluble fraction was treated with 50% saturated $(NH_4)_2SO_4$ and centrifuged. The supernatant was dialyzed to remove excess salt and then loaded on an anion-exchange column (Q-Sepharose, fast flow). Fractions containing G6PDH were combined and concentrated using Amicon ultra centrifuge filter. Concentrated G6PDH was dialyzed to remove salt and then stored at −20 °C. Firefly luciferase was obtained from Promega (Madison, WI) and α-casein from Sigma (St. Louis, MO). GFP was produced as described previously.[8] Mutations in the ClpB linker DNA sequence were introduced with QuikChange site-directed mutagenesis kit (Stratagene/Agilent Technologies).

To produce aggregates of G6PDH, the protein stock (324 μM) was diluted 2-fold with the unfolding buffer (10 M urea, 16% glycerol and 40 mM DTT) and was incubated at 47 °C for 5 min. Then the mixture was diluted 10-fold with the refolding buffer 1 (50 mM Tris/HCl pH 7.5, 20 mM $Mg(OAc)_2$, 30 mM KCl, 1 mM EDTA, and 1mM β-mercaptoethanol) and was incubated at 47 °C for 15 min and then on ice for 2 min. To produce aggregates of luciferase, 216 μM luciferase stock was diluted 300-fold with PBS containing 1 mg/ml BSA and then was incubated at 45 °C for 12 min. To produce aggregated GFP, 4.5 μM protein was heated for 10 min at 80 °C.

Differential Scanning Calorimetry

DSC experiments were performed with a VP-DSC calorimeter (MicroCal Inc., Northampton, Massachusetts). Thermograms were obtained at the 1K/min scan rate for the protein samples at 0.7 mg/ml. For each protein sample, the instrument baseline was obtained first by measuring the thermogram of the dialysis buffer (50 mM Hepes/KOH, pH. 7.5). Subsequently, the buffer in the sample cell was replaced with a ClpB solution and the protein thermogram was measured. Baselines measured with the dialysis buffer were subtracted from the protein scans. The thermal unfolding of ClpB was irreversible, as shown by the lack of endotherms in the repetitive scans of each protein sample.

Analytical Ultracentrifugation

Beckman XL-I analytical ultracentrifuge was used in sedimentation velocity experiments with two-channel analytical cells. Ultracentrifugation was performed at 35,000 rpm and 20 °C for the 2-mg/ml protein samples in 50 mM Hepes/KOH pH 7.5, 0.2 M KCl, 20 mM $MgCl_2$, 1 mM EDTA, 2 mM β-mercaptoethanol. The data were analyzed using the time-derivative method[17,18] and the software distributed with the instrument.

ClpB ATPase activity

The ClpB variants were incubated in assay buffer (100mM Tris/HCl pH 8.0, 1mM DTT, 1mM EDTA, 10mM $MgCl_2$, and 5mM ATP) at 37 °C for 15min without or with 0.1 mg/ml α-casein or 0.04 mg/ml poly-lysine. The concentration of ClpB was 0.05 mg/ml for the basal activity and in the presence of α-casein or 0.005 mg/ml in the presence of poly-lysine. The phosphate concentration generated by ClpB was measured as described before.[2]

Aggregate Reactivation Assays

Aggregated G6PDH (16 μM) was diluted 10-fold with refolding buffer 1 containing 1.5 μM ClpB, 1 μM DnaK, 1 μM DnaJ, 0.5 μM GrpE and 6 mM ATP. The mixture was incubated at 30 °C and aliquots of the mixture were withdrawn to test the recovery of the G6PDH enzymatic activity. Aggregates diluted with refolding buffer without the chaperones were used as control. To measure the G6PDH activity, aliquots from the refolding reaction were incubated in 50 mM Tris/HCl pH 7.8, 5 mM $MgCl_2$, 1.5 mM G6P and 1 mM $NADP^+$ for 10 min followed by the measurement of absorption at 340 nm. Aggregated luciferase (1 μM) was diluted 20-fold with refolding buffer 2 (30 mM Hepes, pH 7.65, 120 mM KCl, 10 mM $MgCl_2$, 6 mM ATP, 1 mM EDTA, 10 mM DTT, 0.1 mg/ml BSA) containing 1.5 μM ClpB, 1 μM DnaK, 1 μM DnaJ, and 0.5 μM GrpE. The mixture was incubated at room temperature and aliquots were withdrawn to test the recovery of the luciferase activity using the luminescence assay system (Promega, Madison, WI). GFP reactivation assays in the absence of the co-chaperones were performed in 20 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM DTT, 0.1 mM EDTA, and 10% glycerol (vol/vol) with 2 mM ATP and 2 mM ATPγS, an ATP regenerating system (20 mM creatine phosphate and 6 μg creatine kinase), 10 mM $MgCl_2$, 10 μL heat-aggregated GFP (heated 10 min at 80°C at 4.5 μM) and 1.0 μM ClpB. GFP reactivation with the co-chaperones included 1.4 μM DnaK, 0.5 μM DnaJ, 0.3 μM GrpE and 4 mM ATP. GFP reactivation was initiated by the addition of Mg-ATP and reactivation was monitored over time at 23 °C using a Perkin-Elmer LS50B luminescence spectrophotometer with a plate reader. Excitation and emission wavelengths were 395 nm and 510 nm, respectively.

ClpB-Aggregate Interaction Assay

Aggregated G6PDH (16 μM) was diluted 10-fold with the refolding buffer 1 containing 1.5 μM ClpB and 6 mM ATPγS. The mixture was incubated at 30 °C with 600 rpm shaking for 20 min and then was applied to the filter device (Millipore Ultrafree-MC Centrifugal Filter Unit with the membrane, pore size 0.1 μm). After 5 min incubation at room temperature, the filter device was centrifuged at 13,000 rpm for 4 min to get the flow-through fractions. The filter device was washed with the refolding buffer 1 containing ATPγS at 30 °C for 5 min and then re-centrifuged. Next, 1x SDS loading buffer was added to the filter device and the filter device was incubated at 50 °C for 5 min with shaking. Then, it was centrifuged to get the eluate fractions, which were applied to SDS-PAGE. Aggregated luciferase (7 μM) was diluted 20-fold with refolding buffer 2 (30 mM Hepes, pH 7.65, 120 mM KCl, 10 mM $MgCl_2$, 5 mM ATPγS, 1 mM EDTA, 10 mM DTT) containing 1.5 μM ClpB and processed as described above. The Coomassie-stained band intensity was determined with the BandScan software (http://bandscan.software.informer.com).

Molecular Dynamics Simulations

The peptides corresponding to the ClpB linker sections (see Fig. 2B) were simulated using classical MD techniques and the GROMACS simulation package (version 4.0)[19] in the isothermal-isobaric (*NpT*) ensemble at a p of 1.0 atm and a T of 303.15 K in 0.15 M NaCl using the v-rescale thermostat[20] and the Berendsen barostat[21], with relaxation times of 0.1 ps and 5.0 ps, respectively, with a 4.5×10^{-5} bar^{-1} compressibility. The protein and ions

were modeled using the Kirkwood-Buff Force Field (KBFF.chem.k-state.edu)[22,23] in explicit solvent with the SPC/E water model.[24] Truncated octahedron boxes defined the systems. Periodic boundary conditions and the minimum image convention were employed. System volumes ranged from 48.7 nm^3 to 126.6 nm^3, with the box sizes chosen so that the image distance was the length of the fully extended linker plus three nanometers. All bond lengths were constrained using the Settle[25] and Linear Constraint Solver (LINCS)[26] algorithms for water and non-water molecules, respectively. The use of bond constraints allowed for a two fs time step to be used for the integration of the equations of motion, which was performed using the Leap-Frog algorithm.[27] The particle-mesh Ewald technique was used to calculate electrostatic interactions with cutoff distances of 1.0 nm and 1.5 nm for the real space electrostatic and van der Waals interactions, respectively, a convergence parameter of 3.123 nm^{-1}, cubic interpolation, a maximum fast Fourier transform grid spacing of 0.12 nm for the reciprocal space sum, and tinfoil boundary conditions.[28] The neighbor list was updated every ten steps. The initial linker structures were built using Swiss PDB Viewer.[29] Since these simulations modeled the behavior of the full-length linker, the N-terminus was blocked with an acetyl group (ACE) and the C-terminus with an *N*-methyl group (NHM). The initial conformation of all linker sequences was fully extended (backbone dihedral angles $\varphi = \psi = 180°$). The steepest descent method was used to perform ≥1000 steps of energy minimization. This was followed by 100 ps of equilibration and one microsecond of production for each linker sequence. Configurations were saved every 1.0 ps for analysis. The ACE to NHM center-of-mass distance was measured for each peptide to quantify the effective end-to-end distance. The errors of the median end-to-end distances, which were calculated as the standard deviation of four 250-ns subaverages, ranged from 0.02 nm (smallest error, occurring in the 2P mutant simulation) to 0.08 nm (largest error, occurring in the DelG mutant simulation). The formation of Arg-Glu salt bridge was scored if the distance between the Arg C^{ε} and the Glu C^{δ} were less than 0.5 nm.

Results

Molecular Dynamics of the ClpB Linker Fragment

To obtain insight into the possible effects of the engineered mutations on the conformational ensemble of the ClpB linker, we performed molecular dynamics (MD) simulations. We simulated the MRGGES peptide and its seven variants, which correspond to the modified linker segment (see Fig. 2B). Since linker modifications did not include residues that could support long-range interactions, we assumed that the internal dynamics of the segment as well as its end-to-end distributions would not be affected by the attachment to the rest of the protein or that all variants would be affected in a similar way. In particular, conformational fluctuations of the simulated linker fragments may become attenuated in the full-length ClpB due to steric conflicts between the N-terminal domains within the hexamer.

Fig. 3 shows snapshots at every 0.05 μs of the 1-μs long simulations of each peptide sequence. The simulation shows that the wt linker segment fluctuates around a quasi-helical structure that is preserved to a high extent after insertion of an additional glycine (1G variant). This structure appears to be stabilized by a salt bridge between Arg and Glu (see Methods) which breaks and reforms repeatedly, but is present for ~40% of the 1-μs simulation of the wt peptide. The conformational fluctuations of the linker peptide become more pronounced in 2G, 3G, and 4G variants with significant twisting of the peptide backbone and strong deviations from the quasi-helical structure (see Fig. 3). In contrast, upon deletion of the two glycines (DelG) or their replacement with Pro (1P, 2P), the conformational fluctuations are suppressed, which is documented by most parts of the peptide backbone retaining its conformation over time. However, the overall structure of 1P and 2P is different from the wt linker because the Arg-Glu salt bridge is present for only 27% of the simulation time in 1P and absent during the simulation of 2P.

To investigate possible effects of the linker sequence on the mutual orientation of the domains connected by the linker, we calculated distributions of the dihedral angle between the N-terminal and C-terminal dipole moment vectors in the simulated peptides. As shown in Fig. S1, the wt peptide shows two preferred configurations with the dihedral angle of ~30° and ~−120°. As expected, DelG shows a narrow probability distribution, which demonstrates its limited capability to support a domain rotation, but the DelG-preferred dihedral angle of ~170° differs significantly from those of the wt sequence. In contrast, all the remaining modified peptides show broad angular distributions with multiple probable conformations. Importantly, none of the modified linkers approximates the angular preference of the wt sequence.

The simulations were also used to calculate the end-to-end distance of the linker segment and the magnitude of fluctuations in the effective peptide length. As shown in Fig. S2, the median end-to-end distance of the wt linker segment increased by ~30% for 1G, 2G, 3G, 4G, and 2P variants and decreased by ~33% in the DelG variant, but the substantial distance fluctuations suggest that the effective “reach” of the N-terminal domain might not significantly increase or decrease in the modified linker variants on ClpB.

Biochemical Properties of the Linker-Modified ClpB Variants

After purifying the linker-modified variants of ClpB, we tested their structural integrity. Fig. S3 shows the thermal denaturation profiles of wt ClpB and its variants obtained with differential scanning calorimetry (DSC). As has been shown before[30], the thermal denaturation of ClpB is irreversible, as expected for a multi-domain protein, and the heat capacity profiles show multiple transitions spanning a range of temperatures between ~50 and 70 °C. Importantly, the high-temperature transition at 64 °C corresponds to the thermal unfolding of the N-terminal domain.[30] As shown in Fig. S3, the DSC profiles for the ClpB linker variants closely parallel that of wt ClpB, which indicates that the changes in sequence of the linker do not affect the thermodynamic stability of the other ClpB domains and that all the ClpB variants are stable and folded at the temperature of the assays (20–37 °C). This result is expected as the linker does not fold into a stable structure and does not structurally participate in folding of the domains it connects.

Self-association into cylinder-shaped oligomers is required for the activity of ClpB.[31] We tested possible effects of the linker modifications on the oligomerization of ClpB by performing sedimentation velocity experiments of wt ClpB and the three most radically modified variants: 4G, DelG, and 2P. The sedimentation velocity was measured at the protein concentration for which the wt oligomers are stable even in the absence of nucleotides.[6,32] As shown by the apparent distributions of the sedimentation coefficients g(s*) in Fig. S4, the solutions of wt ClpB and the modified linker variants contain ~16S oligomers, which indicates that the changes in the linker sequence do not alter the ClpB’s ability to self-associate.

Since the ATP-binding sites in ClpB are located at the subunit interfaces, the oligomerization is linked to nucleotide binding and the ATPase activity.[31] As shown in Fig. 4, the basal ATPase of the linker variants is similar to that of wt ClpB, which agrees with the previous conclusion that all ClpB linker variants are capable of oligomerization. Interestingly, the ClpB ATPase activation by casein is weaker by ~40% in all the modified linker variants (Fig. 4). Another known ATPase activator, poly-lysine, activates all ClpB variants with no obvious trend in comparison to wt ClpB. Altogether, the data in Fig. 4 show that all modifications of the linker sequence partially inhibit the capability of ClpB to respond to a pseudo-substrate casein.

Chaperone Activity of the Linker-Modified ClpB Variants

We next asked if the ClpB linker modification might affect the aggregate reactivation activity. We tested two aggregated substrates: glucose-6-phosphate dehydrogenase (G6PDH) and firefly luciferase, whose reactivation requires ClpB as well as the co-chaperones DnaK/DnaJ/GrpE (KJE)[2,9], and green fluorescent protein (GFP), which can be also reactivated by ClpB without KJE.[8] As has been shown before,[9,15] the ClpB variant lacking the N-terminal domain (DelN, see Fig. 5) did not reactivate large aggregates of G6PDH as efficiently as the full-length ClpB, which indicates that the N-terminal domain participates in the aggregate reactivation. Interestingly, the ClpB variants with modified linker sequence showed a significant decrease in the reactivation rates of the G6PDH aggregates (see Fig. 5). Indeed, the linker variants: 3G, 4G, DelP, 1P, and 2P reactivated aggregated G6PDH as inefficiently as ClpB DelN, which suggests that those linker sequence modifications effectively inactivated the N-terminal domain during the G6PDH reactivation.

The N-terminal domain of ClpB supports also the reactivation of thermally denatured aggregated luciferase.[31] Indeed, ClpB DelN reactivated aggregated luciferase less efficiently than the full-length ClpB (see Fig. 6). Among the linker variants, 3G and 4G showed the lowest luciferase reactivation rates while 1G, 2G, DelG, 1P, and 2P were less affected.

We also tested the reactivation of small aggregates of green fluorescent protein (GFP) by the ClpB linker variants in the presence of KJE and under conditions that do not require the co-chaperones: in the presence of ATP with ATPγS.[8] As shown in Fig. S5, the linker modifications did not inhibit the ClpB activity during the GFP reactivation either with or without KJE. Interestingly, the reactivation of aggregated GFP was more efficient with ClpB DelN than with the full-length ClpB. Altogether, the linker modifications appear to partially inhibit the reactivation rate for those aggregates, which require the N-terminal domain of ClpB for optimal disaggregation (Figs. 5, 6).

Next, we tested whether the reduced aggregate reactivation rates observed with the modified linker ClpB variants might be linked to a less efficient binding of the chaperone to the aggregates. We incubated ClpB with the native or aggregated G6PDH or luciferase in the presence of ATPγS, which induces a stable binding of the chaperone to the substrate.[33] The aggregates were then separated from soluble proteins using filtration and analyzed with SDS-PAGE. As shown in Figs. 7A and 8A, background amounts of ClpB were retained on the filter in the absence of the aggregates (Fig. 7A, left half of the gel; Fig. 8A, first lane). The amount of ClpB increased substantially in the presence of aggregates, which demonstrates interactions of the chaperone with the substrate. In repeated experiments, we consistently observed that the amounts of the ClpB linker variants bound to the G6PDH aggregates were lower than those of wt ClpB. Band density analysis (Fig. 7B) showed the strongest inhibition of substrate binding for 4G and 2P variants (~50% decrease compared to wt ClpB). However, the efficiency of binding of those two most affected variants to the aggregated G6PDH was higher than that of ClpB DelN (Fig. 7B and Fig. S6).

The linker modifications in ClpB affected to a lower degree the efficiency of binding to the aggregated luciferase than to the aggregated G6PDH with the strongest inhibition found for DelG and DelN (Fig. 8B). The GFP aggregates (see Fig. S5) were not retained on the filter used in the binding experiments (data not shown), which indicates their smaller average size, as compared to the aggregates of G6PDH and luciferase.

Discussion

In this work, we discovered that the sequence of the unstructured linker, which connects the N-terminal domain of ClpB with its D1 AAA+ module, supports the disaggregase activity of

ClpB. This result demonstrates that the linker, which does not even fold into a stable structure, participates nevertheless in the mechanism of the ATP-driven aggregate reactivation. Previous studies suggested that the position and orientation of the substrate-interacting N-terminal domain of ClpB may undergo significant fluctuations.[5,15] Since the N-terminal domain of ClpB does not form stable physical contacts with D1[12], its apparent mobility in hexameric ClpB must arise from the conformational fluctuations of the linker.

The MD simulations (Figs. 3, S1, and S2) suggest that the linker sequence modifications may affect the conformational space explored by the N-terminal domain of ClpB by modulating the domain's rotation. It should be noted that the simulation predicted two preferred orientations of the wt linker segment (see Fig. S1), which is consistent with the two distinct orientations of the N-terminal domain found in the ClpB crystal structure.[5]

Although it has been suggested that the high mobility of the N-terminal domain supports the ClpB activity[15,16], we found that the highly mobile glycine insertions (see Fig. 3) did not enhance the functionality of ClpB. In contrast, its functionality was partially inhibited by all changes in the linker sequence (see Figs. 4, 5, 6), regardless of their potential implications for the conformational properties of the linker, as suggested by the MD simulations. Thus, our results imply that selective pressure on the ClpB linker sequence may be crucial for maintaining the optimal efficiency of aggregate reactivation by ClpB. Many AAA+ ATPases contain substrate-binding domains attached to the AAA+ modules with seemingly non-conserved linker segments.[4,34] Our results for ClpB suggest that the conformational space of the substrate-binding domains might be tightly controlled by the unstructured linkers in other members of the AAA+ ATPase family. Indeed, either elongation or shortening of the linker in ClpA (the ATPase component of the protease ClpAP) causes a reduction of its activity.[35]

Predictably, the linker sequence modifications affect the reactivation of large aggregates of G6PDH and luciferase, which depend on the N-terminal domain of ClpB (Figs. 5, 6), but not small aggregates of GFP, which do not require the N-terminal domain for efficient reactivation (Fig. S5). The higher reactivation rate of aggregated GFP with ClpB DelN as compared to wt ClpB might be linked to the higher ATPase activity of the former[15,31] or to a possible obstruction of access of GFP aggregates to the wt ClpB channel. The effects of linker sequence modifications appear to depend on the substrate: the partial reactivation inhibition is stronger for G6PDH than for luciferase (Figs. 5, 6). It is striking that the flexibility-enhancing modifications (3G, 4G) as well as the flexibility-suppressing modifications (DelG, 1P, 2P) show a similar degree of inhibition during the G6PDH reactivation (Fig. 5).

How does the linker support the chaperone activity of ClpB? The defects in the aggregate reactivation rate shown by ClpB with the modified linker can be explained in part by the defects in substrate binding (Figs. 7, 8). One possible interpretation of this result is that the aggregates bind directly to the linker and the changes in the linker sequence interfere with aggregate binding by removing some essential physical contacts. We do not favor this explanation for the following reasons. First, the linker occupies a space between two bulky domains: the N-terminal domain and D1 (see Fig. 1B)[5] and is not readily accessible for interactions with a particle larger than the ClpB hexamer, such as aggregated G6PDH.[9] Second, the linker modifications do not involve negatively charged residues, which were shown to support the aggregate binding by ClpB.[9,36] A more likely explanation of the results in Figs. 7 and 8 is that the modification of the linker may have affected the position and/or orientation of the substrate-binding N-terminal domain of ClpB. This notion is also supported by the defects of all modified ClpB variants in their response to casein (see Fig. 4). Unlike poly-lysine, which does not require the N-terminal domain of ClpB to activate the

ATPase, casein activates the ClpB ATPase through interactions with the N-terminal domain.[30]

The aggregate binding defects of the linker variants shown in Figs. 7 and 8 cannot fully account for the observed decreases in the aggregate reactivation rates (Figs. 5, 6). The ATPγS-induced ClpB binding to aggregates may correspond to the initial step of disaggregation before a substrate becomes inserted into the ClpB channel.[37] The full substrate engagement by the ATP-dependent motions of the channel loops requires assistance from the KJE co-chaperones[38,39], which may stimulate ClpB through interactions with its middle domain.[40] Since the ring of the N-terminal domains surrounds the channel entrance (see Fig. 1), the modifications of the linker sequence may transform the conformational space accessible to the N-terminal domain in such a way that the efficiency of substrate insertion into the channel and its translocation become sub-optimal. Future studies may identify couplings between the linker-supported motions of the N-terminal domain and the efficiency of substrate insertion into the ClpB channel. In particular, crosslinking of the substrates with the channel loops in the linker-modified ClpB variants should be investigated.

Multi-domain proteins often contain unstructured regions that connect stably folded structural units. The inter-domain linkers are often considered irrelevant to protein function beyond maintaining a continuity of the multi-domain polypeptides. Our results demonstrate that the linker in ClpB supports the chaperone function. This conclusion is consistent with the emerging model of the inter-domain allostery in multi-domain proteins.[41] The sequence of an inter-domain linker may play a role in protein function because it may encode an ensemble of conformations that allow propagation of allosteric signals between different domains. In the case of ClpB, the sequence of the linker may support conformational fluctuations that help ClpB in finding the binding sites at the surface of an aggregate and transfer the substrate from the N-terminal domain to the D1/D2 channel.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health (R01GM079277, ARRA Supplement to PES and MZ), Terry C. Johnson Center for Basic Cancer Research, the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and the Kansas Agricultural Experiment Station (contribution 12-234-J). EAP was supported by the KSU NSF GK-12 program under grant NSF DGE-0841414 and the NSF GRF program under grant NSF DGE-0750823.

References

1. Glover JR, Lindquist S. Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. Cell. 1998; 94:73–82. [PubMed: 9674429]
2. Zolkiewski M. ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from Escherichia coli. J Biol Chem. 1999; 274:28083–28086. [PubMed: 10497158]
3. Goloubinoff P, Mogk A, Zvi AP, Tomoyasu T, Bukau B. Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. Proc Natl Acad Sci U S A. 1999; 96:13732–13737. [PubMed: 10570141]
4. Neuwald AF, Aravind L, Spouge JL, Koonin EV. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res. 1999; 9:27–43. [PubMed: 9927482]

5. Lee S, Sowa ME, Watanabe YH, Sigler PB, Chiu W, Yoshida M, Tsai FT. The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell. 2003; 115:229–240. [PubMed: 14567920]
6. Akoev V, Gogol EP, Barnett ME, Zolkiewski M. Nucleotide-induced switch in oligomerization of the AAA+ ATPase ClpB. Protein Sci. 2004; 13:567–574. [PubMed: 14978298]
7. Weibezahn J, Tessarz P, Schlieker C, Zahn R, Maglica Z, Lee S, Zentgraf H, Weber-Ban EU, Dougan DA, Tsai FT, Mogk A, Bukau B. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. Cell. 2004; 119:653–665. [PubMed: 15550247]
8. Doyle SM, Shorter J, Zolkiewski M, Hoskins JR, Lindquist S, Wickner S. Asymmetric deceleration of ClpB or Hsp104 ATPase activity unleashes protein-remodeling activity. Nat Struct Mol Biol. 2007; 14:114–122. [PubMed: 17259993]
9. Barnett ME, Nagy M, Kedzierska S, Zolkiewski M. The amino-terminal domain of ClpB supports binding to strongly aggregated proteins. J Biol Chem. 2005; 280:34940–34945. [PubMed: 16076845]
10. Chow IT, Barnett ME, Zolkiewski M, Baneyx F. The N-terminal domain of Escherichia coli ClpB enhances chaperone function. FEBS Lett. 2005; 579:4242–4248. [PubMed: 16051221]
11. Lee S, Sielaff B, Lee J, Tsai FT. CryoEM structure of Hsp104 and its mechanistic implication for protein disaggregation. Proc Natl Acad Sci U S A. 2010; 107:8135–8140. [PubMed: 20404203]
12. Tek V, Zolkiewski M. Stability and interactions of the amino-terminal domain of ClpB from Escherichia coli. Protein Sci. 2002; 11:1192–1198. [PubMed: 11967375]
13. Squires CL, Pedersen S, Ross BM, Squires C. ClpB is the Escherichia coli heat shock protein F84. 1. J Bacteriol. 1991; 173:4254–4262. [PubMed: 2066329]
14. Park SK, Kim KI, Woo KM, Seol JH, Tanaka K, Ichihara A, Ha DB, Chung CH. Site-directed mutagenesis of the dual translational initiation sites of the clpB gene of Escherichia coli and characterization of its gene products. J Biol Chem. 1993; 268:20170–20174. [PubMed: 8376377]
15. Nagy M, Guenther I, Akoyev V, Barnett ME, Zavodszky MI, Kedzierska-Mieszkowska S, Zolkiewski M. Synergistic cooperation between two ClpB isoforms in aggregate reactivation. J Mol Biol. 2010; 396:697–707. [PubMed: 19961856]
16. Mizuno S, Nakazaki Y, Yoshida M, Watanabe YH. Orientation of the amino-terminal domain of ClpB affects the disaggregation of the protein. FEBS J. 2012; 279:1474–1484. [PubMed: 22348341]
17. Stafford, WF., III, editor. Sedimentation boundary analysis of interacting systems: Use of the apparent sedimentation coefficient distribution function. Boston: Birkhauser; 1994. p. 119-137.
18. Stafford WF 3rd. Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. Anal Biochem. 1992; 203:295–301. [PubMed: 1416025]
19. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theor Comp. 2008; 4:435–447.
20. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys. 2007:126.
21. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR. Molecular-Dynamics with Coupling to an External Bath. J Chem Phys. 1984; 81:3684–3690.
22. Weerasinghe, S.; Gee, MB.; Kang, M.; Bentenitis, N.; Smith, PE. Developing Force Fields from the Microscopic Structure of Solutions: The Kirkwood–Buff Approach. In: Feig, M., editor. Modeling Solvent Environments: Applications to Simulations of Biomolecules. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA; 2010.
23. Ploetz EA, Bentenitis N, Smith PE. Developing force fields from the microscopic structure of solutions. Fluid Phase Eq. 2010; 290:43–47.
24. Berendsen HJC, Grigera JR, Straatsma TP. The Missing Term in Effective Pair Potentials. J Phys Chem. 1987; 91:6269–6271.
25. Miyamoto S, Kollman PA. Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. J Comput Chem. 1992; 13:952–962.

26. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. J Comput Chem. 1997; 18:1463–1472.
27. Hockney RW, Goel SP, Eastwood JW. Quiet High-Resolution Computer Models of a Plasma. J Comput Phys. 1974; 14:148–158.
28. Darden T, York D, Pedersen L. Particle Mesh Ewald - an N. Log(N) Method for Ewald Sums in Large Systems. J Chem Phys. 1993; 98:10089–10092.
29. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis. 1997; 18:2714–2723. [PubMed: 9504803]
30. Liu Z, Tek V, Akoev V, Zolkiewski M. Conserved amino acid residues within the amino-terminal domain of ClpB are essential for the chaperone activity. J Mol Biol. 2002; 321:111–120. [PubMed: 12139937]
31. Barnett ME, Zolkiewska A, Zolkiewski M. Structure and activity of ClpB from Escherichia coli., Role of the amino-and -carboxyl-terminal domains. J Biol Chem. 2000; 275:37565–37571. [PubMed: 10982797]
32. Zolkiewski M, Kessel M, Ginsburg A, Maurizi MR. Nucleotide-dependent oligomerization of ClpB from Escherichia coli. Protein Sci. 1999; 8:1899–1903. [PubMed: 10493591]
33. Nagy M, Wu HC, Liu Z, Kedzierska-Mieszkowska S, Zolkiewski M. Walker-A threonine couples nucleotide occupancy with the chaperone activity of the AAA+ ATPase ClpB. Protein Sci. 2009; 18:287–293. [PubMed: 19177562]
34. Hanson PI, Whiteheart SW. AAA+ proteins: have engine, will work. Nat Rev Mol Cell Biol. 2005; 6:519–529. [PubMed: 16072036]
35. Cranz-Mileva S, Imkamp F, Kolygo K, Maglica Z, Kress W, Weber-Ban E. The flexible attachment of the N-domains to the ClpA ring body allows their use on demand. J Mol Biol. 2008; 378:412–424. [PubMed: 18358489]
36. Schlieker C, Weibezahn J, Patzelt H, Tessarz P, Strub C, Zeth K, Erbse A, Schneider-Mergener J, Chin JW, Schultz PG, Bukau B, Mogk A. Substrate recognition by the AAA+ chaperone ClpB. Nat Struct Mol Biol. 2004; 11:607–615. [PubMed: 15208691]
37. Zolkiewski M, Zhang T, Nagy M. Aggregate reactivation mediated by the Hsp100 chaperones. Arch Biochem Biophys. 2012; 520:1–6. [PubMed: 22306514]
38. Haslberger T, Weibezahn J, Zahn R, Lee S, Tsai FT, Bukau B, Mogk A. M domains couple the ClpB threading motor with the DnaK chaperone activity. Mol Cell. 2007; 25:247–260. [PubMed: 17244532]
39. Acebron SP, Martin I, del Castillo U, Moro F, Muga A. DnaK-mediated association of ClpB to protein aggregates. A bichaperone network at the aggregate surface. FEBS Lett. 2009; 583:2991–2996. [PubMed: 19698713]
40. Miot M, Reidy M, Doyle SM, Hoskins JR, Johnston DM, Genest O, Vitery MC, Masison DC, Wickner S. Species-specific collaboration of heat shock proteins (Hsp) 70 and 100 in thermotolerance and protein disaggregation. Proc Natl Acad Sci U S A. 2011; 108:6915–6920. [PubMed: 21474779]
41. Ma B, Tsai CJ, Haliloglu T, Nussinov R. Dynamic allostery: linkers are not merely flexible. Structure. 2011; 19:907–917. [PubMed: 21742258]
42. Li J, Sha B. Crystal structure of the E. coli Hsp100 ClpB N-terminal domain. Structure. 2003; 11:323–328. [PubMed: 12623019]
43. Li J, Sha B. Crystal structure of E. coli Hsp100 ClpB nucleotide-binding domain 1 (NBD1) and mechanistic studies on ClpB ATPase activity. J Mol Biol. 2002; 318:1127–1137. [PubMed: 12054807]
44. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993; 234:779–815. [PubMed: 8254673]
45. Sali A, Overington JP. Derivation of rules for comparative protein modeling from a database of protein structure alignments. Protein Sci. 1994; 3:1582–1596. [PubMed: 7833817]
46. Fiser A, Do RK, Sali A. Modeling of loops in protein structures. Protein Sci. 2000; 9:1753–1773. [PubMed: 11045621]

47. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J Comput Chem. 1983; 4:187–217.

48. Dominy BN, Brooks CL. Development of a generalized born model parametrization for proteins and nucleic acids. J Phys Chem B. 1999; 103:3765–3773.

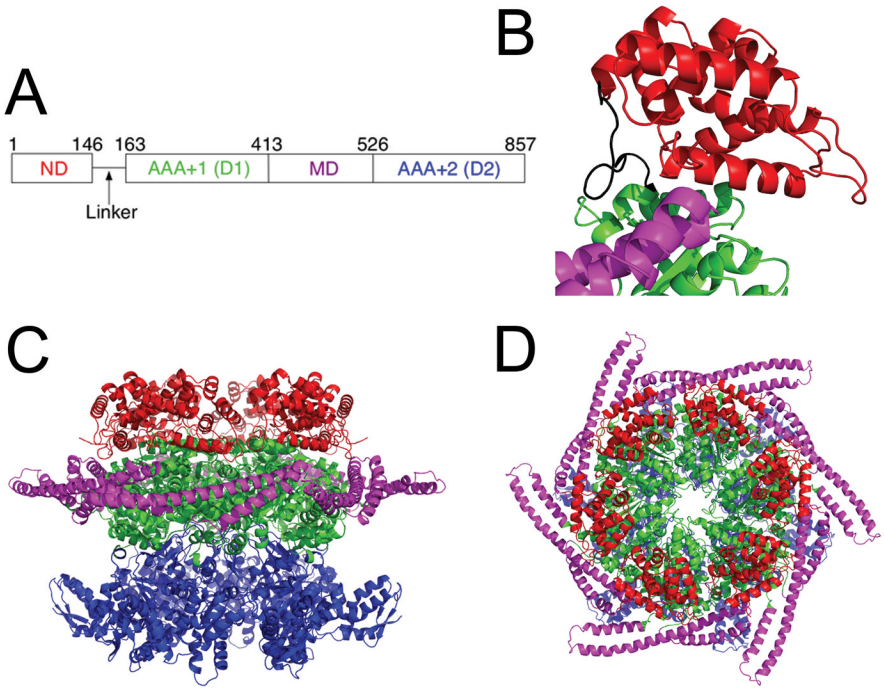

Figure 1.

(A) Domain structure of ClpB. The diagram shows the structural domains of ClpB as determined by X-ray crystallography[5]: the N-terminal domain (ND), D1 and D2 AAA+ modules, and the middle domain (MD). The residue numbers are given for *E. coli* ClpB and the position of the unstructured linker is indicated. (B, C, D) Structural model of ClpB from *E. coli*. The individual domains in each ClpB monomer are indicated with colors as in (A). The homology model of the *E. coli* ClpB monomer was built from the structures of the *E. coli* ClpB N-terminal domain[42], *E. coli* D1 AAA+ module[43], and chain B from the trimeric structure of *T. thermophilus* ClpB[5] using MODELLER software.[44,45] Conformation of the channel loops in D1 and D2 (visible in (D)) was refined with the automated loop-modeling module.[46] (B) Fragment of the ClpB structure showing the connection between ND and D1. The unstructured linker is shown in black. (C, D) Side and top view of the hexameric *E. coli* ClpB. The hexamer structure was obtained by assembling six homology-modeled monomers into a ring and performing a 500-step energy minimization with CHARMM and Generalized Born solvent model.[47,48]

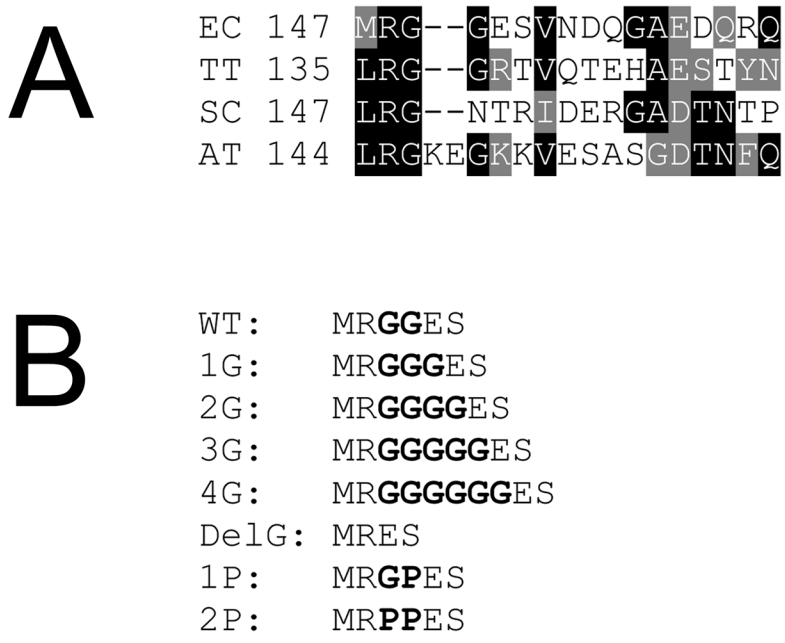

Figure 2.
(A) Sequence alignment of the ClpB N-terminal linkers from *Escherichia coli*, *Thermus thermophilus*, *Saccharomyces cerevisiae*, and *Arabidopsis thaliana*. (B) Modifications of the ClpB linker sequences produced in this work.

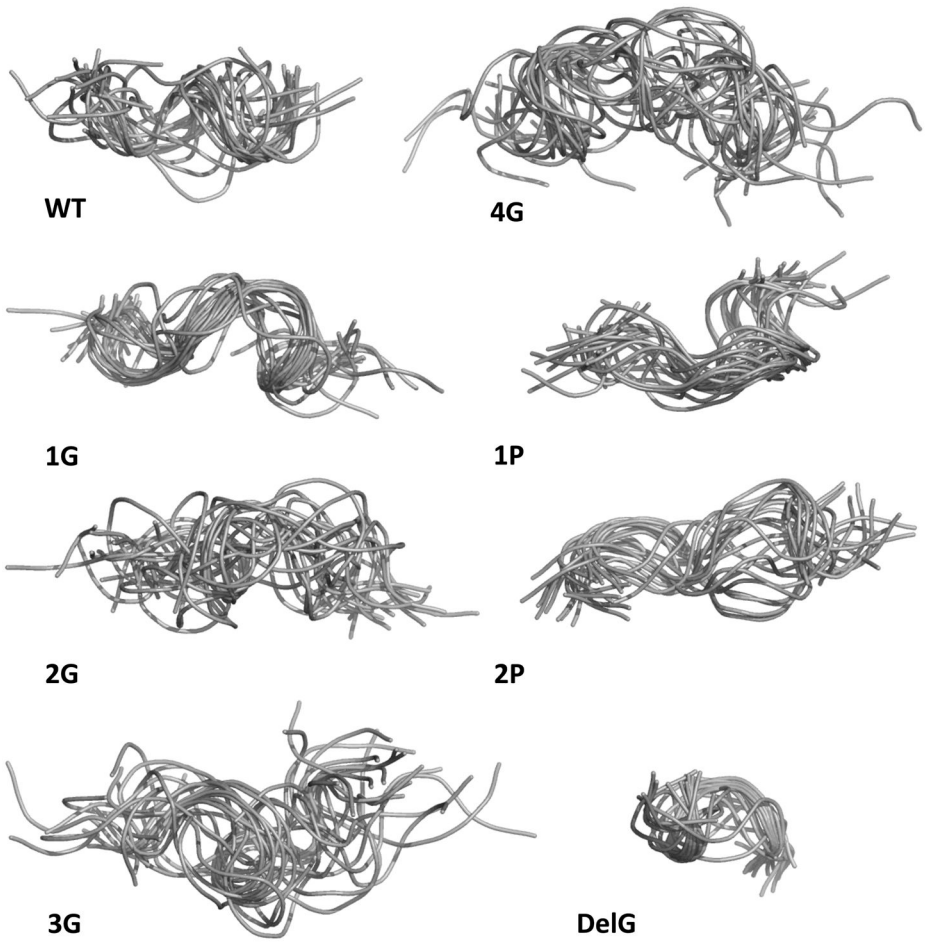

Figure 3.
Molecular dynamics simulation of the ClpB linker segments. Shown are the 0.05-μs backbone snapshots from the 1-μs long MD simulation of each linker sequence shown in Fig. 2B. The snapshots have been overlaid after a rotational and translational fit of the backbone to the starting structure of the peptide.

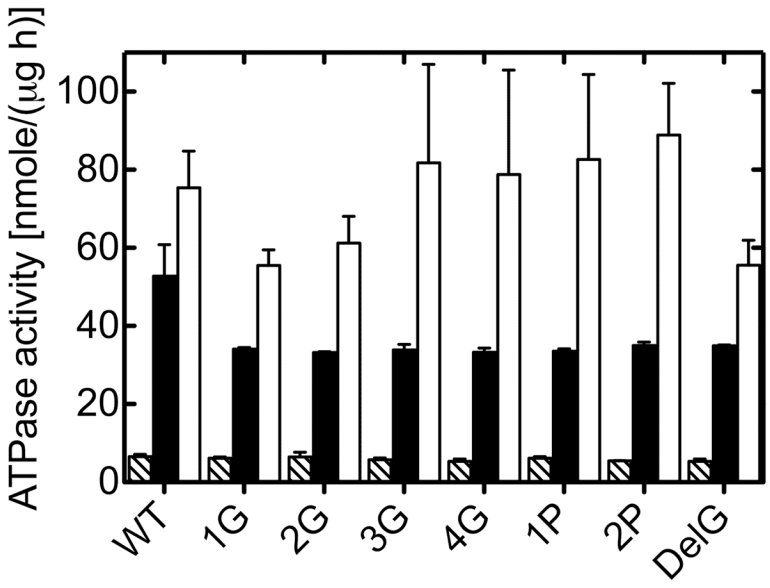

Figure 4.

ATPase activity of the modified linker variants of ClpB. The hydrolysis of ATP catalyzed by the ClpB variants was determined at 37 °C in the absence of other proteins (hatched bars), with α-casein (black bars), and with poly-lysine (white bars). The average values from three separate experiments are shown with the standard deviations.

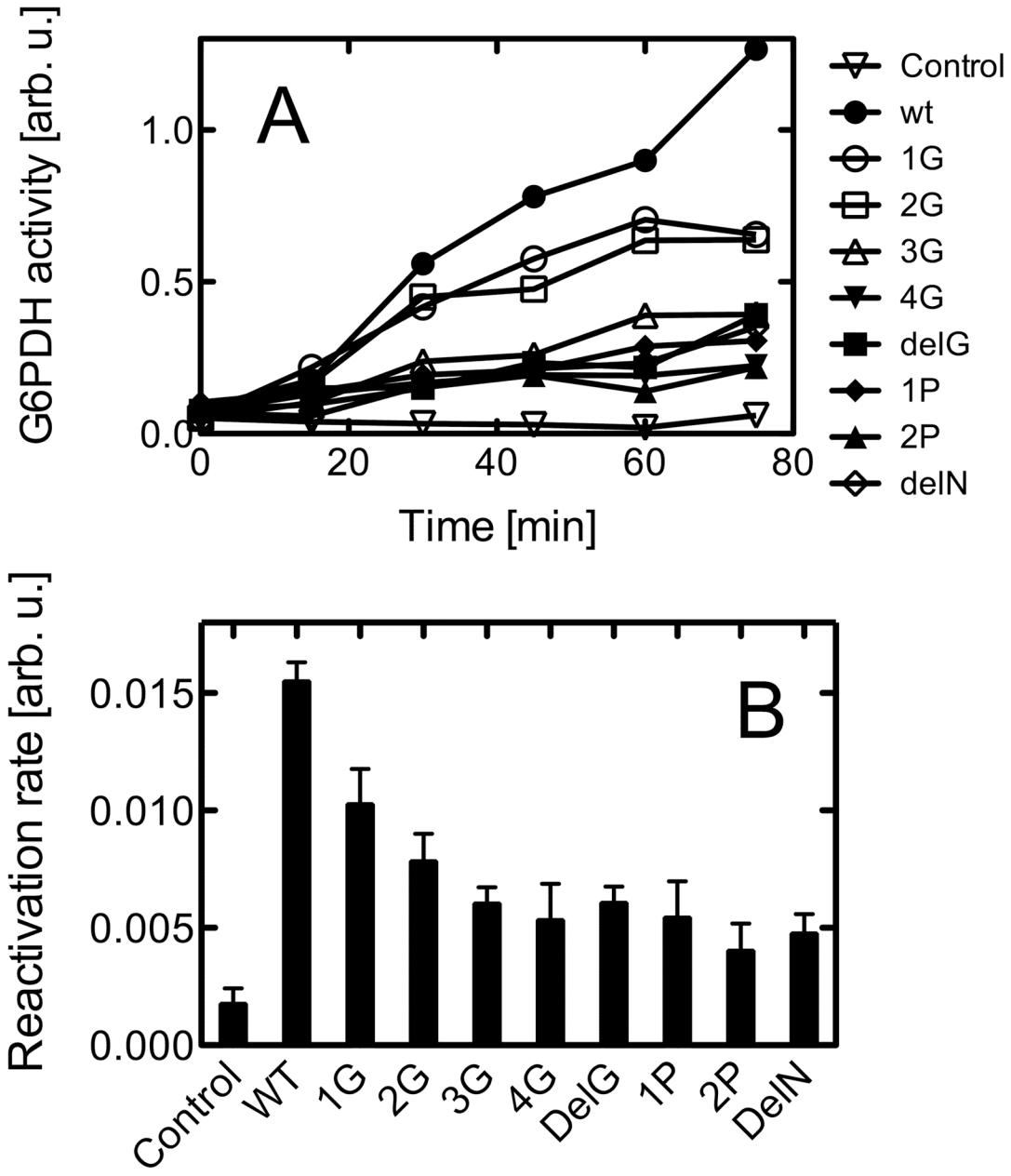

Figure 5.

Reactivation of aggregated glucose-6-phosphate dehydrogenase in the presence of ClpB and DnaK/DnaJ/GrpE. (A) A representative time-course of the reactivation of aggregated G6PDH without the chaperones (control) and with the indicated ClpB variants and DnaK/DnaJ/GrpE. (B) Initial rates of G6PDH reactivation (from the linear slopes of the data in (A)). The average values from three independent experiments are shown with the standard deviations.

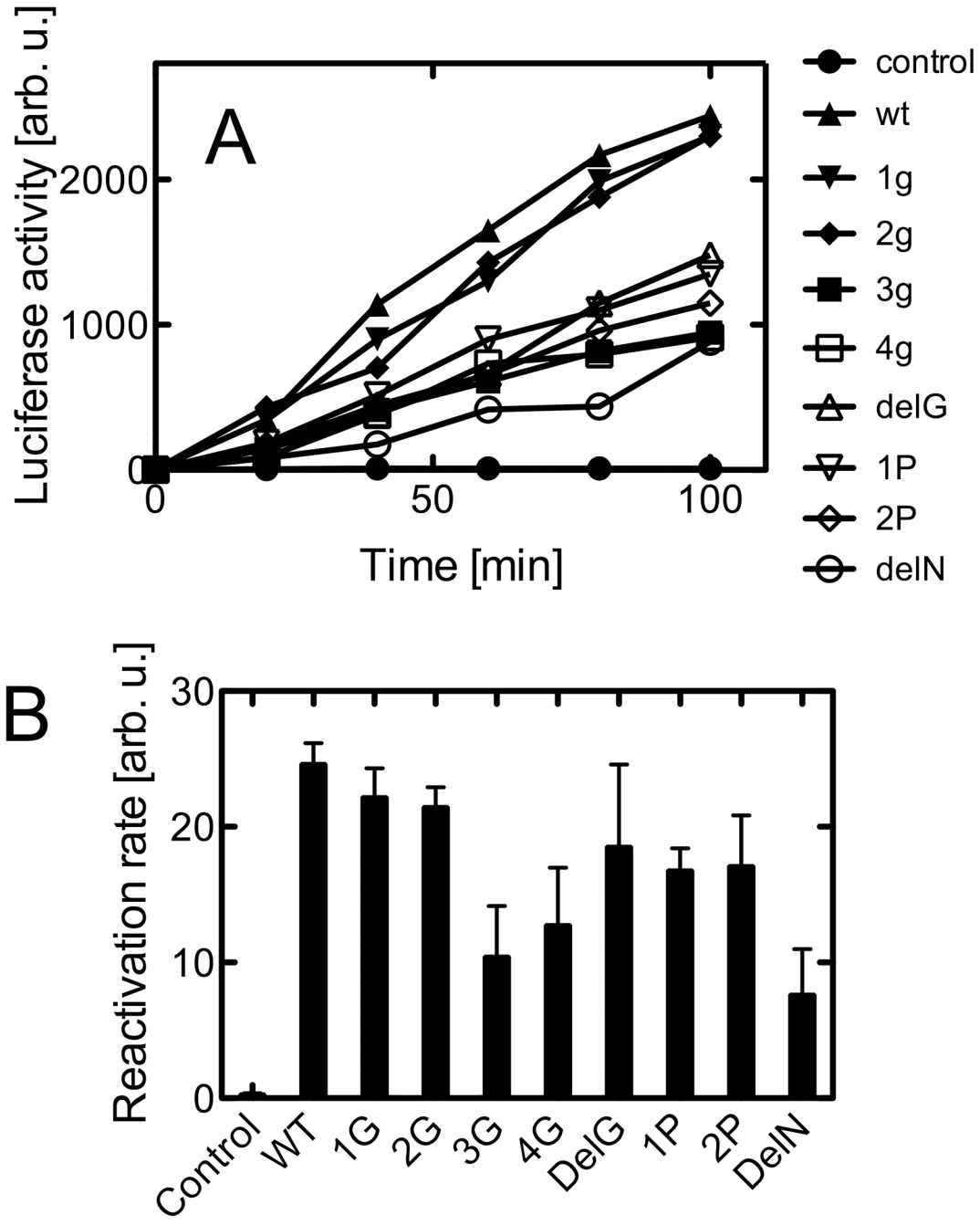

Figure 6.

Reactivation of aggregated firefly luciferase in the presence of ClpB and DnaK/DnaJ/GrpE. (A) A representative time-course of the reactivation of aggregated luciferase without the chaperones (control) and with the indicated ClpB variants and DnaK/DnaJ/GrpE. (B) Initial rates of luciferase reactivation (from the linear slopes of the data in (A)). The average values from three independent experiments are shown with the standard deviations.

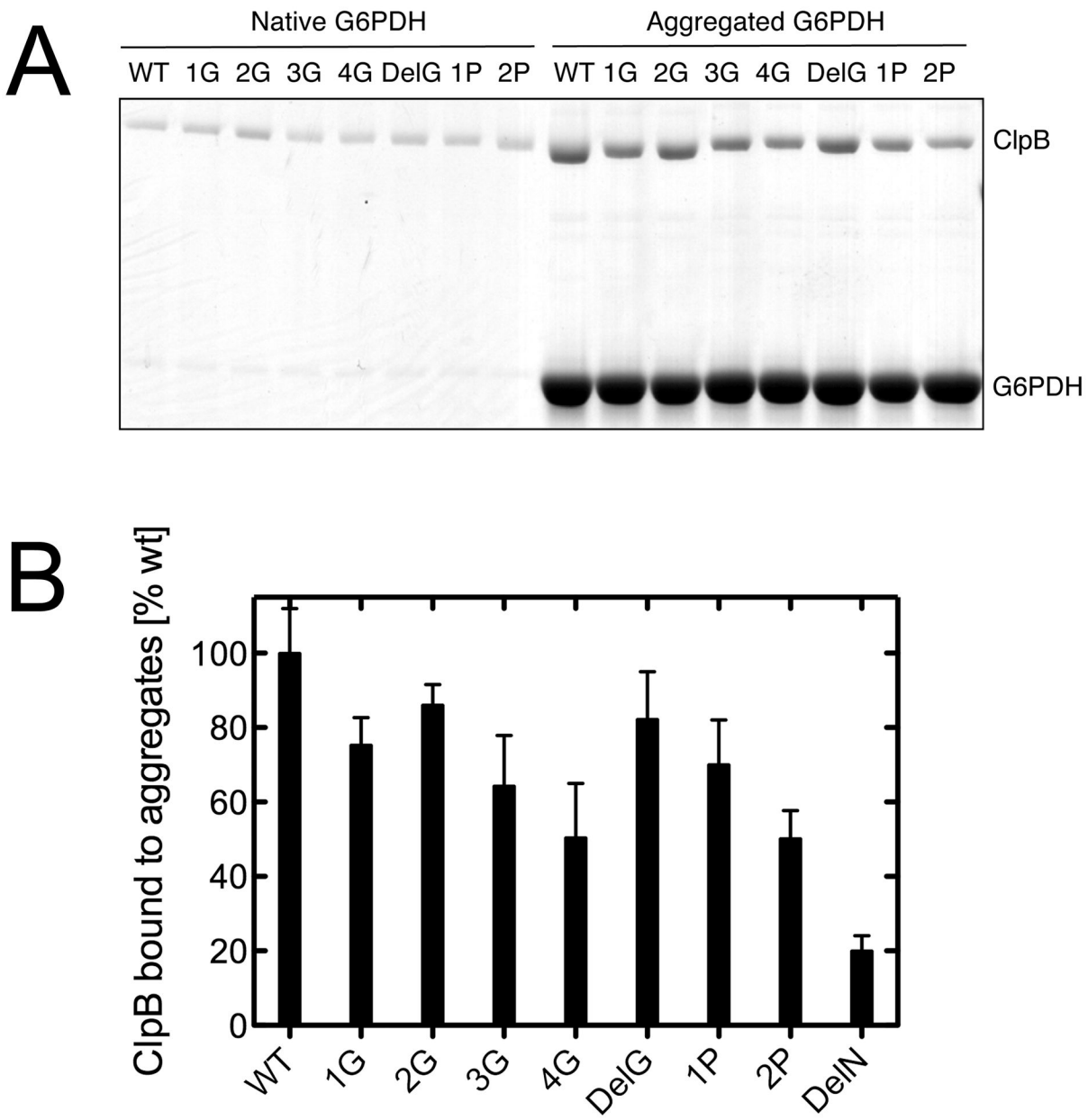

Figure 7.

Interactions of ClpB with aggregated G6PDH. Wt ClpB and its modified linker variants were incubated with the native or aggregated G6PDH in the presence of ATPγS. The solutions were passed through a 0.1- μm filter and the fractions retained on the filter were analyzed by SDS-PAGE with Coomassie stain. (A) A representative gel from the filtration experiment. (B) Image-intensity analysis of the ClpB band in the panel A normalized for the amount of G6PDH retained on the filter. The average relative amounts of ClpB bound to the aggregates from three independent experiments are shown with the standard deviations.

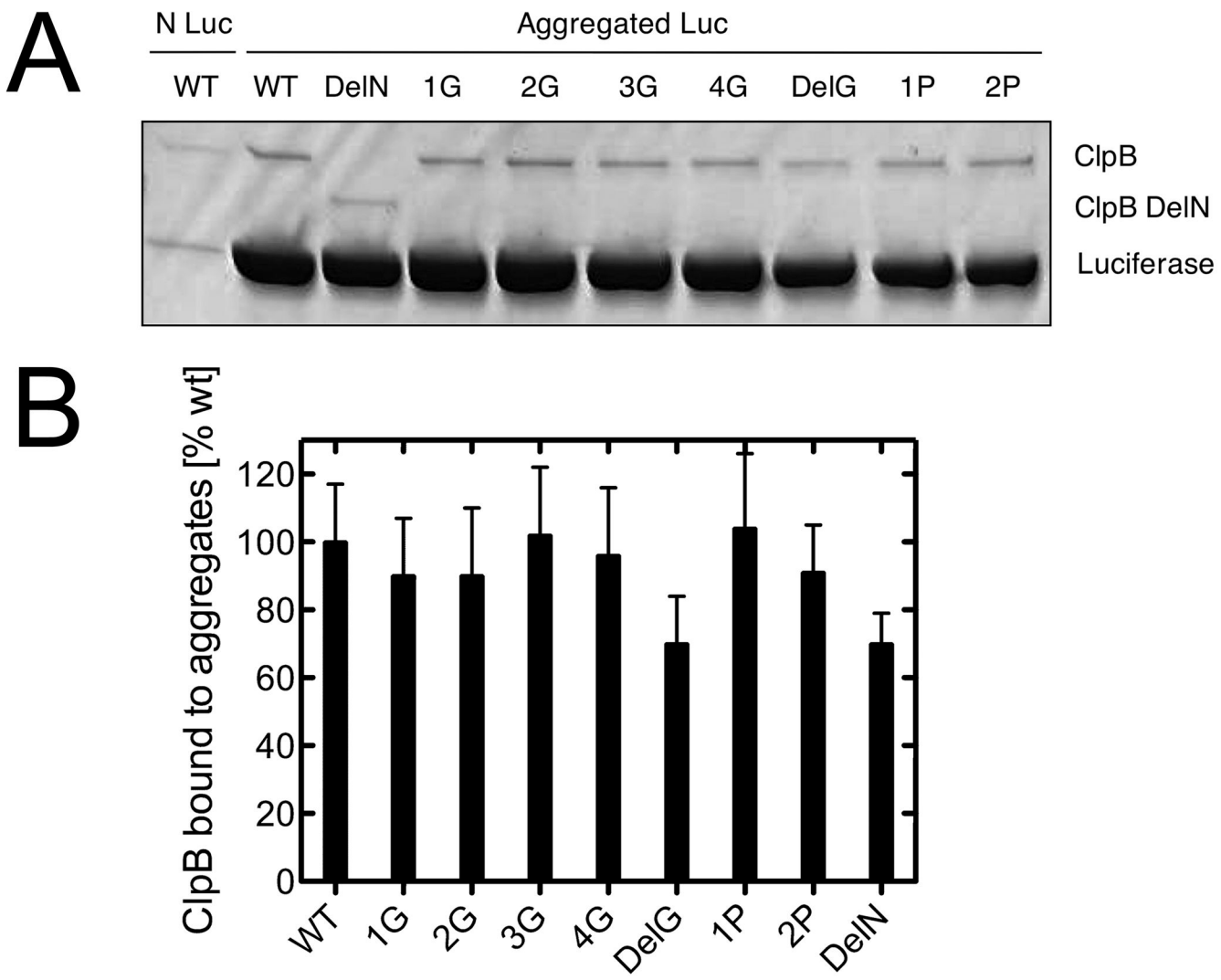

Figure 8.

Interactions of ClpB with aggregated luciferase. Wt ClpB and its modified linker variants were incubated with the native (N) or aggregated luciferase in the presence of ATPγS. The solutions were passed through a 0.1- μm filter and the fractions retained on the filter were analyzed by SDS-PAGE with Coomassie stain. (A) A representative gel from the filtration experiment. (B) Image-intensity analysis of the ClpB band in the panel A normalized for the amount of luciferase retained on the filter. The average relative amounts of ClpB bound to the aggregates from three independent experiments are shown with the standard deviations.