Published in final edited form as:

Cell Rep. 2021 May 25; 35(8): 109166. doi:10.1016/j.celrep.2021.109166.

Structural basis for aggregate dissolution and refolding by the *Mycobacterium tuberculosis* ClpB-DnaK bi-chaperone system

Yanting Yin[1,3], **Xiang Feng**[1,3], **Hongjun Yu**[1,3,4], **Allison Fay**[2], **Amanda Kovach**[1], **Michael S. Glickman**[2], **Huilin Li**[1,5,*]

[1]Department of Structural Biology, Van Andel Institute, Grand Rapids, MI, USA

[2]Immunology Program, Sloan Kettering Institute, New York, NY, USA

[3]These authors contributed equally

[4]Present address: Department of Biochemistry and Molecular Biology, School of Basic Medicine and Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[5]Lead contact

SUMMARY

The *M. tuberculosis* (Mtb) ClpB is a protein disaggregase that helps to rejuvenate the bacterial cell. DnaK is a protein foldase that can function alone, but it can also bind to the ClpB hexamer to physically couple protein disaggregation with protein refolding, although the molecular mechanism is not well understood. Here, we report the cryo-EM analysis of the Mtb ClpB-DnaK bi-chaperone in the presence of ATPγS and a protein substrate. We observe three ClpB conformations in the presence of DnaK, identify a conserved TGIP loop linking the oligonucleotide/oligosaccharide-binding domain and the nucleotide-binding domain that is important for ClpB function, derive the interface between the regulatory middle domain of the ClpB and the DnaK nucleotide-binding domain, and find that DnaK binding stabilizes, but does not bend or tilt, the ClpB middle domain. We propose a model for the synergistic actions of aggregate dissolution and refolding by the Mtb ClpB-DnaK bi-chaperone system.

Graphical abstract

*Correspondence: huilin.li@vai.org.

AUTHOR CONTRIBUTIONS

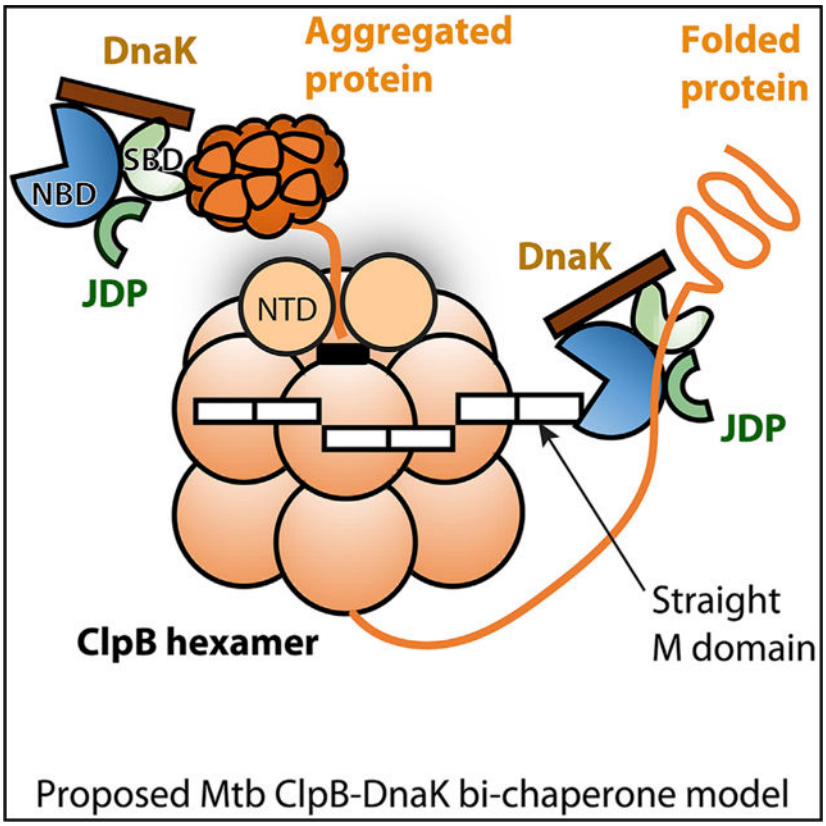
Y.Y., H.Y., M.S.G., and H.L. conceived and designed experiments. Y.Y. and A.K. purified proteins. Y.Y., H.Y., and X.F. performed cryo-EM, image processing and 3D reconstruction, and atomic model building. Y.Y. performed *in vitro* assays. A.F. performed all *in vivo* studies. Y.Y., H.Y., X.F., A.F., M.S.G., and H.L. analyzed the data and wrote the manuscript.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2021.109166.

DECLARATION OF INTERESTS

M.S.G. has received consulting fees and equity from Vedanta Biosciences and consulting fees from Takeda.



In brief

Yin et al. use cryo-EM to analyze the structure of the *Mycobacterium tuberculosis* ClpB-DnaK bi-chaperone system. They find that the Mtb ClpB middle domain does not bend or tilt when interacting with DnaK. They therefore propose that the Mtb DnaK facilitates protein folding following protein disaggregation by ClpB.

INTRODUCTION

Plants, bacteria, and fungi rescue aggregated proteins using a powerful ClpB/DnaK bi-chaperone system (Mogk et al., 2015). ClpB proteins form a family of AAA^+ ATPases that disaggregate and solubilize aggregated proteins (Doyle and Wickner, 2009; Lee et al., 2004; Maurizi and Xia, 2004; Shorter and Southworth, 2019). These disaggregases are traditionally called ClpB in bacteria, and the analogous enzyme in eukaryotes is called Hsp100. We will use ClpB throughout most of this paper in which we discuss the *Mycobacterium tuberculosis* (Mtb) hexamer. The bacterial ClpBs are composed of an N-terminal domain (NTD) and two nucleotide-binding domains (NBD1 and NBD2), and they form a ring-shaped hexamer. As expected of any canonical AAA^+ protein, both NBD1 and NBD2 are subdivided into an α/β domain (α/β domain 1 in NBD1 and α/β domain 2 in NBD2) and a helix domain (HD1 and HD2) (Lee et al., 2003), with each NBD containing the Walker A and Walker B motifs that are required for nucleotide binding and hydrolysis. An important feature that distinguishes the disaggregases from other ring-shaped ATPases is

the unique middle domain (M domain) inserted between helix 3 and helix 4 of the HD1 in NBD1 (Carroni et al., 2014; Desantis and Shorter, 2012; Lee et al., 2003). The M domain is composed of three α helices that are divided into two motifs: helix I and the first half of helix II form motif 1, and helix III and the second half of helix II form motif 2. These two motifs may have different functions (Oguchi et al., 2012).

Several recent structures have been published of the hexamers in complex with model substrate proteins, including the *E. coli* ClpB (Deville et al., 2019; Rizo et al., 2019), the Mtb ClpB (Yu et al., 2018), the *Saccharomyces cerevisiae* Hsp104 (Gates et al., 2017; Lee et al., 2019; Yokom et al., 2016), and the thermophilic fugus *Chaetomium thermophilum* Hsp104 (Heuck et al., 2016). These structures have revealed that these hexamers couple the ATP binding and hydrolysis to the threading of polypeptide through their central channels in order to disaggregate the substrate proteins. The substrate is pulled into the central channel by the coordinated action of 12 pore loops, two from each of the six protomers, with each loop having a highly conserved Tyr residue in the α/β subdomains 1 and 2. All ClpB disaggregases studied so far assemble an asymmetrical ring or a spiral around a substrate with a clear seam, because the six protomers are in different nucleotide and substrate translocation states (Shorter and Southworth, 2019). It was shown that the ClpB/Hsp104 hexamers alone, in the absence of any other co-chaperones, were able to engage and unfold GFP (Heuck et al., 2016) and translocate on a polypeptide deterministically with a strong force of more than 50 pN and with high processivity (Avellaneda et al., 2020).

Many bacterial pathogens, including *E. coli* and *M. tuberculosis*, produce two ClpB isoforms, a full-length and an NTD-truncated form, due to the presence of a conserved internal translation start site (Chow and Baneyx, 2005; Kannan et al., 2008; Nagy et al., 2010; Tripathi et al., 2018). The ClpB NTD is an all α-helical domain connected by a linker to NBD1 (Lee et al., 2003; Rosenzweig et al., 2015). The NTD-truncated ClpB is able to protect the bacteria during heat shock and reactivate aggregates, just like their wild-type counterparts, but their substrate affinity is an order of magnitude lower (Barnett et al., 2005; Chow et al., 2005; Clarke and Eriksson, 2000; Tripathi et al., 2018). Therefore, the ClpB NTD is considered to have an accessory function in the initial substrate recognition. Consistent with this idea, in the cryoelectron microscopy (cryo-EM) structure of the *E. coli* ClpB hexamer bound to casein, two to three NTDs were ordered and arranged spirally to guide the substrate into the central channel (Deville et al., 2017; Rizo et al., 2019). However, the NTDs of the Mtb ClpB bound to casein were disordered in the cryo-EM structure (Yu et al., 2018). Therefore, it is uncertain whether the Mtb ClpB NTD facilitates substrate entry in a similar manner.

Hsp70 is a multifunctional molecular chaperone that catalyzes the folding of nascent proteins, the assembly of protein complexes, and the prevention of protein misfolding and aggregation (Mayer and Gierasch, 2019; Rosenzweig et al., 2019). In bacteria, the Hsp70 proteins are traditionally referred to as DnaK, which collaborate with the Hsp40 protein DnaJ and the nucleotide exchange factor GrpE (Mayer and Gierasch, 2019). The current understanding is that DnaJ recruits an ATP-bound DnaK in open conformation to an unfolded protein substrate (Kampinga and Craig, 2010). Substrate binding stimulates ATP hydrolysis and substrate folding by DnaK and the chaperone moves into an ADP-bound,

closed conformation (Yang et al., 2017). The GrpE protein then binds the closed DnaK and stimulates the release of ADP and folded substrate, priming the DnaK for next round of substrate binding and folding (Bracher and Verghese, 2015). DnaK has an N-terminal nucleotide-binding domain (NBD) and a C-terminal substrate-binding domain (SBD) (Figures 1A and 1B) (Flaherty et al., 1990; Mayer and Gierasch, 2019; Rosenzweig et al., 2019; Zhu et al., 1996). The NBD is composed of four subdomains: IA, IIA, IB, and IIB. The SBD is composed of a β sheet subdomain (SBDβ) containing the peptide binding pocket and an α-helical subdomain (SBDα) that functions like a lid to cover the bound substrate (Zhu et al., 1996). The NBD and SBD are connected by a conserved linker loop (Bertelsen et al., 2009).

DnaK functions both upstream and downstream of ClpB (Goloubinoff et al., 1999; Schlieker et al., 2004). The bacterial ClpB-DnaK bi-chaperone system functions synergistically by physically coupling the two distinct processes of protein disaggregation and protein refolding (Fernández-Higuero et al., 2018; Mogk et al., 2015, 2018; Saibil, 2013; Zolkiewski, 1999). The major regulation point for DnaK is the ClpB M domain. A previous solution NMR titration of the NBD2-truncated, monomeric *T. thermophilus* ClpB fragment with DnaK identified two DnaK NBD residues, Arg-55 and Lys-285, that form hydrogen bonds with Tyr-484 and Tyr-494 of the ClpB M-domain motif 2 (Rosenzweig et al., 2013). However, it is unclear whether this fragment-based interaction reflects the regulatory interface of the full-length ClpB hexamer during ATP-driven substrate disaggregation. Furthermore, it is generally assumed that the M-domain orientation determines the functional state of ClpB: a horizontal orientation sustained by the head-to-tail interaction of neighboring M domains is the repressed state, and a tilted orientation that allows the M domain to interact with DnaK is the derepressed (activated) state (Carroni et al., 2014; Heuck et al., 2016; Oguchi et al., 2012). In this scenario, DnaK regulates ClpB by inducing a wave-like progression of M domains during the ClpB catalytic cycle (Carroni et al., 2014; Mogk et al., 2015). However, such a scenario seems to be at odds with the observations that the *E. coli* BAP-variant ClpB hexamer (a double Walker B mutant) bound to a peptide substrate has well-ordered and horizontally arranged M domains that are in a repressed-like state that disfavors DnaK binding (Deville et al., 2017). Yet, the same *E. coli* ClpB had weak M-domain densities, indicating disorders in either wild-type or K476C proteins (Rizo et al., 2019). In our own cryo-EM studies of the wild-type Mtb ClpB bound to casein and ATPγS, the M domains are flexible, particularly the M-domain motif 2 (Yu et al., 2018). For these reasons, how the M-domain orientation correlates with the ClpB functional states and how DnaK binds ClpB to regulate the activity need further investigation.

DnaK is not essential in normal growth conditions in *E. coli* but is essential in mycobacteria for cell growth and protein folding, thus forming a unique susceptibility node in the network of the mycobacterial proteostasis (Fay and Glickman, 2014). Mycobacteria stressed by host immunity or chemotherapy use the ClpB chaperone to asymmetrically sequester irreversibly oxidized proteins into daughter cells, leading to the rejuvenation of the daughter cells that have a lesser burden of such proteins (Vaubourgeix et al., 2015). The Mtb ClpB-DnaK bi-chaperone system has been reconstituted *in vitro*, revealing the unique and essential role of the J domain proteins DnaJ1 and DnaJ2 (Lupoli et al., 2016). We reconstituted the Mtb ClpB-DnaK complex *in vitro* and performed cryo-EM analysis, revealing a ClpB

conformation not observed in the absence of DnaK, the straight M-domain conformation when bound to DnaK and ATPγS, as well as the detailed ClpB-DnaK interface. We further characterized the ClpB-DnaK interface *in vitro* and *in vivo*, using ITC assay, aggregated luciferase reaction assay, heat-sensitivity assay, and protein localization and colocalization with aggregated proteins in the cell. By combining the existing structural knowledge and our observations, we propose a revised model for the cooperative functions of the Mtb bi-chaperone system.

RESULTS

In vitro reconstitution of the Mtb ClpB-DnaK bi-chaperone system

We separately expressed and purified full-length Mtb ClpB and DnaK in *E. coli* (Figures S1A and S1B) (see STAR Methods). We reconstituted the bi-chaperone system by including purified ClpB, DnaK, and cofactors DnaJ1, DnaJ2, and GrpE in the presence of denatured luciferase and 3 mM ATP. As expected, ClpB and DnaK synergistically reactivated denatured luciferase (Figures 1C and 1D). We note that the observed synergism between the two chaperones is small. It is possible that the *in vitro* system is incomplete, or luciferase is not an ideal substrate of the bi-chaperone system. Mtb ClpB alone in the absence of DnaK and other co-factors had no protein refolding or reactivation activity, in agreement with previous studies (Lupoli et al., 2016; Zolkiewski, 1999). Because recent structural and functional studies have demonstrated that ClpB/Hsp100 alone are able to engage and unfold protein substrates (Avellaneda et al., 2020; Deville et al., 2019; Gates et al., 2017; Heuck et al., 2016; Lee et al., 2019; Rizo et al., 2019; Yokom et al., 2016; Yu et al., 2018), the lack of reactivation activity suggests that ClpB threads the denatured luciferase through the central channel but is unable to refold the solubilized polypeptide into a functional protein.

For cryo-EM, we assembled the ClpB-DnaK bi-chaperone in complex with the model substrate casein *in vitro* by mixing the three proteins at a molar ratio of 1:1.2:1.2, ClpB hexamer:DnaK: casein, in the presence of 3 mM ATPγS. In cryo-EM 2D class averages, we found that most of the ClpB particles (69%) did not bind DnaK, and the remaining 31% that did contained DnaK molecules (Figures 1E, S1C, and S1D). Three DnaK molecules had relatively strong density in 3D reconstruction, indicating high occupancy; the other visible DnaK densities were weak, likely a result of low occupancy or flexible binding (Desantis et al., 2014; Doyle et al., 2015). *M. tuberculosis* expresses two J domain proteins, DnaJ1 and DnaJ2, (Lupoli et al., 2016), but DnaJ2 alone is sufficient to support the ClpB-DnaK activity *in vitro* (Lupoli et al., 2016). Addition of DnaJ2 into the assembly did not reveal extra density in 2D averages, suggesting that DnaJ2 binding on DnaK in the bi-chaperone system is highly flexible. We therefore focused our cryo-EM analysis on the binary ClpB-DnaK.

Mtb ClpB takes on a new transitional conformation upon DnaK binding

The 3D classification and refinement of selected ClpB particles containing DnaK density led to the reconstruction of three cryo-EM maps at average resolutions of 3.1, 3.5, and 3.3 Å (Figures S1E and S2). These structures accounted for about 30%, 20%, and 50% of the particle population used in 3D reconstruction respectively. In all three maps, the DnaK densities were very weak, so atomic models were built only in the ClpB hexamer region,

using the published Mtb ClpB structure as a template (Figures S3A–S3C; Tables S1 and S2). We found that protomers P3, P4, P5, and P6 of the three structures engage the substrate peptide with their respective pore loops similarly; the differences are found in protomers 1 and 2, the seam region of the hexamer ring, and their respective M domains (Figures 2A–2C). The 3.1 and 3.3 Å structures had essentially the same overall shape, bound nucleotides, and substrate peptide engagement as the previously observed low-affinity conformation 1 and the high-affinity conformation 2 of the Mtb ClpB determined in the absence of DnaK (Yu et al., 2018). The exception was that the structures we determined here in the presence of DnaK had improved densities in the P1 and P2 seam region and in the M domain. Therefore, we have referred to the 3.1 Å structure as conf I and the 3.3 Å structure as conf II. The 3.5 Å structure is a new conformation, with a disordered P1 pore loop of the NBD1 domain having little contact with the substrate peptide as in the low-affinity conf I, but with a well-ordered P1 pore loop of the NBD2 domain that tightly binds the substrate peptide, resembling the conf II (Figures 2D and S3D). Based on the nucleotide and substrate binding of the seam protomer P1, we assigned the 3.5 Å structure as a transitional state between conf I and II and refer to as conformation T. In conf I, NBD1 is in the ADP/apo state and NBD2 in the ATP state and none of the NBD pore loops engage substrate. In conf II, both NBD1 and NBD2 are in the ATP state, and both pore loops are bound to substrate. In conf T state, the NBD2 has engaged substrate but NBD1 has not yet exchanged ADP for ATP. Transitioning from the CT to CII, NBD1 would exchange ADP for ATP leading to its pore loop to bind to the substrate.

Because the bound nucleotides in conf T are the same as in conf I (Figures S3A–S3C), we suggest that structural differences between conf 1 and T are a result of DnaK binding. Therefore, DnaK likely modulates the ClpB catalytic cycle. Superposition of conf I, T, and II by aligning their respective P6 protomers revealed an overall similarity in the P3–6 region and key differences in P1 and P2 (Video S1). Overall, NBD2 and the region around the M domain in conf T resemble those of conf II, and the NBD1 of conf T resembles that of conf I. The transition from conf I to T is accompanied by an ordering of the M-domain motif 1 and coordinated movements of the NBD1 HD1 subdomain and the NBD2 α/β subdomain 2 in the P1 protomer. The NBD1 α/β domain 1 shifts up by about 15 Å to engage the substrate as ClpB further transitions from conf T to conf II (Figures 2A and 2D). Notably, the moving components of conf T are always in between the corresponding features of conf I and conf II, further supporting the assignment of conf T as an intermediate (Figure 2C). The M-domain motif 2 was disordered in all three conformations, and the motif 1 moved only slightly, suggesting that M domains do not undergo large movements during the ClpB catalytic cycle.

The conserved TGIP loop is essential for ClpB function

The movements of NBD1 and NBD2 in each protomer requires large conformational changes of their linker loop (Figures 2A–2D). The NBD1-NBD2 linker is not well conserved in the ClpB family, except for the TGIP region that is highly conserved (Figure 2B). The TGIP loop immediately follows the last helix (helix 4) of HD1 (HD1-Helix 4), which is itself preceded by the inserted M domain. This architecture provides a rationale on

how DnaK binding on the ClpB M domain may modulate HD1 and the TGIP loop and suggests a crucial role of the TGIP loop in ClpB function.

We found that TGIP-AAAA mutant of ClpB could be expressed and purified from *E. coli* (Figure S4A). This mutation did not significantly affect the ATP hydrolysis activity but completely abolished ClpB's ability to reactivate aggregated luciferase (Figures 2E and S4E). We tested the *in vivo* functionality of a ClpB that lacked the TGIP motif by complementing *M. smegmatis* D*clpB* with a plasmid encoding either wild-type ClpB or ClpB TGIP-AAAA. We found that the Mtb *clpB* allele failed to complement to the level of *M. smegmatis clpB* heat resistance at 53°C and proceeded with utilizing the *M. smegmatis clpB* allele, because their sequence identity is about 91% and all the residues of interest are conserved. Despite equivalent protein expression of all constructs by western blot (Figure 2F), the ClpB TGIP-AAAA mutant was unable to complement the heat sensitivity of the *ΔclpB* strain (Figure 2G). When examined by fluorescence microscopy in live cells, mutant ClpB fused to mCitrine was partially soluble, with some aggregation (Figure 2H). We assessed the mutant ClpB's ability to localize to protein aggregates by induced expression in the *M. smegmatis* of mCerulean-AL09 (Fay and Glickman, 2014), a fluorescent aggregating protein fusion. Similar to wild-type ClpB-mCitrine, the mutant ClpB proteins co-localized to aggregate foci (Figure S4F). We note that some of the co-localized mutant ClpB proteins may be just co-aggregation with AL09, because the mutant ClpB proteins alone were partially aggregating (Figure S4G). Taken together, our study has revealed a role of the TGIP loop in mediating ClpB's disaggregation function.

The ClpB NTD initiates substrate engagement via hydrophobic interactions

The ClpB N-terminal domain is known to affect substrate engagement (Rosenzweig et al., 2015), but the exact mechanism has been unclear. In our previous work on ClpB bound to casein, the NTD was not resolved (Yu et al., 2018). We observed strong NTD densities in the side views of the class averages of ClpB bound to DnaK (Figure 1E), but the NTDs were averaged out in the whole-complex 3D reconstructions. Therefore, we applied focused 3D classification on the NTD ring on the main dataset (417,884 particles) and obtained a 3D map at 3.2 Å overall resolution, which contained a well-defined NTD density with secondary structure features in protomer 3 (P3) (Figures 3A and S5). We speculate that the substrate may primarily engage only one of the six available NTDs and that the substrate bound NTD is stabilized and becomes visible in the 3D map. An atomic model of the NTD was built based on the crystal structure of a thermophilic fungus Hsp100 protein (Michalska et al., 2019).

ClpB NTD is composed of eight α helices, with α helices 1 (H1) and 5 (H5) adjacent to the substrate peptide entry (Figure 3B). In H1 and H5, six residues—Gln-11, Thr-15, Leu-18, Ser-22, Leu-88, and Thr-92—that likely engage the substrate peptide are either hydrophobic or polar, perhaps accounting for their ability to interact with the unfolded/aggregated substrate peptide. We examined the effect of replacing these residues with arginine. We do not know the effects of Q11R and T15R mutations, because ClpB carrying these mutations could not be expressed and purified in *E. coli.* For the four mutant proteins that we were able to purify (L18R, S22R, L88R, and T92R), their ATP hydrolysis activities were essentially

preserved in the presence of substrate (Figure S4E). We found that the S22R and T92R mutations significantly reduced the ClpB-mediated aggregate luciferase reactivation (Figure 3C). Leu-18 and Leu-88 were closest to the substrate entry, and their replacement by arginine (L18R and L88R) totally abolished ClpB activity. The lower-than-DnaK-alone activity of these two mutant ClpB may be due to the inactive proteins binding and shifting soluble luciferases into aggregates, or the inactive proteins sequestering DnaK hence preventing its activity, or an ATP depleting effect of the ineffective proteins leading to impaired luciferase activity. These two hydrophobic residues are likely primarily responsible for recruiting the aggregated substrate into the ClpB central chamber.

The identified substrate binding surface in the NTD is consistent with that in the *E. coli* ClpB (Deville et al., 2017; Rizo et al., 2019) (Figure S4H). However, a recent study suggested that a hydrophobic groove formed by Leu-97, Leu-101, Phe-140, and Val-113 in the Mtb ClpB NTD may bind substrate (Tripathi et al., 2018). This region is away from the substrate initial engagement site identified in the structural studies (Figure S4I), and it is unclear how this site is involved in substrate binding. We presume that an aggregated substrate may make multiple contacts with the NTD such that both sites may contribute to substrate recognition. In support of this possibility, we know that the hydrophobic residues in H1 and H5 and those in the loops between H3-H4 and between H4-H5 of the *T. thermophilus* ClpB NTD, as well as the acidic residues of the *E. coli* ClpB NTD, interact with substrates (Barnett et al., 2005; Rosenzweig et al., 2015).

M domain does not bend, and M-domain motif 1 mediates multiple charge-charge interactions

DnaK binding significantly stabilized the ClpB M domain such that the entire ClpB M-domain density was visible in some 2D class averages (see Figure 1E), but the M domain and the bound DnaK had weak densities in the whole-particle 3D reconstruction. We therefore performed focused 3D classification on the P6 M domain and the bound DnaK on the main dataset (417,884 particles) to select the particles with clear density on this region and then a full-particle refinement to obtain the 3D map of ClpB with a clear DnaK density (Figure S5). This map had an overall resolution of 3.7 Å and well-defined M-domain density, but the DnaK density was at much lower resolution of 1–2 nm. The ClpB M domain contains three α helices (helix I, II, and III) and can be divided into motif 1, composed of helix I and first half of helix II, and motif 2, composed of helix III and the second half of helix II. It appears that the DnaK binding greatly improved the M-domain density such that the atomic model could be built for the entire M domain, even including the loop connecting helix I and helix II (Figure 4A). In the middle of the M domain, at the boundary between motif 1 and motif 2, a strong hydrophobic interaction involving the conserved Pro-410, Ile-413, Leu-459, and Trp-463 stabilized motif 1 and oriented motif 2 outward in order to interact with DnaK (Figure 4A). Unexpectedly, the M domain of P6 in Mtb ClpB conf T, in the presence of its functional partner DnaK and substrate casein, was largely horizontal, projected outward away from the ClpB main body, and did not bend or tilt upward. This is very different from the bent M domain structure that tilts up toward the NBD1 as observed in the K476C M-domain mutant *E. coli* ClpB that mimics the transient state of ClpB activation by Hsp70, as well as the horizontal conformation stabilized by the head-to-tail

interactions between motif 1 and motif 2 of neighboring M domains (Deville et al., 2019) (Figures 4B and S6A).

The complete M-domain model reveals numerous charge-charge interactions between M-domain motif 1 and the ClpB main body (Figures 4A and 4C). Specifically, the P6 M-domain helix I interacts with the upper P5 NBD1 via two clusters of charge interactions. These involve Asp-184 and Arg-188 of P5 NBD1 and Arg-422 and Asp-414 of P6 M-domain helix I in one cluster, and Arg-352 of P5 NBD1 and Glu-426 of P6 M-domain helix I in the other cluster (Figure 4A). Although Asp-414 and Glu-426 had no side-chain densities, Asp-414, Arg-188, Glu-426, and Arg-352 are essential for ClpB function (Yu et al., 2018), suggesting all these residues interact with NBD1 during the functional cycle. Consistently, we found here that the R422A mutant ClpB could not cooperate with DnaK to refold a protein substrate (Figure 4D). The loop of the P6 M-domain motif 1 appears to be strategically located to form two patches of intermolecular charge-charge interactions with P5 NBD1 and one patch of intramolecular charge-charge interactions with P6 NBD2 (Figure 4C). The first intermolecular patch is triangular, involving Glu-433 and Glu-434 in the loop region of the P6 M-domain motif 1 and the P5 NBD1 Arg-365, which stabilizes the motif 1 loop. The second intermolecular patch is a salt bridge between P6 M-domain motif 1 Arg-441 and the P5 NBD1 Asp-368. We also observed an intramolecular charge patch involving the P6 M-domain motif 1 Glu-436 and Glu-440 and P6 NBD2 Arg-780. This intramolecular interaction may coordinate NBD2 and the M domain and is consistent with the finding that NBD2 and the M domain of conf T resemble those of conf II.

A previous mutagenesis screen demonstrated that Glu-432 of the *E. coli* ClpB, equivalent to the Mtb ClpB Glu-433, is essential for function (Oguchi et al., 2012). To test the importance of the other conserved charge-charge interactions, we either individually reversed the charges or mutated the residues to alanine, constructing ClpB variants having an R365A, E434K, D368R, or E436R mutation. All mutant ClpBs assembled into hexamers and could be purified by gel filtration chromatography (Figure S4A). None of these mutations significantly affected the ClpB ATPase activity (Figure S4E). We found all mutations abolished the aggregated luciferase reactivation function of the Mtb ClpB with the exception of E434K, which only reduced ClpB function (Figure 4D). Because Glu-434 is involved in the triangular charge interaction with Arg-365 and Glu-433, the E434K mutation simply changed the “two negatives versus one positive” interaction to a “one negative versus two positives” interaction. Thus, the triangular charge interaction was maintained in this mutant ClpB, perhaps explaining the dampened effect of E434K on the Mtb ClpB function. Taken together, DnaK binding greatly stabilized the ClpB M domain of P6 in conf T, leading to a nearly complete M-domain atomic model and the revelation that the ClpB M-domain motif 1 is involved in an extensive network of charge-charge interactions that are crucial to ClpB function.

The DnaK interface with the Mtb ClpB hexamer

Previous NMR study of the *T. thermophilus* system using truncated protein domains has shown that the ClpB M domain motif II interacts with DnaK NBD (Rosenzweig et al., 2013). At a lower display threshold, four DnaK densities can be observed bound to four

ClpB protomers, with the fourth DnaK density (bound to P1) being the weakest (Figure S7). We masked, subtracted the signal of the rest, and then locally refined the three ClpB MD-DnaK NBD with stronger densities on ClpB protomers P4, P5, and P6 and obtained three improved maps at a resolution of about 7 Å, sufficient to show some features of secondary structural elements. The resolution improvement from 1 to 2 nm in the full-particle refinement indicates that the limited resolution of the local region is likely due to the long and partially flexible M domain onto which DnaK is attached, rather than due to unstable interaction between DnaK NBD and ClpB M domain. These 3D maps represent a visualization of the ClpB-DnaK bi-chaperone system (Figures 5A and 5B). We found all three DnaK-bound M domains were in the straight configuration. However, only the NBD of DnaK was visualized; the substrate binding domain was missing in all three maps. The three maps were highly similar, suggesting a similar binding interface. We flexibly docked the crystal structure of the Mtb DnaK NBD into a selected EM map (on protomer P6) with Rosetta program to derive a plausible model for the ClpB-DnaK interaction. The structure model was further fitted into the other two locally refined maps and combined with the structural model of the ClpB. A composite map was then produced by combining ClpB density map with the three local maps (Figures 5A and 5B; Video S2). The structural model indicates that the ClpB-DnaK interaction is mediated by two residues of the DnaK NBD (Met-229 and Tyr-257) and three residues of the ClpB M-domain motif 2 (Leu-496, Arg-503, and Tyr-504) (Figure 5C). Interestingly, the Mtb ClpB M-domain interface on DnaK NBD is limited to the subdomain IIB, different from the *T. thermophilus* system in which ClpB M domain interacts with both the IB and IIB subdomains of DnaK NBD (Figure S6B) (Rosenzweig et al., 2013).

To validate the interface, we individually replaced those five residues with alanine and measured aggregated luciferase reactivation with the mutant ClpB and DnaK proteins. We found that the purified Y257A and M229A mutant DnaK proteins were active in ATP hydrolysis (Figure S6C) but had luciferase reactivation activity similar to those in the absence of ClpB (Figure 5D). Two of the ClpB mutations, L496A and Y504A, but not R503A, abolished the ClpB-mediated activity (Figure 5E). ClpB R503 is close to but does not contact DnaK D227 in our structural model. These results suggest that the four mutations disrupted the ClpB-DnaK interaction. In agreement with our observation, a Y503A mutation in *T. thermophilus* or *E. coli* ClpB, equivalent to Mtb ClpB Y504A, has been reported to abolish their respective ClpB interactions with DnaK (Rosenzweig et al., 2013). R503A did not greatly affect the ClpB activity, likely because the interface is not primarily mediated by the charge pair.

We next validated the ClpB-DnaK binding interface by isothermal titration calorimetry. The purified ClpB M-domain motif 2 protein (residues 467–518) bound to the purified DnaK NBD (residues 1–359) with a stoichiometry of about 1 and a K_D of about 79 μM (Figure S8), which is largely consistent with the reported K_D of 25 μM derived from an NMR titration using full-length DnaK and the NBD2-truncated *T. thermophilus* ClpB (Rosenzweig et al., 2013). There was no detectable binding between the Mtb M-domain motif 2 carrying a L496A or Y504A mutation and the WT DnaK NBD, or between WT ClpB M-domain motif 2 protein and the DnaK NBD carrying either a M229A or Y257A mutation. However, the Mtb ClpB M-domain motif 2 with R503A interacted with WT DnaK NBD with a K_D of

about 80 μM. These *in vitro* binding results are consistent with the aggregate luciferase reactivation assay.

We further examined the functional importance of the ClpB-DnaK interface in *M. smegmatis*. We found that ClpB proteins carrying the L496A and Y504A mutations and the DnaK proteins carrying Y257A and M229A were synthesized in similar amounts and were distributed in cells similarly to their respective WT counterparts, and all four mutant proteins co-localized with the model protein aggregate AL09 (Figure S9). These observations suggest that the interface mutations do not affect protein expression or localization to protein aggregates. We then tested bacterial survival as a measure of the dynamic functionality of these interface mutations under heat stress (Fay and Glickman, 2014). The DnaK M229A mutation (Figure 5F) and the two ClpB mutations (Figure 5G) rendered *M. smegmatis* heat sensitive at 53°C. The viability and heat sensitivity implicate these interface mutations in DnaK-ClpB cooperation during protein denaturation. It is unclear why the DnaK Y257A mutant was protected against heat nearly as well as the WT. Overall, our experimental results underscored the importance of these identified residues in mediating the physical interaction and the functional synergism between ClpB and DnaK.

DISCUSSION

Our cryo-EM analysis of the Mtb ClpB-DnaK bi-chaperone in the presence of a model substrate, and ATPγS has revealed an additional conformation of ClpB and the crucial role of the conserved TGIP loop in ClpB function. The ClpB M domain mediates ClpB and ClpB-DnaK functions, but the mechanism of regulation has not been well understood (Mogk et al., 2015; Rosenzweig et al., 2013; Seyffer et al., 2012). We have derived a plausible interface between the ClpB M domain and DnaK NBD and validated the contact surfaces by mutagenesis and functional assays. Importantly, we found that DnaK binding stiffens the coiled coils of the M domain, rather than bending and tilting that domain as observed in a hyperactive mutant *E. coli* ClpB in the absence of DnaK (Rosenzweig et al., 2019; Saibil, 2013). We suggest that the straight conformation is active, because (1) we used slowly hydrolyzable ATP analog (ATPγS) and found some ATPγS were hydrolyzed in the structure, (2) a stretch of substrate peptide was observed in the ClpB translocation channel and engaged with one NTD of ClpB where the substrate was initially recruited; and (3) the same setup was shown to activate denatured luciferase *in vitro*. However, the atomic level interface between ClpB and DnaK, whether the DnaK-induced stabilization and straightening of the ClpB M domain is a universal mechanism (considering the strong conservation in the Hsp100-Hsp70 system) or this is species-specific mechanism requires further investigation, and whether the straight M domain is an active configuration await further investigation.

Structural alignment of the Mtb model with the *E. coli* DnaK-DnaJ structure (PDB: 5NRO) indicates that DnaJ binds DnaK NBD at the opposite end of where the M domain binds (Figure S6D). Therefore, the ClpB M domain and DnaJ may not interfere each other on DnaK. The *E. coil* GrpE interacts with DnaK via an extensive interface involving DnaK subdomains IA, IB, and IIB (Mayer and Gierasch, 2019). Because the Mtb ClpB M domain in our model contacts the DnaK IIB domain, if the interaction between GrpE and DnaK is

similar in the Mtb system, GrpE may interfere with ClpB M-domain binding. Indeed, alignment of our Mtb model with the *E. coil* GrpE-DnaK structure (PDB: 1DKG) shows that GrpE binds DnaK NBD from below, but the binding site may partially overlap with that of the ClpB M domain (Figure S6D). Therefore, DnaK may need to dissociate from ClpB in order to interact with GrpE for nucleotide exchange. Further work is needed to clarify this possibility.

A potentially important observation of this work is that the Mtb ClpB M domain may bind to only a single subdomain, specifically, the lobe II subdomain B (IIB) of the DnaK NBD. The interface is determined to a relatively low resolution of sub-nm. An atomic level interface awaits further investigation. But if the modeled interface is true, the Mtb ClpB-DnaK binding mode would be different from that reported in *T. thermophilus*, in which the ClpB M domain contacts both motifs IB and IIB of the DnaK NBD (Rosenzweig et al., 2013). The different binding modes may have significant functional consequence because lobes I and II of DnaK NBD toggle between the closed and open conformations to allow nucleotide and substrate binding and release (Mayer and Gierasch, 2019; Rosenzweig et al., 2019). The ClpB M domain binding to both IB and IIB would bridge the two NBD lobes and hinder their conformational toggling. In contrast, in *M. tuberculosis*, ClpB M-domain binding to only one DnaK motif (NBD IIB) would not impede the conformational toggling or DnaK activity, consistent with the observed synergism of the Mtb bi-chaperone system.

In our cryo-EM 3D reconstruction of the Mtb ClpB-DnaK, the DnaK SBD was averaged out due to its flexibility or the multiple conformations of the domain. It is known that the SBD undergoes a very large conformational change upon ATP hydrolysis and substrate binding (Bertelsen et al., 2009; Kityk et al., 2012, 2018; Swain et al., 2007). Because the long arm of the ClpB M domain positions the DnaK far away from the main body of the ClpB, and also because the DnaK SBD is positioned furthest from ClpB (Video S3), SBD is unobstructed by the ClpB hexamer and is free to swing through a large range to bind substrate and then to release the folded product (Figure 6A). Because the Mtb DnaK is positioned to the side of the ClpB hexamer far away from the substrate entry on the top of the hexamer, we suggest that some aggregates may be recruited by the ClpB NTD without the need for DnaK, while others may need help from DnaK that are not bound to ClpB M domain, and that DnaK docked on the ClpB M domain may cooperate with DnaJ1/2 to bind and fold polypeptide emerging from the bottom of the ClpB hexamer (Figure 6B). This scenario accounts for the synergism between ClpB and DnaK in the luciferase reaction, is consistent with the fact that Mtb ClpB alone engages and unfolds—but does not fold—protein substrates (Figure 1D), and agrees with the recent stopped-flow FRET study of the *E. coli* ClpB-DnaK system demonstrating that DnaK facilitates the release of polypeptide substrate from ClpB (Durie et al., 2018).

It is well established that DnaK can function upstream of ClpB to target aggregates to ClpB. This function was described for a hybrid variant (ClpVB) of the *E. coli* ClpB in which the ClpB substrate-targeting NTD was replaced by the NTD of the *Vibrio cholerae* Hsp100 protein ClpV (Seyffer et al., 2012). For certain hard-to-unfold aggregates such as prion fibrils, the initial substrate binding and targeting function of DnaK is essential for the ClpB-DnaK system (Winkler et al., 2012). In our model, the DnaK molecules that function down

stream of ClpB by binding on ClpB M domain are distinct from the DnaK molecules that function upstream of ClpB to target substrate to ClpB (Figure 6B). The existing model relies on the ClpB M domain to bend and tilt upward such that a DnaK molecule docked on the ClpB M domain is positioned near the ClpB substrate entry; whereas our structural analysis has demonstrated that the Mtb ClpB M domain is straight and positions the bound DnaK far away from the ClpB hexamer. Further study is required to clarify what features observed in the Mtb ClpB-DnaK system is generalizable to other organisms.

Finally, we observed that up to four DnaK molecules bind to a ClpB hexamer. It is unclear whether DnaK functions processively on ClpB M domain or hops on and off. It is also unknown the exact stoichiometry of ClpB versus DnaK. These important questions await further investigation.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Huilin Li (Huilin.Li@vai.org).

Materials availability—Plasmids and bacterial strains generated in this study are available by contacting the lead contact.

Data and code availability—The 3D cryo-EM maps of the Mtb ClpB in conformers I (3.1 Å), T (3.5 Å), and II (3.3 Å) have been deposited in the Electron Microscopy Data Bank (EMDB) under accession codes EMD-21554, EMD-21556, and EMD-21555, and their corresponding atomic models have been deposited in the Protein Data Bank under accession codes 6W6G, 6W6I, and 6W6H. The EMDB codes for the focused 3D classification derived 3D maps of the Mtb ClpB in conformation T with the NTD density (3.2 Å) and with the full M-domain and DnaK NBD densities (3.7 Å) are EMD-21557 and EMD-21553, respectively, and the PDB accession codes for ClpB with NTD and ClpB with full M domain bound to the DnaK NBD are 6W6J and 6W6E, respectively. The EMDB codes for the Mtb ClpB in conformation T with three DnaK NBDs are EMD-23206 and the PDB accession code is 7L6N. All relevant data are available from the authors.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

For protein production, *E. coli* BL21(DE3) competent cells were cultured at 37°C in LB medium OD_{600} of about 0.8 before cooling to 16°C. Protein expression was induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16 h.

We used *M. smegmatis* for *in vivo* study. Mycobacterial cells were grown to an OD_{600} of 0.2–0.3 in 7H9-albumin-dextrose-NaCl with 0.05% Tween 80 overnight at 37°C and diluted to an OD_{600} of 0.1 in fresh medium.

METHOD DETAILS

Protein production and purification—The construction of the N-terminal 6xHis-SUMO tagged Mtb ClpB and DnaK, their expression in *E. coli*, and the purification

procedure were described previously (Lupoli et al., 2016; Yu et al., 2018). Briefly, *E. coli* BL21(DE3) competent cells (Agilent) transformed with the N-terminal His-SUMO–tagged Mtb ClpB and DnaK were cultured in LB medium supplemented with 100 μg/mL ampicillin and were grown to OD_{600} of about 0.8 before cooling to 16°C. Protein expression was induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16 h. Collected frozen cells were lysed in buffer containing 25 mM Tris, pH 8.0, 40 mM NaCl, 10 mM $MgCl_2$ (Buffer A) and the EDTA-free protease inhibitor cocktail tablets. After centrifugation, the proteins were purified from the supernatant using nickel-nitrilotriacetic acid (Ni-NTA) affinity chromatography. We used buffer A plus 300 mM imidazole (Buffer B) to elute protein. The His-SUMO tag was cleaved by incubation with His-tagged protease Ulp1 at a 10:1 (w/w) ratio of ClpB:Ulp1 or DnaK:Ulp1 and dialysis in buffer A overnight. The protease and the cleaved His-SUMO tag were removed from target proteins by Ni-NTA affinity chromatography, and the flow-through was concentrated and subjected to a gel-filtration purification step (Superose increase 6) in buffer A. The fractions corresponding to the estimated elution peak were pooled for further use. The protein concentration of ClpB at this stage was about 8 mg/mL and the protein concentration of DnaK was around 4 mg/mL.

For the production of the N-terminal His-SUMO tagged DnaJ1, DnaJ2, and GrpE, *E. coli* BL21 cultures were grown with shaking at 37°C to OD_{600} of 0.8 in LB medium supplemented with 50 μg/mL ampicillin (Lupoli et al., 2016). Then cells were cooled to 16°C and induced with 0.2 mM IPTG for additional 16 h with shaking. Cells were harvested by centrifugation (3,100 × g, 10 min, 4°C) and the pellets were resuspended on ice with 20 mL of buffer C (25 mM Tris pH 8.0, 400 mM NaCl, 10 mM $MgCl_2$) supplemented with the EDTA-free protease inhibitor cocktail. For DnaJ1-expressing cells, buffer A (25 mM HEPES pH 7.5, 400 mM NaCl, 10 mM $MgCl_2$) was used with the same inhibitors added. Other purification steps are the same as the procedure used for ClpB and DnaK purification.

For the production of the N-terminal His-tagged Hsp20, the growth step was carried out as described (Lupoli et al., 2016). The purification followed the reported protocol except that at the dialysis step, the eluate samples were concentrated by using a 10-kDa MWCO Amicon Ultra Centrifugal Filter Device to 500 μL and subjected to gel filtration (Superdex 200 increase) in buffer C. The protein concentration of Hsp20 at this stage was about 4.2 mg/ml. All proteins were each flash frozen in liquid nitrogen and stored at −80°C.

Specimen preparation and cryo-EM data acquisition—We tried to purify the ClpB–DnaK complex by gel filtration but binding affinity was weak and did not survive chromatography. We then resorted to directly imaging the *in vitro* mixture. We reconstituted the ternary ClpB–DnaK–casein by preincubating 1 molar equivalent of the ClpB hexamer with 1.2 molar equivalent of κ-casein in 3 mM ATPγS on ice for 1 h, then added 1.2 molar equivalent of DnaK and incubated on ice for another 1 h. To examine the DnaJ2 binding, we mixed ClpB, casein, DnaK, and DnaJ2 at a molar ratio of 1:1.2:1.2:1.2 under the same conditions. Protein concentrations were determined by Bradford assay. The sample was diluted to 8 μM for cryo-EM grid preparation. A ClpB–DnaK–casein sample (3 μL) was deposited onto a freshly glow-discharged C-flat 2/2 300 mesh grid and blotted using a Thermo Fisher Scientific (TFS) Vitrobot Mark IV (FEI) with standard Vitrobot filter paper. Blotting time was set to 6 s, blotting force was set to 5, and the blotting was done under

100% humidity at 6°C. Grids were flash-frozen in liquid ethane and stored in liquid nitrogen.

The cryo-EM dataset was collected using a 300-kV TFS Titan Krios electron microscope operated at a nominal magnification of × 130,000 with defocus values from −1.0 to −2.0 μm on a Gatan K2 detector with a pixel size of 1.074 A per pixel at the sample level. Dose rate was two electrons per A^2 per second per frame, and 40 frames were recorded in a movie. We recorded two full datasets with 8753 and 5697 raw movie micrographs, respectively.

Data processing and 3D reconstruction—Movie micrographs were motion-corrected using the program MotionCorr-2 (Zheng et al., 2017). Contrast transfer function parameters of each aligned micrograph were calculated using CTFFIND-4.18 CTF (Mindell and Grigorieff, 2003). All remaining steps were performed using RELION-3 (Zivanov et al., 2018). Templates for automatic picking were generated from a 2D average of about 1500 manually picked particles. 2D classification was then performed and contamination and junk particles were removed following 2D classification analysis. For the two datasets, some 2,000,000 and 1,347,200 particles were selected and used for further 2D classification. Particles belong to selected 3D classes were combined, and a second round of 2D classification was performed to further exclude “bad” particles. We had 630,100 particles remaining for another round of 3D classification. A total of 417,884 particles (65%) were selected for final 3D classification using a fine angular sampling step of 1.8° and 12 local search steps, leading to three 3D classes that reached 3.3 Å, 3.5 A, and 3.1 Å overall resolution after refinement and postprocessing. These 3D maps were referred to as conf II, conf T, and conf I, respectively. Resolution of the map was estimated by the gold-standard Fourier shell correlation at a correlation cutoff value of 0.143 (Rosenthal and Henderson, 2003). Local 3D resolution was calculated using the program ResMap (Kucukelbir et al., 2014).

Separately, the main dataset (417,844 particles) was used to perform focused 3D classification, with a mask focusing either on the ClpB NTD region or around regions with the strong M-domain and DnaK density. Two masks were generated manually in UCSF Chimera (Pettersen et al., 2004), one for the NTD ring mask and the other for the ClpB M domain plus the DnaK. We treated the M-domain and DnaK densities as a single identity because our goal was to study their contact interface. Partial signal subtraction and 3D classification were performed without image alignment in Relion-3. The particles belonging to the classes with improved ClpB NTD or with improved ClpB M domain plus DnaK were selected for full-particle 3D refinement, leading to a 3D map with defined ClpB NTD density at 3.2 Å average resolution and a 3D map with defined densities for the M-domain and the DnaK NBD at 3.7 Å average resolution. However, the resolution of the density map of the interface between M-domain and DnaK-NBD is low (1–2 nm). To get better resolution of the interface, focused refinement of the three separately extracted M-domain-DnaK density maps led to three 3D maps at 7.4 Å, 7.9 Å and 6.8 Å resolution, respectively.

Model building—Atomic models of confs I, T, and II were built based on the published structure of the Mtb ClpB in complex with ATPγS and casein (PDB ID 6DJU). We used protomer P5 that is tightly bound to the polypeptide as the starting model (Yu et al., 2018).

For the protomers tightly bound to the polypeptide in the new 3D maps, the starting model with both NBD1 and NBD2 was docked into the map as a single rigidbody using Chimera. For the mobile protomers that were loosely associated with the substrate polypeptide, NBD1 (residues 159 – 550) and NBD2 (551–848) were separated and docked as individual rigidbodies. For those protomers with good M-domain densities, the atomic model of the M-domain motif 1 model was rigid-body docked into the corresponding EM densities. The docked models were then manually adjusted and rebuilt in the Coot program (Emsley et al., 2010). The resulting models were subjected to real-space refinement using the phenix.real_space_refine in the PHENIX program (Adams et al., 2010). Some protomers had clear helical densities for the ClpB M-domain motif 2. We built atomic model for the M-domain motif 2 starting from rigidbody fitting with the *E. coli* ClpB M-domain (PDB ID 4CIU) followed by manual adjustment using Coot and finally by all-atom refinement in the PHENIX. The quality of the final models was assessed using the MolProbity program (Duke University) (Chen et al., 2010).

For the ClpB NTD region, after focused 3D refinement, density map of the ClpB NTD on protomer P3 reached about 5 Å. This allowed for the main chain placement for most regions of the domain. The initial model was built using the published structure of the Hsp104 NTD from *Calcarisporiella thermophila* (PDB ID 6AZY) as a template. The template was identified using the HHpred program (Söding et al., 2005). Fitting of the template and the density map was done in the Rosetta program (Wang et al., 2016). The structure was manually adjusted in the Coot and all-atom refined by PHENIX. The validation and statistics of the model of ClpB conf T with NTD are listed in Table S2.

For the M domain–DnaK region, after signal subtraction, focused 3D classification, and local refinement, the resolution of the three 3D maps of the M-domain–DnaK reached around 7 Å, which is insufficient for modeling, because the crystal structure of the Mtb DnaK NBD was available (PDB ID 4RTF). We performed flexible docking of the crystal structure into the best local map (P6) with the Rosetta program. After aligning the DnaK NBD structure to the density map, a total of 200 structure models were generated using the density map as guidance. During the process, both the interface and the individual proteins were allowed to sample different conformations. The structure with the best Rosetta score was selected as the structure model of the complex. The structural model was then combined with the structural model of ClpB conf T, the modeling and validation statistics of ClpB conf T with the M domain and DnaK are listed in Table S2. The structural model of the DnaK–M domain was also fitted into the density maps of the other two branches (P4, P5) and combined with the structural model of ClpB conf T. The resulting ClpB-three DnaKs structure model was then used to guide the combination of ClpB conf T map with the three locally-refined DnaK–M domain maps by using the “Combine-focused-maps” module in the Phenix program (Adams et al., 2010). The modeling and validation statistics of ClpB conf T with three DnaKs are also listed in Table S2.

All structural figures were prepared using PyMOL (Schrödinger, LLC.) and Chimera.

Protein reactivation assay using denatured firefly luciferase—We followed the protocol as described previously (Lupoli et al., 2016; Mogk et al., 2003; Nillegoda et al.,

2015; Yu et al., 2018). Briefly, luciferase (100 nM) was heated with Hsp20 (400 nM) in 5 μL aliquots for 5 min at 42°C in 50 mM Tris pH7.5, 150 mM KCl, 20 mM $MgCl_2$, and 2 mM DTT (buffer E) in nonstick tubes then cooled on ice for 5 min. For protein refolding reactions, we added chaperones and/or cofactors (ClpB 2 μM; DnaK 6 μM; and DnaJ1, DnaJ2, and GrpE at 2 μM each) into buffer E plus 1 mg/mL BSA, reactions were initiated by adding 2 mM ATP to a final reaction volume of 20 μL and the tubes incubated at 25°C for 30 min. Luminescence was measured by placing 2 μL samples into OptiPlate™ 96-well plates and adding 100 μL of luciferase reagent. Light emission was measured using EnVision plate reader. Native luciferase activity was measured using luciferase + Hsp20 prepared in the same manner without the denaturation step. Reactions with or without ClpB and mutants were done in triplicate in two independent experiments. The percentage of luciferase activity was determined as the percentage of the reaction luminescence after subtracting background signal relative to native luminescence for each, and then normalized relative to reactions containing wild-type ClpB that was measured in the same experiment.

ATPase assay—ATPase activity of DnaK and its mutants and ClpB with or without casein were determined by measuring the production of inorganic phosphate using the Malachite Green Phosphate Assay Kit. Measurement of the DnaK ATPase activity was started by adding 0.25 mM ATP to solution containing 8 μg WT or mutant DnaK proteins, 25 mM Tris-Cl pH8, 40 mM NaCl, 10 mM $MgCl_2$. ATPase activity of ClpB with or without casein was determined by adding 0.25 mM ATP and 20 μM casein to solution containing 1.2 μM WT or mutated ClpB (25 mM Tris, pH 8.0, 40 mM NaCl, 10 mM $MgCl_2$). Reactions were performed for 30 min at 37°C. 80 μL reaction products were collected at 0 min and 30 min, then added to 20 μL Malachite Green reagent in a 96-well plate. The amount of inorganic phosphate released was measured by the absorbance change at wavelength 620 nm.

Isothermal titration calorimetry—Binding measurements were performed on a MicroCal PEAQ ITC at 25°C. Concentrated N-terminal 6xHis-tag ClpB M-domain motif II, DnaK NBD, and their corresponding mutant proteins were subjected to a gel-filtration purification step in buffer A before the binding assay. Purity of the target proteins was confirmed by SDS-PAGE. The WT DnaK NBD or mutant proteins (3426 μM) were first loaded into a syringe, and WT or mutant ClpB M-domain motif 2 (250 μM) were then added into the cell. Nineteen injections (2 μL each, separated by 150 s) were performed after an initial injection of 0.4 μL.

Heat sensitivity assay—Cultures were grown to an OD_{600} of 0.2–0.3 in 7H9-albumin-dextrose-NaCl (ADS) with 0.05% Tween 80 overnight at 37°C and diluted to an OD_{600} of 0.1 in fresh medium. Fifty-microliters of each were incubated at 53°C or 37°C for 1.5 h. Serial dilutions of each were then spotted on 7H10 smeg (0.5% glycerol and 0.5% dextrose added) and incubated at 37°C to calculate CFU.

Immunoblots—For protein detection the following antibodies where used: Anti-DnaK (1:20,000), anti-ClpB (1:10,000 (Vaubourgeix et al., 2015)), and anti-SigA. Peptide-specific anti-SigA was affinity purified by using the respective peptide and SulfoLink

Immobilization Kit for Peptides. For immunoblotting, antibodies were used with anti-Rabbit IgG-HRP at 1:10,000.

***In vivo* cell image collection—**All images were acquired using a Zeiss Axio Observer Z1 microscope equipped with Definite focus, Stage top incubator (Insert P Lab-Tek S1, TempModule S1), Colibri.2 and Illuminator HXP 120 C light sources, a Hamamatsu ORCA-Flash4.0 CMOS Camera and a Plan-Apochromat 100 × /1.4 oil DIC objective. Zeiss Zen software was used for acquisition and image export. The following filter sets and light sources were used for imaging: for mCitrine: YFP (46 HE, Colibri2.0 505 LED), and for mCerulean: CFP (47 HE, HXP 120 C). For all images were cells added to a 1.5% low melting point agarose 7H9smeg pad prepared in a 17 × 28 mm geneframe (Thermo Scientific, AB-0578) (Fay and Glickman, 2014). For co-localization of mCerulean-AL09 and mCitrine-tagged ClpB or DnaK proteins, aggregates were induced with the addition of 50 ng/ml ATc (anhydrotetracycline) for 5 h prior to imaging. The plasmid pET12a-AL-09 FL used in this study was a gift from Marina Ramirez-Alvarado (Addgene plasmid # 47081).

QUANTIFICATION AND STATISTICAL ANALYSIS

All data shown in Figures 1, 2, 3, 4, 5, S4, and S6 represent the average ± SEM from three replicates in two independent experiments. ITC data were analyzed by the software that came with the MicroCal PEAQ-ITC equipment. Statistical analyses were performed in GraphPad Prism software version 8. Quantification, statistical analysis, and validation are implemented in the software packages used for 3D reconstruction and model refinement.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

Cryo-EM data were collected at the David Van Andel Advanced Cryo-Electron Microscopy Suite at the Van Andel Research Institute. We thank Gongpu Zhao and Xing Meng for help with data collection and David Nadziejka for technical editing of the manuscript. This study was supported by the US National Institutes of Health grants R01 AI070285 (to H.L.) and U19 AI111143 and P30 CA008748 (to M.S.G.).

REFERENCES

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr 66, 213–221. [PubMed: 20124702]

Avellaneda MJ, Franke KB, Sunderlikova V, Bukau B, Mogk A, and Tans SJ (2020). Processive extrusion of polypeptide loops by a Hsp100 disaggregase. Nature 578, 317–320. [PubMed: 31996849]

Barnett ME, Nagy M, Kedzierska S, and Zolkiewski M (2005). The amino-terminal domain of ClpB supports binding to strongly aggregated proteins. J. Biol. Chem 280, 34940–34945. [PubMed: 16076845]

Bertelsen EB, Chang L, Gestwicki JE, and Zuiderweg ERP (2009). Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proc. Natl. Acad. Sci. USA 106, 8471–8476. [PubMed: 19439666]

Bracher A, and Verghese J (2015). The nucleotide exchange factors of Hsp70 molecular chaperones. Front. Mol. Biosci 2, 10. [PubMed: 26913285]

Carroni M, Kummer E, Oguchi Y, Wendler P, Clare DK, Sinning I, Kopp J, Mogk A, Bukau B, and Saibil HR (2014). Head-to-tail interactions of the coiled-coil domains regulate ClpB activity and cooperation with Hsp70 in protein disaggregation. eLife 3, e02481. [PubMed: 24843029]

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, and Richardson DC (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr 66, 12–21. [PubMed: 20057044]

Chow IT, and Baneyx F (2005). Coordinated synthesis of the two ClpB isoforms improves the ability of Escherichia coli to survive thermal stress. FEBS Lett. 579, 4235–4241. [PubMed: 16038902]

Chow IT, Barnett ME, Zolkiewski M, and Baneyx F (2005). The N-terminal domain of Escherichia coli ClpB enhances chaperone function. FEBS Lett. 579, 4242–4248. [PubMed: 16051221]

Clarke AK, and Eriksson MJ (2000). The truncated form of the bacterial heat shock protein ClpB/HSP100 contributes to development of thermotolerance in the cyanobacterium Synechococcus sp. strain PCC 7942. J. Bacteriol 182, 7092–7096. [PubMed: 11092876]

Desantis ME, and Shorter J (2012). The elusive middle domain of Hsp104 and ClpB: location and function. Biochim. Biophys. Acta 1823, 29–39. [PubMed: 21843558]

Desantis ME, Sweeny EA, Snead D, Leung EH, Go MS, Gupta K, Wendler P, and Shorter J (2014). Conserved distal loop residues in the Hsp104 and ClpB middle domain contact nucleotide-binding domain 2 and enable Hsp70-dependent protein disaggregation. J. Biol. Chem 289, 848–867. [PubMed: 24280225]

Deville C, Carroni M, Franke KB, Topf M, Bukau B, Mogk A, and Saibil HR (2017). Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. Sci. Adv 3, e1701726. [PubMed: 28798962]

Deville C, Franke K, Mogk A, Bukau B, and Saibil HR (2019). Two-Step Activation Mechanism of the ClpB Disaggregase for Sequential Substrate Threading by the Main ATPase Motor. Cell Rep. 27, 3433–3446. [PubMed: 31216466]

Doyle SM, and Wickner S (2009). Hsp104 and ClpB: protein disaggregating machines. Trends Biochem. Sci 34, 40–48. [PubMed: 19008106]

Doyle SM, Shastry S, Kravats AN, Shih YH, Miot M, Hoskins JR, Stan G, and Wickner S (2015). Interplay between E. coli DnaK, ClpB and GrpE during protein disaggregation. J. Mol. Biol 427, 312–327. [PubMed: 25451597]

Durie CL, Duran EC, and Lucius AL (2018). Escherichia coli DnaK Allosterically Modulates ClpB between High- and Low-Peptide Affinity States. Biochemistry 57, 3665–3675. [PubMed: 29812913]

Emsley P, Lohkamp B, Scott WG, and Cowtan K (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr 66, 486–501. [PubMed: 20383002]

Fay A, and Glickman MS (2014). An essential nonredundant role for mycobacterial DnaK in native protein folding. PLoS Genet. 10, e1004516. [PubMed: 25058675]

Fernández-Higuero JA, Aguado A, Perales-Calvo J, Moro F, and Muga A (2018). Activation of the DnaK-ClpB Complex is Regulated by the Properties of the Bound Substrate. Sci. Rep 8, 5796. [PubMed: 29643454]

Flaherty KM, DeLuca-Flaherty C, and McKay DB (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. Nature 346, 623–628. [PubMed: 2143562]

Gates SN, Yokom AL, Lin J, Jackrel ME, Rizo AN, Kendsersky NM, Buell CE, Sweeny EA, Mack KL, Chuang E, et al. (2017). Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. Science 357, 273–279. [PubMed: 28619716]

Goloubinoff P, Mogk A, Zvi AP, Tomoyasu T, and Bukau B (1999). Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. Proc. Natl. Acad. Sci. USA 96, 13732–13737. [PubMed: 10570141]

Harnagel A, Lopez Quezada L, Park SW, Baranowski C, Kieser K, Jiang X, Roberts J, Vaubourgeix J, Yang A, Nelson B, et al. (2021). Nonredundant functions of Mycobacterium tuberculosis chaperones promote survival under stress. Mol Microbiol. 115, 272–289. [PubMed: 32996193]

Heuck A, Schitter-Sollner S, Suskiewicz MJ, Kurzbauer R, Kley J, Schleiffer A, Rombaut P, Herzog F, and Clausen T (2016). Structural basis for the disaggregase activity and regulation of Hsp104. eLife 5, e21516. [PubMed: 27901467]

Kampinga HH, and Craig EA (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat. Rev. Mol. Cell Biol 11, 579–592. [PubMed: 20651708]

Kannan TR, Musatovova O, Gowda P, and Baseman JB (2008). Characterization of a unique ClpB protein of Mycoplasma pneumoniae and its impact on growth. Infect. Immun 76, 5082–5092. [PubMed: 18779336]

Kityk R, Kopp J, Sinning I, and Mayer MP (2012). Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones. Mol. Cell 48, 863–874. [PubMed: 23123194]

Kityk R, Kopp J, and Mayer MP (2018). Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. Mol. Cell 69, 227–237. [PubMed: 29290615]

Kucukelbir A, Sigworth FJ, and Tagare HD (2014). Quantifying the local resolution of cryo-EM density maps. Nat. Methods 11, 63–65. [PubMed: 24213166]

Lee S, Sowa ME, Watanabe YH, Sigler PB, Chiu W, Yoshida M, and Tsai FTF (2003). The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell 115, 229–240. [PubMed: 14567920]

Lee S, Sowa ME, Choi JM, and Tsai FT (2004). The ClpB/Hsp104 molecular chaperone-a protein disaggregating machine. J. Struct. Biol 146, 99–105. [PubMed: 15037241]

Lee S, Roh SH, Lee J, Sung N, Liu J, and Tsai FTF (2019). Cryo-EM Structures of the Hsp104 Protein Disaggregase Captured in the ATP Conformation. Cell Rep. 26, 29–36. [PubMed: 30605683]

Lupoli TJ, Fay A, Adura C, Glickman MS, and Nathan CF (2016). Reconstitution of a Mycobacterium tuberculosis proteostasis network highlights essential cofactor interactions with chaperone DnaK. Proc. Natl. Acad. Sci. USA 113, E7947–E7956. [PubMed: 27872278]

Maurizi MR, and Xia D (2004). Protein binding and disruption by Clp/Hsp100 chaperones. Structure 12, 175–183. [PubMed: 14962378]

Mayer MP, and Gierasch LM (2019). Recent advances in the structural and mechanistic aspects of Hsp70 molecular chaperones. J. Biol. Chem 294, 2085–2097. [PubMed: 30455352]

Michalska K, Zhang K, March ZM, Hatzos-Skintges C, Pintilie G, Bigelow L, Castellano LM, Miles LJ, Jackrel ME, Chuang E, et al. (2019). Structure of Calcarisporiella thermophila Hsp104 Disaggregase that Antagonizes Diverse Proteotoxic Misfolding Events. Structure 27, 449–463. [PubMed: 30595457]

Mindell JA, and Grigorieff N (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol 142, 334–347. [PubMed: 12781660]

Mogk A, Deuerling E, Vorderwülbecke S, Vierling E, and Bukau B (2003). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. Mol. Microbiol 50, 585–595. [PubMed: 14617181]

Mogk A, Kummer E, and Bukau B (2015). Cooperation of Hsp70 and Hsp100 chaperone machines in protein disaggregation. Front. Mol. Biosci 2, 22. [PubMed: 26042222]

Mogk A, Bukau B, and Kampinga HH (2018). Cellular Handling of Protein Aggregates by Disaggregation Machines. Mol. Cell 69, 214–226. [PubMed: 29351843]

Nagy M, Guenther I, Akoyev V, Barnett ME, Zavodszky MI, Kedzierska-Mieszkowska S, and Zolkiewski M (2010). Synergistic cooperation between two ClpB isoforms in aggregate reactivation. J. Mol. Biol 396, 697–707. [PubMed: 19961856]

Nillegoda NB, Kirstein J, Szlachcic A, Berynskyy M, Stank A, Stengel F, Arnsburg K, Gao X, Scior A, Aebersold R, et al. (2015). Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. Nature 524, 247–251. [PubMed: 26245380]

Oguchi Y, Kummer E, Seyffer F, Berynskyy M, Anstett B, Zahn R, Wade RC, Mogk A, and Bukau B (2012). A tightly regulated molecular toggle controls AAA+ disaggregase. Nat. Struct. Mol. Biol 19, 1338–1346. [PubMed: 23160353]

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem 25, 1605–1612. [PubMed: 15264254]

Rizo AN, Lin J, Gates SN, Tse E, Bart SM, Castellano LM, DiMaio F, Shorter J, and Southworth DR (2019). Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase. Nat. Commun 10, 2393. [PubMed: 31160557]

Rosenthal PB, and Henderson R (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol 333, 721–745. [PubMed: 14568533]

Rosenzweig R, Moradi S, Zarrine-Afsar A, Glover JR, and Kay LE (2013). Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. Science 339, 1080–1083. [PubMed: 23393091]

Rosenzweig R, Farber P, Velyvis A, Rennella E, Latham MP, and Kay LE (2015). ClpB N-terminal domain plays a regulatory role in protein disaggregation. Proc. Natl. Acad. Sci. USA 112, E6872–E6881. [PubMed: 26621746]

Rosenzweig R, Nillegoda NB, Mayer MP, and Bukau B (2019). The Hsp70 chaperone network. Nat. Rev. Mol. Cell Biol 20, 665–680. [PubMed: 31253954]

Saibil H (2013). Chaperone machines for protein folding, unfolding and disaggregation. Nat. Rev. Mol. Cell Biol 14, 630–642. [PubMed: 24026055]

Schlieker C, Tews I, Bukau B, and Mogk A (2004). Solubilization of aggregated proteins by ClpB/DnaK relies on the continuous extraction of unfolded polypeptides. FEBS Lett. 578, 351–356. [PubMed: 15589844]

Seyffer F, Kummer E, Oguchi Y, Winkler J, Kumar M, Zahn R, Sourjik V, Bukau B, and Mogk A (2012). Hsp70 proteins bind Hsp100 regulatory M domains to activate AAA+ disaggregase at aggregate surfaces. Nat. Struct. Mol. Biol 19, 1347–1355. [PubMed: 23160352]

Shorter J, and Southworth DR (2019). Spiraling in Control: Structures and Mechanisms of the Hsp104 Disaggregase. Cold Spring Harb. Perspect. Biol 11, a034033. [PubMed: 30745294]

Söding J, Biegert A, and Lupas AN (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244–W248. [PubMed: 15980461]

Swain JF, Dinler G, Sivendran R, Montgomery DL, Stotz M, and Gierasch LM (2007). Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. Mol. Cell 26, 27–39. [PubMed: 17434124]

Tripathi P, Parijat P, Patel VK, and Batra JK (2018). The amino-terminal domain of *Mycobacterium tuberculosis* ClpB protein plays a crucial role in its substrate disaggregation activity. FEBS Open Bio 8, 1669–1690.

Vaubourgeix J, Lin G, Dhar N, Chenouard N, Jiang X, Botella H, Lupoli T, Mariani O, Yang G, Ouerfelli O, et al. (2015). Stressed mycobacteria use the chaperone ClpB to sequester irreversibly oxidized proteins asymmetrically within and between cells. Cell Host Microbe 17, 178–190. [PubMed: 25620549]

Wang RYR, Song Y, Barad BA, Cheng Y, Fraser JS, and DiMaio F (2016). Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. eLife 5, e17219. [PubMed: 27669148]

Winkler J, Tyedmers J, Bukau B, and Mogk A (2012). Hsp70 targets Hsp100 chaperones to substrates for protein disaggregation and prion fragmentation. J. Cell Biol 198, 387–404. [PubMed: 22869599]

Yang J, Zong Y, Su J, Li H, Zhu H, Columbus L, Zhou L, and Liu Q (2017). Conformation transitions of the polypeptide-binding pocket support an active substrate release from Hsp70s. Nat. Commun 8, 1201. [PubMed: 29084938]

Yokom AL, Gates SN, Jackrel ME, Mack KL, Su M, Shorter J, and Southworth DR (2016). Spiral architecture of the Hsp104 disaggregase reveals the basis for polypeptide translocation. Nat. Struct. Mol. Biol 23, 830–837. [PubMed: 27478928]

Yu H, Lupoli TJ, Kovach A, Meng X, Zhao G, Nathan CF, and Li H (2018). ATP hydrolysis-coupled peptide translocation mechanism of *Mycobacterium tuberculosis* ClpB. Proc. Natl. Acad. Sci. USA 115, E9560–E9569. [PubMed: 30257943]

Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, and Agard DA (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332. [PubMed: 28250466]

Zhu X, Zhao X, Burkholder WF, Gragerov A, Ogata CM, Gottesman ME, and Hendrickson WA (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. Science 272, 1606–1614. [PubMed: 8658133]

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJH, Lindahl E, and Scheres SHW (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife 7, e42166. [PubMed: 30412051]

Zolkiewski M (1999). ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from Escherichia coli. J. Biol. Chem 274, 28083–28086. [PubMed: 10497158]

Highlights

- Cryo-EM structural analyses of the *M. tuberculosis* ClpB-DnaK bi-chaperone system
- Mtb ClpB adopts three conformations in the presence of DnaK
- The ClpB middle domain does not bend or tilt when bound to DnaK
- The ClpB middle domain may interact with only the IIB subdomain of DnaK NBD

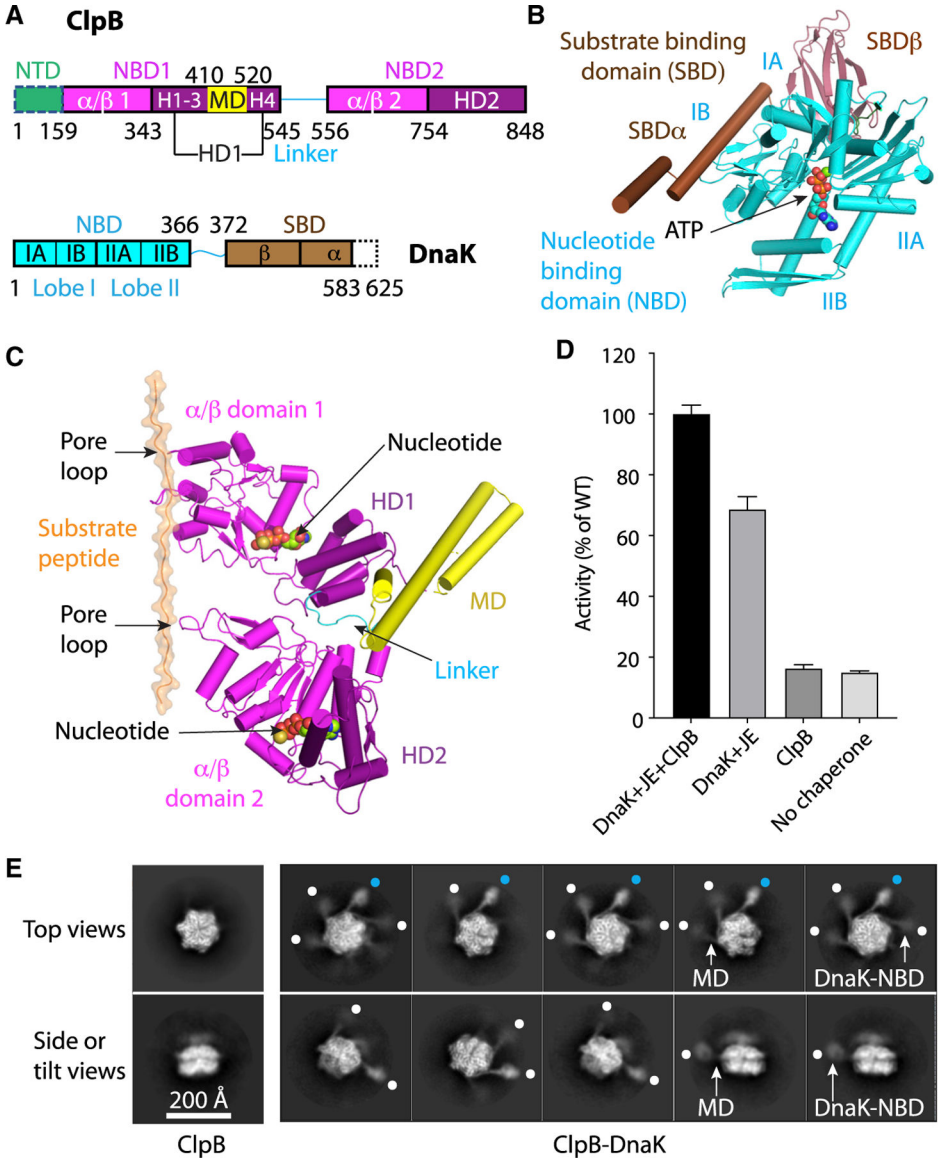


Figure 1. Assembly of the Mtb ClpB-DnaK complex *in vitro*

(A) The domain architecture of ClpB (upper) and DnaK.

(B) Structure of *E. coli* DnaK (PDB: 4B9Q).

(C) Structure of ClpB protomer with full M domain determined in this study.

(D) Denatured luciferase reactivation assay. J = DnaJ1 and DnaJ2 and E = GrpE. ClpB and DnaK synergistically refold luciferase. Reactions were done in triplicate in two independent experiments. Data are represented as mean ± SEM.

(E) Representative 2D class averages of ClpB alone (left) and DnaK-bound ClpB (right) in the presence of both casein and ATPγS. A dot is adjacent to each observed DnaK density. In the top views, cyan dots represent the strongest DnaK density in each class average.

See also Figure S1.

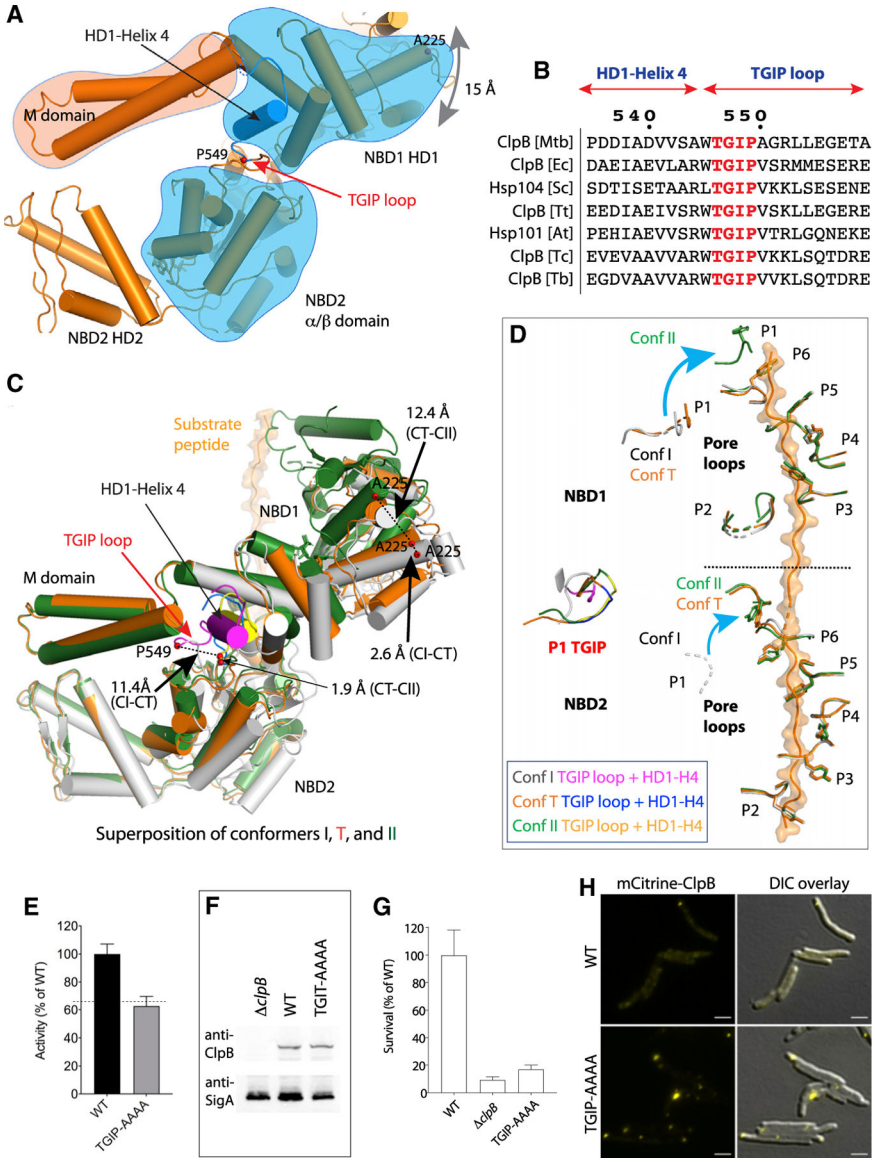


Figure 2. The TGIP loop enables conformational changes among conf I, conf T, and conf II

(A) Cartoon view of a ClpB protomer in conf T. The last helix (Helix 4) of helical domain HD1 of NDB1 (HD1-Helix 4) connects M domain and the TGIP loop.

(B) Sequence alignment of ClpB homologs around the TGIP region. Mtb, *M. tuberculosis*; Ec, *E. coli*; Sc, *Saccharomyces cerevisiae*; Tt, *Thermus thermophilus*; At, *Arabidopsis thaliana*; Tc, *Trypanosoma cruzi*; Tb, *Trypanosoma brucei*.

(C) Comparison of the three P1 protomers in conf I (gray), conf T (orange), and conf II (dark green) aligned with their respective protomer P6. The positions of conf T are always in between those of conf I and conf II. Residues A225 and P549 are shown as red spheres in each conformation to better visualize the positional changes.

(D) A comparison of the pore loops in conf I, conf T, and conf II; the conformations are aligned by their respective protomer P6. Dashed curves mark disordered loops. The black dotted line separates NBD1 and NBD2.

(E) Luciferase aggregate reactivation assay using WT and TGIP mutant ClpB. The dotted line at 63% defines the activity in the absence of ClpB (i.e., the background activity of DnaK and cofactors). Data are represented as mean ± SEM.

(F) Expression levels of the WT and mutant ClpB proteins as probed by western blot.

(G) Survival of the MGM6971 (ClpB-mCitrine WT), MGM7023 (Δ*clpB*), and MGM6988 ($ClpB_{TGIP\text{-}AAAA}$-mCitrine) strains after exposure to 53°C for 1.5 h, normalized to the wild-type (MGM6971). All experiments were conducted in triplicate in two independent experiments; data are represented as mean ± SEM.

(H) Localization of the WT (MGM6971) and mutant $ClpB_{TGIP\text{-}AAAA}$ (MGM6988) in *M. smegmatis*.

See also Figures S2–S4 and Table S1.

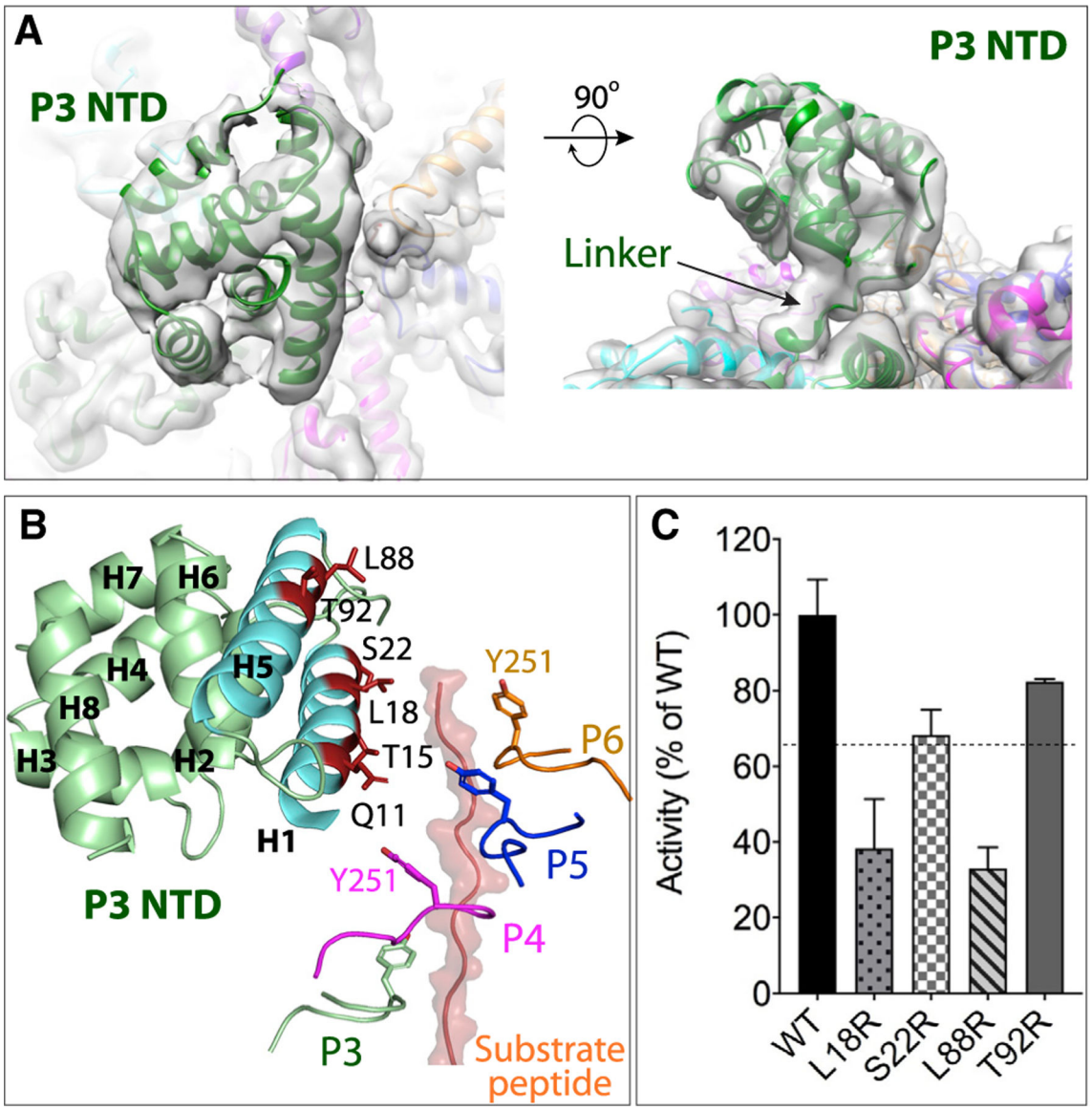


Figure 3. NTD of one ClpB protomer in conf T becomes ordered in the presence of DnaK and substrate casein

(A) Top and side views of the NTD density of protomer P3 docked with the atomic model. The arrow points to the density connecting NTD to the main body of P3 protomer.

(B) Cartoon view of the atomic model showing the substrate entry region and highlighting the P3 NTD and peptide-translocating pore loops. Highlighted in red are ClpB NTD residues that surround and likely interact with the substrate: Gln-11, Thr-15, Leu-18, and Ser-22 in helix H1, and Leu-88 and Thr-92 in helix H5.

(C) Luciferase aggregate reactivation assay using purified ClpB with mutated residues surrounding the substrate entry. Dotted line at 63% defines the activity of reactions lacking ClpB (i.e., background refolding activity of DnaK and cofactors). Experiments were conducted in triplicate, and data are represented as mean ± SEM.

See also Figure S5 and Table S2.

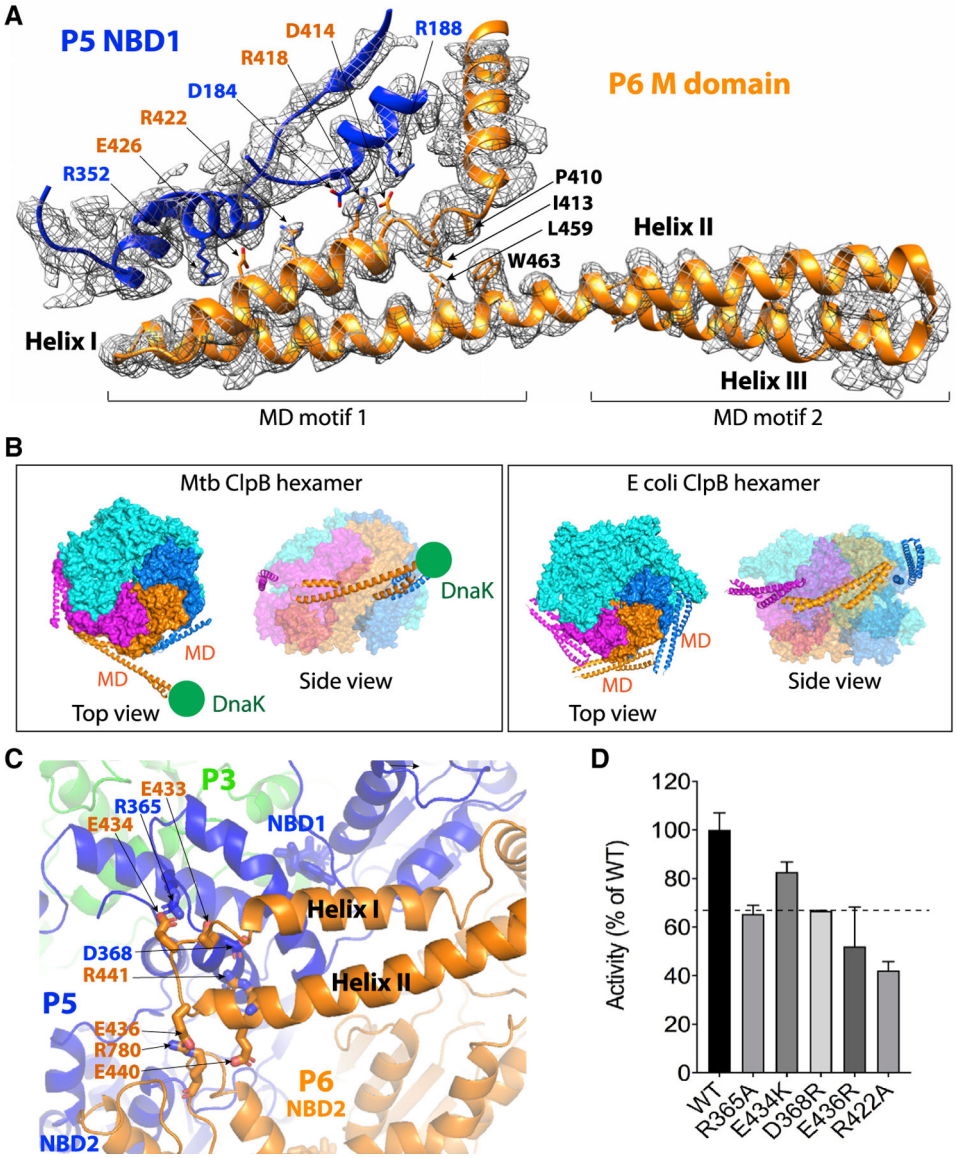


Figure 4. The ClpB M-domain motif 1 mediates multiple charge-charge interactions within and between protomers in the T conformation

(A) Density and the atomic model of the P6 M domain (orange) and NBD1 of neighboring P5 (blue), viewed from the top of ClpB hexamer. Salt bridges between P6 M-domain motif 1 helix I (Glu-426, Arg-422, Arg-418, and Asp-414) and P5 NBD1 (Arg-352, Asp-184, and Arg-188) are shown as sticks. In M-domain motif 1, helix I and II pack hydrophobically with each other, involving Pro-410, Ile-413, Leu-459, and Tyr-463 shown as sticks.

(B) Conformation of the M domain of Mtb ClpB bound to DnaK (left) compared to M domain of *E. coil* ClpB (right, DWB mutant, PDB: 6RN3). Note that only the M domain of Mtb ClpB bound to DnaK is fully stabilized such that both motif 1 and motif 2 are resolved.

(C) P6 M-domain motif 1 forms salt bridges with P5 NBD1 and P6 NBD2 viewed from the side of the ClpB hexamer focusing on the junction region between M-domain helices I-II and between NBD1-NBD2.

(D) Mutations of the five residues involved in charge-charge interactions with NBD1 and NBD2 affect ClpB activity. Dashed line at 63% defines yield of reactions lacking ClpB (i.e., background refolding activity of DnaK and cofactors alone). Experiments were conducted in triplicate and mean ± SEM are plotted.
See also Figures S4 and S5 and Table S2.

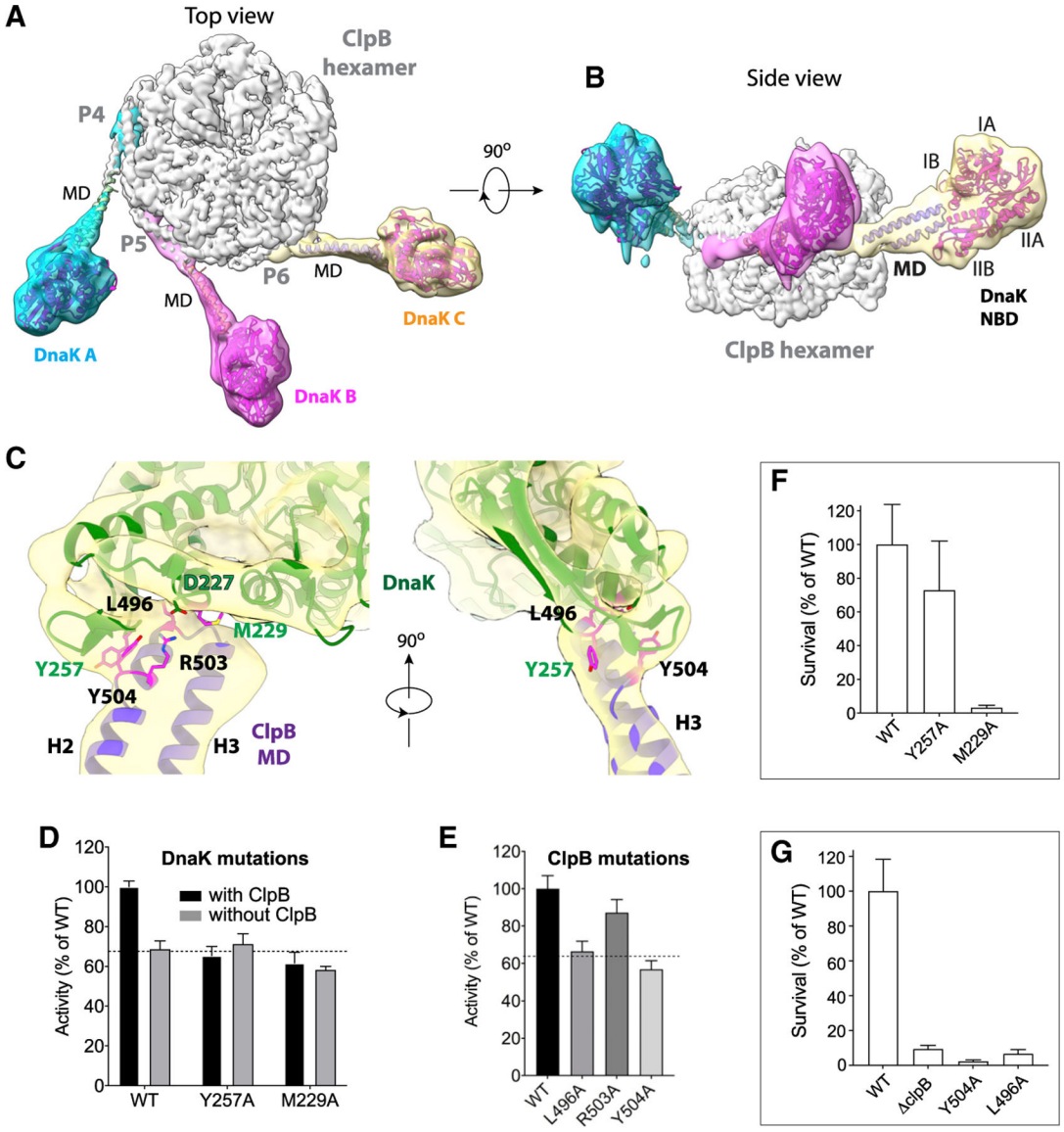


Figure 5. Cryo-EM map of ClpB-DnaK and their interface

(A and B) Top and side views of the combined 3D maps of ClpB and the three DnaK NBDs, derived individually by masking and local refinement. Structural models of the three DnaK NBDs are fitted in the density maps and shown in cartoons. DnaK NBD subdomains IA, IIA, IB, and IIB are labeled in (B).

(C) Enlarged view of the interface between ClpB M domain and DnaK NBD fitted in the 3D density map after local refinement. Interfacial residues are labeled. H1, H2, and H3 refer to helix 1, 2, and 3 of the ClpB M domains. Positions that are tested in mutational assay are color magenta.

(D and E) Mutational analysis of the key residues between DnaK and ClpB using a luciferase-based aggregate reactivation assay. Dashed line at 65% defines the activity in reactions lacking ClpB. Experiments were conducted in triplicate and mean ± SEM plotted.

(F) Percentage of survival of the MGM6003 (DnaK-mCitrine WT), MGM6972 ($DnaK_{Y257A}$-mCitrine), and MGM6973 ($DnaK_{M229A}$-mCitrine) strains after exposure to 53°C for 1.5 h, normalized to wild-type (MGM6003). Endogenous *dnaK* was deleted.
(G) Percentage of survival of MGM6971 (ClpB-mCitrine WT), MGM7023 (Δ*clpB*), MGM6989 ($ClpB_{Y504A}$-mCitrine), and MGM6991 ($ClpB_{L496A}$-mCitrine) strains after exposure to 53°C for 1.5 h, normalized to the wild-type (MGM6971). Endogenous *clpB* was deleted. All experiments were conducted in triplicate and mean ± SEM plotted.
See also Figures S6–S9 and Table S2.

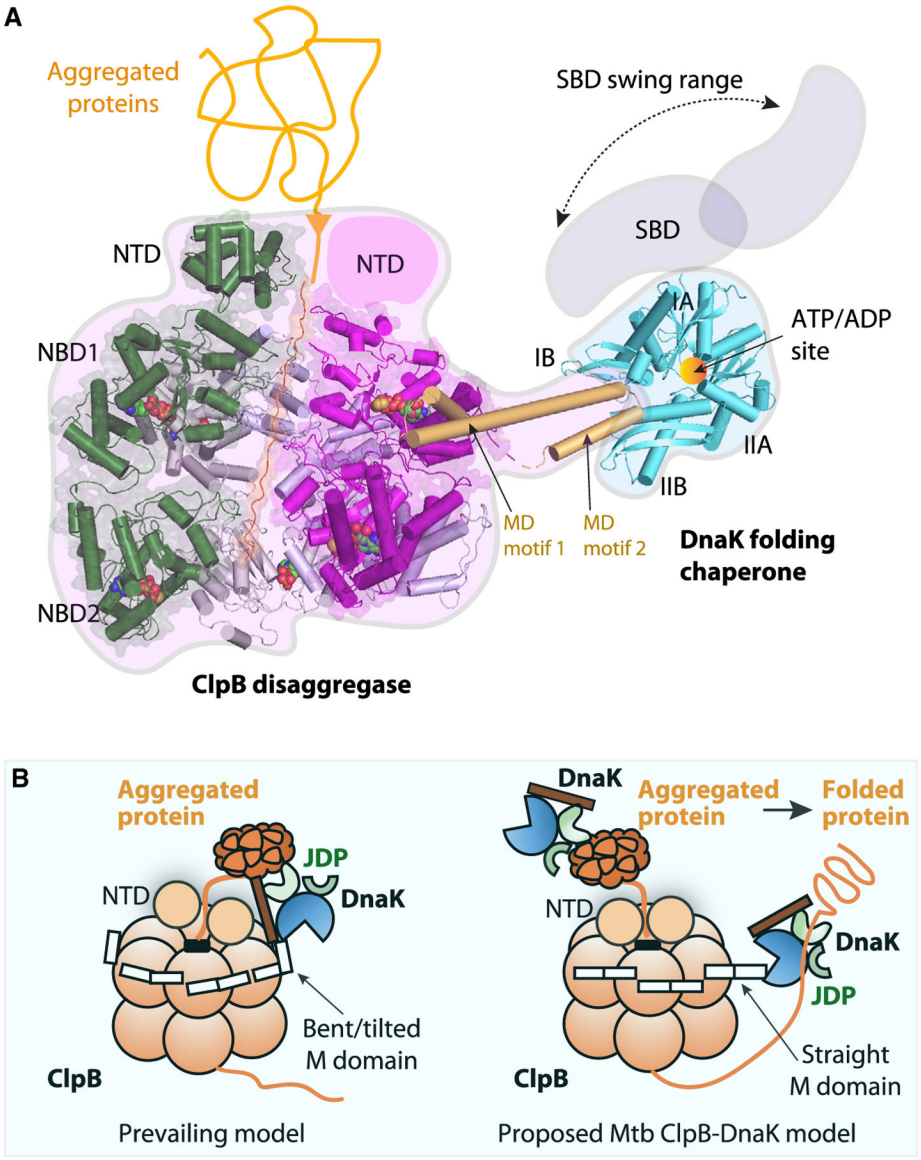


Figure 6. Coordination between Mtb ClpB and DnaK

(A) A model of the Mtb ClpB-DnaK bi-chaperone system. Range of movement of the substrate binding domain (SBD) of DnaK is sketched in gray, and an aggregate substrate is sketched in orange. (B) Comparison of the current model (left) and the revised model (right) of the ClpB-DnaK system. JDP, J-domain protein. The current concept is that DnaK binding to ClpB tilts the M domain, resulting in activation of the disaggregase and positioning of DnaK to facilitate substrate binding together with the NTD of ClpB. In the model proposed for the Mtb system, M domain does not bend and motif 2 does not tilt upward, so DnaK is located too far to aid substrate entry while bound to the M domain. Instead, DnaK is well positioned to capture unfolded polypeptide from ClpB and facilitates substrate folding. Aggregates in Mtb can be targeted to ClpB either by ClpB NTD or by DnaK that is not bound to ClpB M domain.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-DnaK	rabbit antisera (Harnagel et al., 2021	
anti-ClpB	rabbit antisera (Vaubourgeix et al., 2015)	
anti-SigA	affinity purified antisera, EZBiolab	
anti-Rabbit IgG-HRP	ThermoScientific	Cat#656120
Bacterial and virus strains		
E. coli BL21 (DE3)	Sigma-Aldrich	Cat#69450-3
Chemicals, peptides, and recombinant proteins		
ATP	Roche	Cat#10127531001; CAS: 51963-61-2
Malachite green	Sigma-Aldrich	Cat#M9636; CAS: 123333-61-9
Ni-NTA agarose	QIAGEN	Cat#30230
Superdex 200 10/300 GL column	GE Healthcare	Cat#17108801
Superose increase 6	GE Healthcare	GE17-5172-01
κ-casein	Sigma-Aldrich	Cat#C0406; CAS: 9000-71-9
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Sigma-Aldrich	Cat#I6758; CAS:367-93-1
EDTA-free protease inhibitor cocktail tablets	Roche	04693132001 - cOmplete, EDTA-free Protease Inhibitor Cocktail
Luciferase reagent	Promega	Cat#E1500
Deposited data		
ClpB conf I	This manuscript	PDB: 6W6G; EMD-21554
ClpB conf T	This manuscript	PDB: 6W6I; EMD-21556
ClpB conf II	This manuscript	PDB: 6W6H; EMD-21555
ClpB conf T with NTD	This manuscript	PDB: 6W6J; EMD-21557
ClpB conf T with one DnaK NBD	This manuscript	PDB: 6W6E; EMD-21553
ClpB conf T with three DnaK NBDs	This manuscript	PDB: 7L6N; EMD-23206
Experimental models: Organisms/strains		
MGM6003	Fay and Glickman, 2014	Δ*dnaK* attB::*dnaK*-mCitrine kan
MGM6971	Δ*clpB* (Vaubourgeix et al., 2015) transformed with pAJF1081	Δ*clpB* attB::Ms *clpB*-mCitrine kan
MGM6972	MGM6002 (Fay and Glickman 2014) transformed with pAJF1087 and verified as StrepS	Δ*dnaK* attB::*dnaK*(Y257A)-mCitrine kan
MGM6973	MGM6002 transformed with pAJF1088 and verified as StrepS	Δ*dnaK* attB::*dnaK*(M229A)-mCitrine kan
MGM6988	Δ*clpB* (Vaubourgeix et al., 2015) transformed with pAJF1092	Δ*clpB* attB::Ms *clpB* (TGIP-AAAA)-mCitrine kan
MGM6989	Δ*clpB* (Vaubourgeix et al., 2015) transformed with pAJF1093	Δ*clpB* attB::Ms *clpB*(Y504A)-mCitrine kan
MGM6991	Δ*clpB* (Vaubourgeix et al., 2015) transformed with pAJF1095	Δ*clpB* attB::Ms *clpB*(L496A)-mCitrine kan
MGM7011	MGM6003 transformed with pAJF522	Δ*dnaK* attB::*dnaK*-mCitrine kan pAJF522 (TetON mCerulean-AL09 hygR)

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MGM7012	MGM6972 transformed with pAJF522	Δ*dnaK* attB::*dnaK*(Y257A)-mCitrine kan pAJF522 (TetON mCerulean-AL09 hygR)
MGM7013	MGM6973 transformed with pAJF522	Δ*dnaK* attB::*dnaK*(M229A)-mCitrine kan pAJF522 (TetON mCerulean-AL09 hygR)
MGM7014	MGM6971 transformed with pAJF539	Δ*clpB* attB::Ms *clpB*-mCitrine kan pAJF539 (TetON mCerulean-AL09 strepR)
MGM7015	MGM6988 transformed with pAJF539	ΔclpB attB::Ms clpB (TGIP-AAAA)-mCitrine kan pAJF539 (TetON mCerulean-AL09 strepR)
MGM7016	MGM6989 transformed with pAJF539	Δ*clpB* attB::Ms *clpB*(Y504A)-mCitrine kan pAJF539 (TetON mCerulean-AL09 strepR)
MGM7018	MGM6991 transformed with pAJF539	Δ*clpB* attB::Ms *clpB*(L496A)-mCitrine kan pAJF539 (TetON mCerulean-AL09 strepR)
MGM7023	ΔclpB (Vaubourgeix et al., 2015) transformed with pMV306kan	Δ*clpB* attB::kan
Recombinant DNA		
Plasmid expressing Mtb clpB	Lupoli et al., 2016	pEcTL02
Plasmid expressing Mtb dnaK	Lupoli et al., 2016	pEcTL06
Plasmid expressing Mtb dnaJ1	Lupoli et al., 2016	pEcTL03
Plasmid expressing Mtb dnaJ2	Lupoli et al., 2016	pEcTL04
Plasmid expressing Mtb grpE	Lupoli et al., 2016	pEcTL05
Plasmid expressing Mtb Hsp20	Lupoli et al., 2016	pEcTL08
Plasmid expressing Ulp1	Lupoli et al., 2016	pHYRS52, Addgene#31122
pAJF1081	in-fusion reaction of Ms clpB (oAF2011/2012, Mc2155 gDNA), mCitrine (203/1431) and pMV306kn (XbaI/ClaI).	attB::Ms clpB-mCitrine kan
pAJF1087	fusion reaction of pAJF200 pcr linearized with oAF2086/2087.	attB: dnaK(Y257A)-mCitrine kan
pAJF1088	fusion reaction of pAJF200 pcr linearized with oAF2088/2089.	attB: dnaK(M229A)-mCitrine kan
pAJF1092	in-fusion reaction of pAJF1081 pcr linearized with oAF2063/2064.	attB::clpB(TGIP-AAAA)-mCitrine kan
pAJF1093	in-fusion reaction of pAJF1081 pcr linearized with oAF2065/2066.	attB::clpB(Y504A)-mCitrine kan
pAJF1095	in-fusion reaction of pAJF1081 pcr linearized with oAF2069/2070.	attB::clpB(L496A)-mCitrine kan
pAJF522	AL09 (Addgene plasmid # 47081) and mCerulean ligated to pmsg419	TetR-Pmyc-tetO-mCerulean-AL09 oriM hyg
pAJF539	hyg marker removed from pAJF522 and replaced with strep marker	TetR-Pmyc-tetO-mCerulean-AL09 oriM strep
Software and algorithms		
RELION 3	Zivanov et al., 2018	https://www3.mrc-lmb.cam.ac.uk/relion//index.php/Download_&_install
Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera/
Coot program	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
PHENIX program	Adams et al., 2010	https://www.phenix-online.org/download/
PyMOL program	Schrödinger, LLC.	https://pymol.org/2/
HHpred program	Söding et al., 2005	https://toolkit.tuebingen.mpg.de/tools/hhpred
Rosetta program	Wang et al., 2016	https://rosettacommons.org/software

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MolProbity program	Chen et al., 2010	http://molprobity.biochem.duke.edu/
Other		
C-flat 2/2 300 mesh grid	ELECTRON MICROSCOPY SCIENCES	Cat#CF-223C-50
MicroCal PEAQ ITC	Malvern	https://www.malvernpanalytical.com/en/