Processive extrusion of polypeptide loops by a Hsp100 disaggregase

Avellaneda, Mario J.; Franke, Kamila B.; Sunderlikova, Vanda; Bukau, Bernd; Mogk, Axel; Tans, Sander J.

DOI
10.1038/s41586-020-1964-y

Publication date
2020

Document Version
Final published version

Published in
Nature

Article

Processive extrusion of polypeptide loops by a Hsp100 disaggregase

https://doi.org/10.1038/s41586-020-1964-y

Received: 17 May 2019

Accepted: 3 December 2019

Published online: 29 January 2020

Mario J. Avellaneda[1], Kamila B. Franke[2], Vanda Sunderlikova[1], Bernd Bukau[2], Axel Mogk[2] & Sander J. Tans[1,3]*

The ability to reverse protein aggregation is vital to cells[1,2]. Hsp100 disaggregases such as ClpB and Hsp104 are proposed to catalyse this reaction by translocating polypeptide loops through their central pore[3,4]. This model of disaggregation is appealing, as it could explain how polypeptides entangled within aggregates can be extracted and subsequently refolded with the assistance of Hsp70[4,5]. However, the model is also controversial, as the necessary motor activity has not been identified[6–8] and recent findings indicate non-processive mechanisms such as entropic pulling or Brownian ratcheting[9,10]. How loop formation would be accomplished is also obscure. Indeed, cryo-electron microscopy studies consistently show single polypeptide strands in the Hsp100 pore[11,12]. Here, by following individual ClpB–substrate complexes in real time, we unambiguously demonstrate processive translocation of looped polypeptides. We integrate optical tweezers with fluorescent-particle tracking to show that ClpB translocates both arms of the loop simultaneously and switches to single-arm translocation when encountering obstacles. ClpB is notably powerful and rapid; it exerts forces of more than 50 pN at speeds of more than 500 residues per second in bursts of up to 28 residues. Remarkably, substrates refold while exiting the pore, analogous to co-translational folding. Our findings have implications for protein-processing phenomena including ubiquitin-mediated remodelling by Cdc48 (or its mammalian orthologue p97)[13] and degradation by the 26S proteasome[14].

We studied the disaggregase ClpB, a member of the Hsp100 chaperone family, using single-molecule techniques. Maltose-binding protein (MBP) was coupled to DNA handles at both termini and tethered between polystyrene beads, which were trapped and manipulated with optical tweezers (Fig. 1a). After mechanical unfolding of the protein (Fig. 1a, Extended Data Fig. 1a), the applied force was reduced to a value between 5 and 10 pN, high enough to prevent spontaneous refolding (Fig. 1a). Addition of ATP and ClpB(Y503D)—a mutant altered in the regulatory middle (M) domain that does not require Hsp70 (DnaK) binding for ATPase activation[15]—resulted in isolated episodes of contraction in the bead-to-bead distance (Fig. 1b). Zooming in showed that the effective polypeptide contour length L_e was initially approximately 360 amino acids (aa), as expected for fully unfolded MBP, and then decreased linearly to 0, indicating that the C and N termini of MBP were directly adjacent to each other (Fig. 1c, Extended Data Fig. 1b). After a brief pause, L_e increased abruptly back to 360 residues and then immediately decreased again (Fig. 1c). ClpB thus produced processive substrate translocation runs that ended with a loss of ClpB grip. This in turn caused the substrate to slip and be pulled back by the applied force and hence enabled a new run to start.

Translocation was abolished by using ADP instead of ATP; when either of the two ClpB ATPase catalytic centres (E279A or E678A) were mutated, preventing ATP hydrolysis; when the substrate-contacting pore loops (Y251A or Y653A) were mutated; or by deletion of the N-terminal domain that forms the pore entry. Translocation was also observed for the M-domain mutant K476C and for wild-type ClpB with the Hsp70 (DnaK in bacteria) system[16] (Extended Data Fig. 2a–e). ClpB(K476C) and wild-type ClpB translocated at the same speed as ClpB(Y503D), which was unexpected because both stimulated ATP hydrolysis less strongly in bulk[15] (Extended Data Fig. 2f, g). However, they exhibited translocation for a smaller proportion of the time (Extended Data Fig. 2b), suggesting that the differences in hydrolysis rates reflected the fraction of actively translocating ClpB hexamers rather than their individual translocation speed.

Longer polypeptide constructs of two and four tandem MBP repeats displayed longer runs before slipping, with some exceeding 1,000 residues (Extended Data Figs. 1c–f, 3, 4). The speed distribution displayed two peaks (the second at roughly double the speed of the first) and extended beyond 500 aa per second (Fig. 1d), more than tenfold faster than other peptide translocases[17–19]. This distribution appeared similar for the different substrate constructs and for different individual translocation bursts (Fig. 1b, Extended Data Fig. 3d–f). These bursts probably reflected the activity of single ClpB hexamers, because they consisted of continuous run–slip–run activity and were spaced apart by several seconds. Without slowing down, ClpB exerted high forces of more than 50 pN, resulting in the melting of our DNA tethers (Fig. 1e). These data indicated notable speed, processivity and power.

[1]AMOLF, Amsterdam, The Netherlands. [2]Center for Molecular Biology of Heidelberg University, German Cancer Research Center, Heidelberg, Germany. [3]Department of Bionanoscience, Kavli Institute of Nanoscience Delft, Delft University of Technology, Delft, The Netherlands. *e-mail: s.tans@amolf.nl

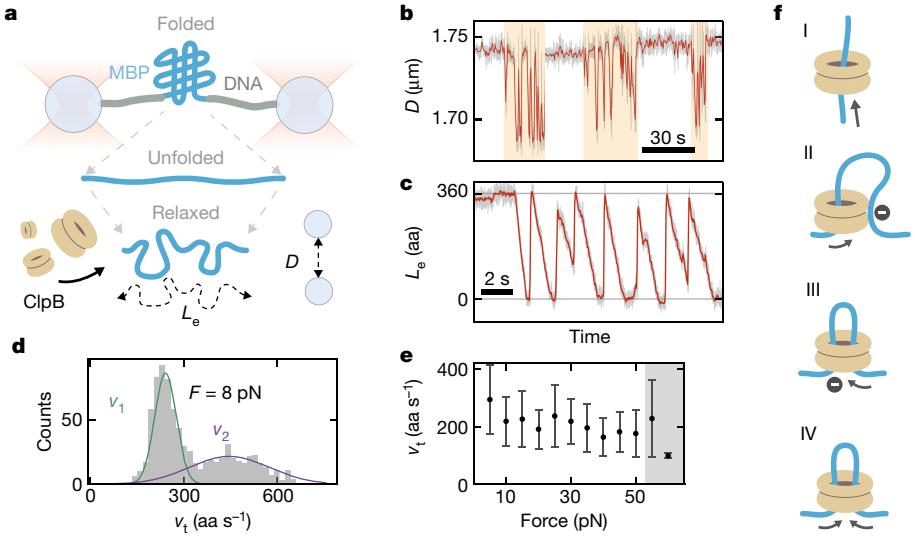


Fig. 1 | ClpB is a processive translocase. a, Tethered MBP was unfolded with optical tweezers, relaxed to a low force that prohibits refolding, and exposed to ClpB and ATP. **b**, Tether contraction bursts (orange regions) with ClpB(Y503D) and ATP. Grey, raw signal (500 Hz); red, filtered signal (2 Hz). **c**, Polypeptide contour length L_e during a contraction burst, as determined from the bead-bead distance (D), force, and worm-like chain model. L_e decreases linearly from 360 aa (MBP fully extended) down to 0 aa (MBP C and N termini directly adjacent), indicating processive translocation by ClpB. Abrupt increases of L_e to 360 aa indicate that ClpB transiently loses grip and substrate slips backwards, pulled by the applied force. Red, filtered signal (20 Hz). **d**, Speed distribution of runs from all MBP substrate constructs in the presence of ClpB(Y503D) and ATP, at a force of approximately 8 pN. Double Gaussian fit shows two mean speeds, $v_2 \approx 2v_1$ with $v_1 = 240 \pm 30$ aa s^{-1} and $v_2 = 450 \pm 130$ aa s^{-1} (mean ± s.d., $n = 800$ runs, 18 molecules). **e**, Mean translocation speed for ClpB(Y503D) versus applied tension ($n = 717$ runs, 8 molecules; see Methods). Grey, DNA melting regime and upper force limit. Data are mean ± s.e.m. **f**, Hypothetical ClpB translocation topologies. Single-strand insertion and translocation (I) does not yield contraction, unless it is immobilized elsewhere on the ClpB surface (II). Dual-strand insertion in a looped topology can also produce contraction, either by single-arm (III) or dual-arm (IV) translocation.

Different hypothetical translocation models or topologies could be considered (Fig. 1f). Even when their termini are not free for insertion, as is the case here, single polypeptide chains can be accommodated into the ClpB pore by rings that open and close or that assemble around them[20,21] (model I). However, this scenario would only produce the observed contractions if a second chain site is immobilized on ClpB (model II), analogous to DNA processing by condensin[22]. Alternatively, the substrate could be inserted as a loop into the central pore, with translocation of one (model III) or both (model IV) arms of the loop.

Testing these models with optical tweezers is difficult. We therefore developed a technique that allows independent measurement of the length of each arm of the polypeptide loop and integrates optical tweezers with ClpB tracking at sub-wavelength resolution using single-molecule fluorescence imaging (Fig. 2a, Extended Data Fig. 5). We chose the construct with two maltose-binding protein (MBP) repeats (2MBP), as it yields longer runs, and exposed it to fluorescently labelled ClpB and ATP after unfolding, while scanning a confocal excitation beam along the tether and beads (Fig. 2b). To limit the parasitic signal emanating from the beads, we developed a protein–DNA coupling protocol that enabled the attachment of long 5-kilobase pair (kbp) DNA handles (see Methods). Single ClpB-binding events were identified by a fluorescent spot appearing between the beads (Fig. 2b), and translocation was observed soon after (Fig. 2c). We next moved to an ATP-only solution to reduce background fluorescence and prevent further ClpB binding, and tracked the spot position using Gaussian fitting (Fig. 2d–f). Combining the tweezers and tracking data yielded the distances between ClpB and each of the MBP termini, and hence the translocation activity on both loop arms independently (Fig. 2a, Methods).

We found various sequences of events: after translocation of the entire chain, the left arm of the loop was released and slipped backwards until the full chain was again extended in *cis*, and subsequently left-arm translocation restarted rapidly (Fig. 2g, h, event sequence A→B→C). A similar sequence on the right side occurred directly afterwards (Fig. 2g, h, A→D→E). We also observed both arms being translocated simultaneously, each at similar velocity (Fig. 2g, i, event F, Extended Data Fig. 5i–k). Consistently, the total translocation speed, which reflects the velocity at which both polypeptide termini approach each other, and is more accurate as it is based only on the signal from the optical tweezers, was then twice as high ($2v$) as in single-arm translocation runs (v) (Fig. 2k, grey region). Model II does not allow for two-arm translocation and hence was not consistent with the data, whereas models III and IV were consistent with the data. Switches between single- and two-arm translocation modes took place after blockage of one arm, typically on ClpB encountering the DNA tether at either terminus. The data also provided direct confirmation that single ClpB rings remained intact and bound during runs, switches and back-slips (Fig. 2h, i).

This scenario was supported by multiple additional observations. First, 64% of the very first runs in a translocation burst initially showed the higher speed ($2v$) before switching to the lower speed (v), compared with 22% when considering all runs (Extended Data Fig. 6a). Indeed, the initial ClpB binding site is probably not directly adjacent to the DNA handles at the termini, and thus both arms are then unobstructed when translocation starts. Initial binding regions estimated from these experiments were consistent with peptide scanning data, although we note that both methods yield rough estimates (Extended Data Fig. 6). Second, experiments at increased resolution showed that lower-speed (v) runs were composed of individual translocation steps of 14.6 ± 0.9 aa, whereas higher-speed ($2v$) runs were in steps of 28 ± 3 aa (Fig. 3a–d, Extended Data Fig. 7a–e). These findings are consistent, since decreases in distance between termini should be twofold larger when both arms are translocated simultaneously. ClpB thus switches between translocation modes by changing the step size rather than the step frequency (Fig. 3e).

We next investigated how these stepping dynamics relate to the structure of ClpB. Each ClpB monomer is thought to move substrates by approximately 2 aa, substantially less than the observed 14-aa or 28-aa steps[12,23]. A proposed concerted action[24] of all subunits together

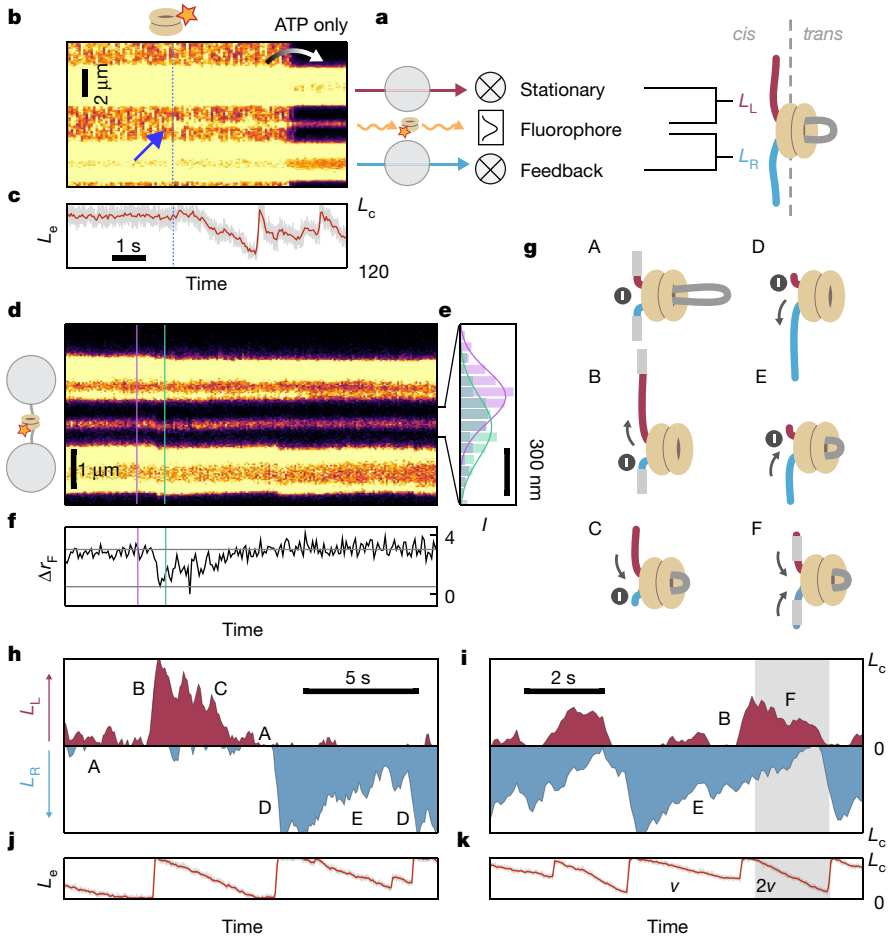


Fig. 2 | Optical tweezers with fluorescence reveals ClpB translocation of both loop arms. a, Principle of approach: one trap is continuously moved to maintain force constant. Bead positions yield polypeptide N-to-C-terminal distance at nanometre precision (expressed in contour length L_e). Confocal fluorescence imaging of ClpB–Atto633 yields its position at sub-wavelength precision using Gaussian fitting. Together, they quantify the lengths of both non-translocated (*cis*) polypeptide arms: L_R (blue) and L_L (purple). **b**, Fluorescence kymograph from scans along beads and tether, showing ClpB binding (blue arrow) and movement to the ClpB-free region. **c**, Concurrent tweezers data of polypeptide contour length L_e, showing translocation start soon after ClpB binding. **d**, Kymograph during translocation. **e**, Photon count of ClpB spot along two scans and Gaussian fits that determine position. **f**, Position of ClpB–Atto633 (in number of pixels), moving suddenly down at back-slip B (**h**) and gradually up during translocation C. Back-slip D does not change ClpB position, because slip is on the right (blue), and left-arm linked to stationary bead remains unchanged (purple). Top line, ClpB at left-hand MBP terminus; bottom line: ClpB is at right-hand terminus and polypeptide is fully in *cis*. Consistently, ClpB deviates from top line when tweezers detects back-slip. **g**, Cartoons indicating positions corresponding to plots in **h** and **i**. A, polypeptide is fully in *trans* (L_e = 0); B and D, back-slip of left and right arm, respectively. C and E, translocation of left and right arm, respectively. F, translocation of both arms. **h**, **i**, L_R and L_L for kymographs shown in **d** and Extended Data Fig. 5g. Grey-shaded region, double-speed translocation. Consistently, both arms shorten simultaneously. **j**, **k**, Total *cis*-polypeptide length, $L_e = L_R + L_L$ from tweezers alone.

would yield a continuous series of 2-aa steps, and thus appears inconsistent with these results. By contrast, the six subunits acting in rapid consecutive manner would produce steps similar to those detected here (approximately 12 aa, or approximately 24 aa when both arms are moving). The pauses between steps could thus reflect a slow transition within the ATP cycle occurring in all subunits[25]. Zooming into the steps should show six substeps of around 2 aa, but these cannot be resolved owing to the particularly high translocation speed and the need to time-average at these length scales. Therefore, we mixed ATP with the poorly hydrolysable ATPγS, as this would be expected to interrupt sequential subunit action moving along the ClpB ring and therefore yield smaller steps. Translocation was more erratic and indeed showed smaller steps well below 14 aa in size (Extended Data Fig. 7g–i), rather than longer pauses only. These data thus supported the sequence-pause model.

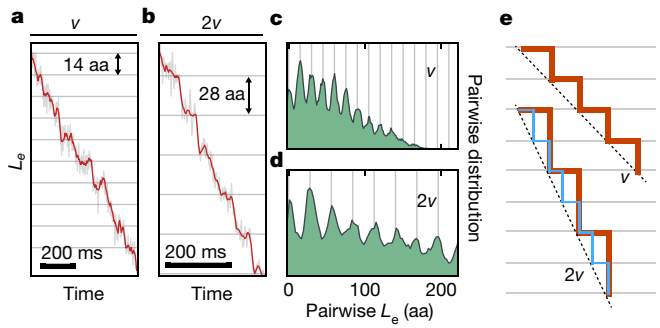


Fig. 3 | Translocation steps by ClpB. a, **b**, Plot of L_e for single-speed (**a**) and dual-speed (**b**) translocation runs. Red, Savitzky–Golay filtering. **c**, **d**, Distribution of L_e difference between any two points for one single-speed run (**c**) or one dual-speed run (**d**). The regularly spaced peaks indicate a step size of 14.6 ± 0.9 aa for the single-speed run; the peak spacing for the dual-speed run is doubled, yielding a step-size of 28 ± 3 aa. Data are mean ± s.e.m. calculated from n = 12 runs. **e**, The data show that the speed is doubled by doubling the step size (red), not the step frequency (blue).

The complex dynamics observed thus far can be further complicated when folded structures are present within the looped polypeptide. Specifically, we found back-slips for 2MBP that were incomplete, with a segment of about 270 aa remaining on the *trans* side of ClpB (Fig. 4a, b). This is exactly the length of one MBP core, suggesting that the polypeptide folded after translocation (in line with its normal folding time of about 1 s)[26] and was subsequently blocked at the *trans* side of the narrow ClpB pore when pulled backwards during a back-slip (Fig. 4d). Consistently, such incomplete back-slips only occurred after full MBP cores were translocated (Fig. 4c) and were not observed for folding-compromised 2MBP mutants[27] or for 1MBP, whose core cannot fold because a key segment remains stuck in the ClpB pore (Extended Data Fig. 8). Of note, substrates refolded at the exit of the ClpB channel, analogous to co-translational folding of nascent chains, and without requiring DnaK.

Conversely, misfolded structures already present within the chain should be blocked at the *cis* side of ClpB during translocation. Such an obstruction of translocation can ramp up local forces and in turn pull apart the blocking misfolded structure. Indeed, we observed such disruption when a partially folded MBP or a small MBP aggregate was exposed to ClpB (Extended Data Fig. 9). The disruption events were directly followed by translocation runs because unfolded polypeptides produced by structure disruption on the *cis* side of ClpB are available for translocation. These data indicate how folded structures present in *cis* and *trans* can affect translocation dynamics in a looped topology.

In conclusion, our study on ClpB shows unambiguously that polypeptide loop extrusion is possible. Free substrate termini may also

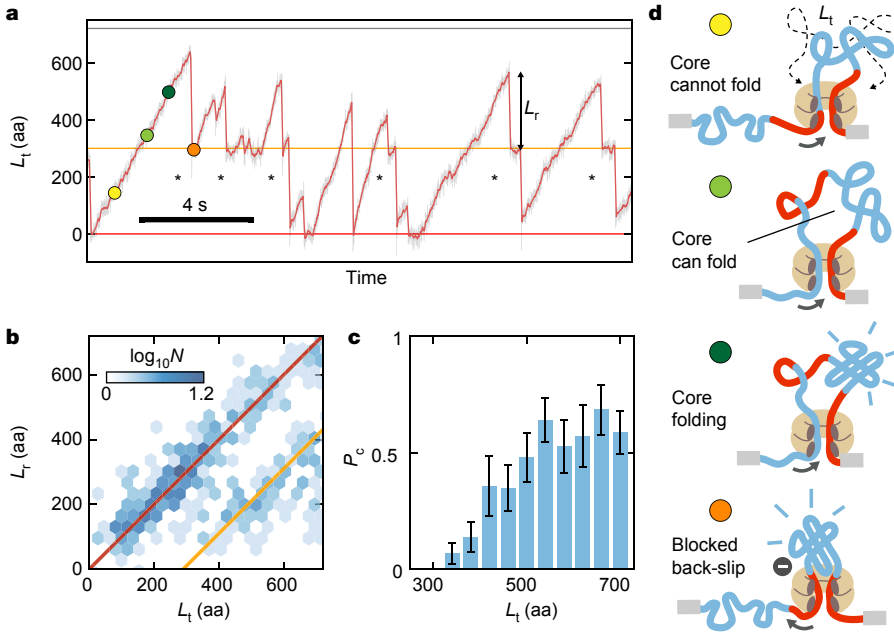


Fig. 4 | Substrate refolding on the ClpB *trans*-side during translocation. a, Translocated polypeptide length (L_t) for 2MBP. Horizontal red and orange lines, $L_t = 0$ and $L_t = 310$ aa, respectively (a 270-aa MBP core in *trans*, and 2 × 20 aa inside the ClpB pore). Stars, back-slip arrested at $L_t = 310$ aa. Circles, events illustrated in **d**. **b**, Red line, translocated length is fully released to the *cis* side upon back-slip ($L_r = L_t$). Orange line corresponds to an MBP core refolding on the *trans* side and then blocked from passing the ClpB pore upon back-slip ($L_r = L_t - 310$ aa). **c**, Probability of core folding (P_c). Consistently, P_c is 0 below 310 aa, when the MBP core is not fully translocated ($n = 266$ runs, 12 molecules, mean ± standard error of binomial distribution, see Methods). **d**, Cartoons of *trans* refolding. Blue, core MBP; red, non-core MBP. MBP cores do not refold on the *cis* side because tweezers constrain the polypeptide chain. In cells, aggregates may constrain substrates from folding. On the *trans* side, the chain is freed from these constraints and an MBP core can fold when fully translocated (light and dark green dots) and thus cannot pass through the ClpB pore when back-slip occurs (orange dot and line).

insert into the ClpB pore and be translocated in a non-loop topology, though we surmise that internal segments of aggregated proteins are targeted more readily and hence translocated as loops. ClpB is fast, processive, generates large forces, and can switch between single- and dual-arm translocation. ClpB thus appears to maintain a tight and long-term grip on both arms, with back-slips indicating a sporadic loss of contact. It remains an open question how the independent handling of two arms is achieved at the structural level. These features of ClpB are relevant to efficient disaggregation (Extended Data Fig. 10). Full dissolution of stable aggregates probably involves multiple ClpB rings and other chaperones such as Hsp70/DnaK acting at different sites, at different moments in time, and involving many random dissociation and re-association events. Nevertheless, ClpB translocation itself is remarkably deterministic and processive once started.

Overall, our findings define loop extrusion as the mechanistic basis of Hsp100 disaggregation, highlight the need for tight regulation of Hsp100 activity and suggest that other polypeptide processing systems such as the Cdc48 (mammalian orthologue p97) segregase, the ribosomal assembly factor Rix7 (mammalian orthologue NVL), and the 26S proteasome may also exploit the capability to handle multiple polypeptide strands in a controlled manner.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-1964-y.

1. Soto, C. & Pritzkow, S. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat. Neurosci.* **21**, 1332–1340 (2018).
2. Chiti, F. & Dobson, C. M. Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **86**, 27–68 (2017).
3. Weibezahn, J. et al. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell* **119**, 653–665 (2004).
4. Mogk, A., Bukau, B. & Kampinga, H. H. Cellular handling of protein aggregates by disaggregation machines. *Mol. Cell* **69**, 214–226 (2018).
5. Kummer, E. et al. Bacterial and yeast AAA+ disaggregases ClpB and Hsp104 operate through conserved mechanism involving cooperation with Hsp70. *J. Mol. Biol.* **428**, 4378–4391 (2016).
6. Shorter, J. & Southworth, D. R. Spiraling in control: structures and mechanisms of the Hsp104 disaggregase. *Cold Spring Harb. Perspect. Biol.* **11**, 034033 (2019).
7. Sousa, R. Structural mechanisms of chaperone mediated protein disaggregation. *Front. Mol. Biosci.* **1**, 12 (2014).
8. Liberek, K., Lewandowska, A. & Ziętkiewicz, S. Chaperones in control of protein disaggregation. *EMBO J.* **27**, 328–335 (2008).
9. Li, T. et al. *Escherichia coli* ClpB is a non-processive polypeptide translocase. *Biochem. J.* **470**, 39–52 (2015).
10. Durie, C. L. et al. Hsp104 and potentiated variants can operate as distinct nonprocessive translocases. *Biophys. J.* **116**, 1856–1872 (2019).
11. Gates, S. N. et al. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. *Science* **357**, 273–279 (2017).
12. Deville, C., Franke, K., Mogk, A., Bukau, B. & Saibil, H. R. Two-step activation mechanism of the ClpB disaggregase for sequential substrate threading by the main ATPase motor. *Cell Rep.* **27**, 3433–3446 (2019).
13. Bodnar, N. O. et al. Structure of the Cdc48 ATPase with its ubiquitin-binding cofactor Ufd1–Npl4. *Nat. Struct. Mol. Biol.* **25**, 616–622 (2018).
14. Dong, Y. et al. Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. *Nature* **565**, 49–55 (2019).
15. Oguchi, Y. et al. A tightly regulated molecular toggle controls AAA+ disaggregase. *Nat. Struct. Mol. Biol.* **19**, 1338–1346 (2012).
16. Fernández-Higuero, J. A., Aguado, A., Perales-Calvo, J., Moro, F. & Muga, A. Activation of the DnaK–ClpB complex is regulated by the properties of the bound substrate. *Sci. Rep.* **8**, 5796 (2018).
17. Maillard, R. A. et al. ClpX(P) generates mechanical force to unfold and translocate its protein substrates. *Cell* **145**, 459–469 (2011).
18. Aubin-Tam, M.-E., Olivares, A. O., Sauer, R. T., Baker, T. A. & Lang, M. J. Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. *Cell* **145**, 257–267 (2011).
19. Olivares, A. O., Nager, A. R., Iosefson, O., Sauer, R. T. & Baker, T. A. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. *Nat. Struct. Mol. Biol.* **21**, 871–875 (2014).
20. Werbeck, N. D., Schlee, S. & Reinstein, J. Coupling and dynamics of subunits in the hexameric AAA+ chaperone ClpB. *J. Mol. Biol.* **378**, 178–190 (2008).
21. Uchihashi, T. et al. Dynamic structural states of ClpB involved in its disaggregation function. *Nat. Commun.* **9**, 2147 (2018).
22. Ganji, M. et al. Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).
23. Yu, H. et al. ATP hydrolysis-coupled peptide translocation mechanism of *Mycobacterium tuberculosis* ClpB. *Proc. Natl Acad. Sci. USA* **115**, E9560–E9569 (2018).
24. Lyubimov, A. Y., Strycharska, M. & Berger, J. M. The nuts and bolts of ring-translocase structure and mechanism. *Curr. Opin. Struct. Biol.* **21**, 240–248 (2011).
25. Moffitt, J. R. et al. Intersubunit coordination in a homomeric ring ATPase. *Nature* **457**, 446–450 (2009).
26. Bechtluft, P. et al. Direct observation of chaperone-induced changes in a protein folding pathway. *Science* **318**, 1458–1461 (2007).
27. Chakraborty, K. et al. Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell* **142**, 112–122 (2010).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Protein expression and purification

All MBP constructs were modified at both termini with cysteine residues using the pET28 vector. Double-mutant MBP harbours two mutations (V8G and Y283D) that hinder folding[28]. Proteins were purified from *Escherichia coli* BL21(DE3) cells. For overexpression, overnight cultures were diluted 1:100 in fresh LB medium supplemented with 50 mg l^{-1} kanamycin, 0.2% glucose and incubated under vigorous shaking at 30 °C. Expression was induced at OD_{600} = 0.6 by addition of 1 mM IPTG and incubation overnight at room temperature. Cells were cooled, collected by centrifugation at 5000*g* for 20 min, flash-frozen and stored at −80 °C. Cell pellets were resuspended in ice-cold buffer A (50 mM potassium phosphate pH 7.5, 150 mM NaCl, 3 mM chloramphenicol, 50 mM Glu–Arg, 10 mM Complete Protease Inhibitor Ultra (Roche), 10 mM EDTA) and lysed using a pressure homogenizer. The lysate was cleared from cell debris by centrifugation at 50,000*g* for 60 min and incubated with Amylose resin (New England Biolabs) that was previously equilibrated in buffer A for 20 min at 4 °C. The resin was washed with buffer A three times by centrifugation and bound proteins were eluted in buffer A supplemented with 20 mM maltose. Purified proteins were aliquoted, flash-frozen in liquid nitrogen and stored at −80 °C. ClpB and variants were overexpressed in *E. coli ΔclpB::kan* cells. Cell pellets were resuspended in LEW buffer (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 5 mM β-mercaptoethanol) and lysed by French press. Cleared supernatants were incubated with Protino Ni-IDA resin and bound proteins were eluted by LEW buffer containing 250 mM imidazole. ClpB containing fractions were pooled and subjected to Superdex S200 16/60 size-exclusion chromatography in MDH buffer (50 mM Tris pH 7.5, 150 mM KCl, 20 mM $MgCl_2$, 2 mM DTT) containing 5% (v/v) glycerol.

ClpB-Atto633 labelling

Labelling of ClpB-E731C variants with Atto633-maleimide was performed in PBS buffer according to the instructions of the manufacturer (ATTO-TEC). Labelled ClpB-E731C was separated from non-reacted Atto633 by size-exclusion chromatography using Superdex S200 HR10/30 in MDH buffer containing 5% (v/v) glycerol.

Attachment of DNA handles to substrates

Protein substrates were buffer-exchanged using a PD-10 desalting column (GE Healthcare) to remove reducing agents and elutants. Next, they were incubated with a 4× excess maleimide-modified oligonucleotides (20 nucleotides) for 1 h at 30 °C. Uncoupled oligos were removed using Amylose resin (NEB). The coupled protein was then ligated to 2.5-kbp DNA tethers presenting a complementary 20-nucleotide single-stranded overhang using T4 ligase for 1 h at room temperature.

Optical tweezers assay

Carboxyl polystyrene beads (CP-20-10, diameter 2.1 μm, Spherotech) were covalently coated with sheep anti-digoxigenin antibody (Roche) via a carbodiimide reaction (PolyLink Protein coupling kit, Polysciences). Approximately 50 ng of the generated construct were incubated with 2 μl beads in 10 μl buffer C (50 mM HEPES pH 7.5, 5 mM $MgCl_2$, 100 mM KCl) for 15 min in a rotary mixer at 4 °C and rediluted in 350 μl buffer C. With our coupling strategy, approximately 50% of the constructs were asymmetrically functionalized with digoxigenin and biotin in each side. In order to create the second connection, we used NeutrAvidin-coated polystyrene beads (NVP-20-5, diameter 2.1 μm, Spherotech). Once trapped, beads were brought into close proximity to allow binding, and tether formation was identified by an increase in force when the beads were moved apart. ClpB was diluted in buffer C to a final concentration of 2 μM. For the ATP experiments, we used an ATP regeneration system (3 mM ATP, 20 ng μl^{-1} pyruvate kinase, 3 mM phosphoenol pyruvate). Experiments were performed in the presence of an oxygen scavenging system[29] (3 units per ml pyranose oxidase, 90 units per ml catalase and 50 mM glucose, all purchased from Sigma-Aldrich) to prevent DNA and protein oxidation damage.

Single-molecule data analysis and ClpB translocation event characterization

Data was recorded at 500 Hz using a custom-built dual trap optical tweezers and a C-Trap (Lumicks). Data was analysed using custom scripts in Python. The optical traps were calibrated using the power spectrum of the Brownian motion of the trapped beads[28], obtaining average stiffness values of $\kappa = 0.39 \pm 0.04$ pN nm^{-1}. Most measurements were taken in an active force-clamp regime, in which one of the traps was moved in response to changes in the force using a proportional–integral–derivative (PID) feedback loop (Extended Data Fig. 5e, f). Individual force-extension curves were identified and fitted to two worm-like-chain (WLC) models in series (Extended Data Fig. 1a), using the approximation of an extensible polymer reported in ref. [30] for the DNA, and the Odijk inextensible approximation for the protein contribution[31].

$$x = L_c^* \left(\frac{4}{3}\left(1 - \frac{1}{\sqrt{\beta^* + 1}}\right) - \frac{10 e^{4\sqrt{\frac{900}{\beta^*}}}}{\sqrt{\beta^*}\left(e^{4\sqrt{\frac{900}{\beta^*}}} - 1\right)^2} + \frac{\beta^{*1.62}}{3.55 + 3.8\beta^{*2.2}} \right) + L_e\left(1 - \frac{1}{2}\sqrt{\beta}\right)$$

Where $\beta^* = \frac{F L_p^*}{k_B T}$, F is the force, T is the temperature and L_p^*, K and L_c^* are the persistence length, stretch modulus and contour length of DNA, respectively; $\beta = \frac{F L_p}{k_B T}$, where L_p and L_e are the persistence and extended length of the protein, respectively. L_c^* values were 906, 1,750 and 3,500 nm for the three different DNA handles used (1.3, 2.5 and 5 kb, respectively), and L_p was 0.75 nm. L_p^* and K were fitted, yielding average values of 30 nm and 700 pN nm^{-1}, respectively. These fitted parameters were then used to compute the instantaneous extended length of the protein (L_e) using the same WLC model (Extended Data Fig. 1b). The translocated length (L_t) was computed by subtracting the extended length (L_e) to the total contour length of the protein (L_c). The unfiltered data (500 Hz) is displayed in all panels in grey. With the exception of Fig. 1b, the red signal always indicates data decimated to 20 Hz.

To classify translocation events, the translocated length signal was smoothed using a Savitzky–Golay filter[32] (Extended Data Fig. 4c, black line), enabling its time derivative to be calculated without large fluctuations. Back-slipping results in a large negative slope in the derivative, which was used as the criteria to separate individual translocation runs (Extended Data Fig. 4d). Subsequent one-dimensional dilation and erosion was performed to remove artefacts. Next, each individual run was similarly treated in order to find local changes in the slope (Extended Data Fig. 4e), setting as threshold a value between the two known speeds (around 70 and 140 nm s^{-1}, Extended Data Fig. 4f). Linear fits were performed in each identified region and reported as the local translocation speed (Extended Data Fig. 4e). Only fits that yielded r values higher than 0.8 were considered unless otherwise stated. Speed distributions were computed using the speeds of all valid runs for each condition.

Translocation-step characterization

To increase the spatial resolution[33], we tethered a single MBP using 1,300-bp DNA handles, 500 μM ATP and high tension (>20 pN). Raw data was smoothed using a Savitzky–Golay filter of 5th order with a window of 21 data points. The difference between every distinct pair of data points was calculated and the sample was binned to compute

the pairwise length distribution. To calculate the periodicity more accurately, we computed the autocorrelation of the pairwise length distribution using the Pearson correlation coefficient for different lag lengths (Extended Data Fig. 7a, b) and its power spectrum using the Welch method[34] (Extended Data Fig. 7c, d). The autocorrelation distribution was fitted using the equation:

$$\left(A\cos\left(\frac{2\pi}{T}x\right)+mx+n\right)Be^{-Cx}$$

where T is the period of the steps and a linear and exponential function have been introduced to account for the decay in the signal (Extended Data Fig. 7a, b). The peak in the power spectrum was fitted to a Gaussian distribution (Extended Data Fig. 7c, d).

Confocal fluorescence measurements

An excitation laser beam with a wavelength of 638 nm and a power of 1.3 mW was scanned along the beads and tether at a line rate of 12 Hz. To avoid parasitic noise from the beads, proteins were tethered using two 5-kbp instead of 2.5-kbp DNA handles. In addition, the 2MBP construct was used in order to observe larger distance changes. Force spectroscopy and confocal microscopy data were synchronized based on the movement of the beads (Extended Data Fig. 5a–d). The edge of the moving bead was tracked using a Gaussian fit (Extended Data Fig. 5b) and overlying it on top of the optical tweezers signal for the bead movement showed a time offset (Extended Data Fig. 5c). In order to determine the value of this offset, we computed the root-mean-square deviation (r.m.s.d.) between the signals for different time offsets (Extended Data Fig. 5d):

$$\text{r.m.s.d.}(\tau)=\sqrt{\frac{\sum_i^{N(\tau)}(X_i(\tau)-x_i(\tau))^2}{N(\tau)}}$$

Where τ is the time offset applied to the tracked signal, $N(\tau)$ is the total number of points, $X(\tau)$ is the position of the bead according to the voltage of the mirror and $x(\tau)$ is the tracked position from the fluorescence kymograph. Minimization of r.m.s.d.(τ) provides an excellent estimate of the time offset between the signals (Extended Data Fig. 5d).

Integration of optical tweezers and imaging signals to compute the length components

After ClpB binding and moving to a region containing only ATP, the fluorescent spot between the beads was fitted to a Gaussian distribution. To reduce the noise of the signal, we averaged the intensity profiles of three scanning lines before fitting. The resulting trajectory yielded the absolute position of ClpB with subpixel precision, which was then converted to nanometres using a factor of 80 nm per pixel.

Next, we computed the position of each bead edge that is closest to ClpB (bottom edge for top bead and vice versa) using the trap position, bead displacement and bead radius. Although it is possible to obtain these positions from the fluorescence kymograph, the optical tweezers data yield higher spatial resolution. We subtracted the ClpB tracked position from the position of bottom edge of the top bead, and we subtracted the position of the top edge of the bottom bead from the ClpB position. These distances contain an arbitrary shift owing to the mismatched reference system of the optical tweezers and confocal fluorescence images. In order to identify the offset, we used the fact that when the polypeptide is completely translocated (information present in the optical tweezers data, such as Fig. 1c or Extended Data Fig. 2), both distances should be equal to each other and equal to half the distance D between the edges of the beads. After correcting for the shift, we obtained the absolute distance between ClpB and each of the beads (D_L and D_R). Since we use a force clamp, any change in distance is solely due to a change in the protein length ($\Delta x_L = \Delta D_L$ and $\Delta x_R = \Delta D_R$, Extended Data Fig. 5e, f). Therefore, we removed the constant DNA contribution and computed the protein contour length from each distance (L_L and L_R) using the WLC model.

Peptide library data and initial ClpB binding location

The MBP peptide library was prepared by automated spot synthesis by JPT Peptide Technologies (PepSpots). The library is composed of 13-mer peptides scanning the MBP primary sequence with an overlap of 10 residues. One micromolar ClpB-NTD (Met1–Ser148) was incubated for 30 min in buffer P (10 mM Tris pH 7.5, 150 mM KCl, 20 mM $MgCl_2$, 5% (w/v) sucrose and 0.005% (v/v) Tween 20) with the library. Afterwards, buffer P was discarded and the membrane was washed with cold TBS (50 mM Tris pH 7.6, 150 mM NaCl). Fractionated western blotting enabled transfer of ClpB-NTD bound to peptide spots onto PVDF membranes and bound ClpB-NTD was detected by use of specific, polyclonal (rabbit) anti-ClpB-NTD serum.

The obtained blot image was divided in regions and the individual intensities were computed (Extended Data Fig. 6d). A Gaussian filter was applied to the resulting distribution to account for sequence overlapping and mirroring was performed for direct comparison with the optical tweezers data (Extended Data Fig. 6e).

ATPase activity assay

MBP–DM was denatured in 50 mM Tris, 25 mM KCl, 10 mM $MgCl_2$, 6 M urea and 2 mM DTT. The ATPase activity of the different variants was determined in 50 mM Tris, 25 mM KCl, 10 mM $MgCl_2$, 0.4 M urea and 2 mM DTT in presence of 2 mM ATP. ATPase measurements were started by addition of substrate.

Additional statistical calculations

Error bars of proportion histograms (Fig. 4c and Extended Data Fig. 6a) were calculated using the standard error of a binomial distribution:

$$\sigma=\sqrt{\frac{p(1-p)}{N}}$$

where p is the success proportion and N is the total number of observations.

Statistical sizes for bar plots are: Fig. 1e: 20, 9, 52, 58, 31, 48, 77, 40, 4, 41, 14 and 5 for each point; Fig. 4c: 41, 30, 29, 14, 23, 23, 25, 19, 14, 19 and 29 for each bar.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request.

Code availability

All data were analysed using a custom Python package that is available online and can be downloaded upon request to the corresponding author.

28. Berg-Sørensen, K. & Flyvbjerg, H. Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.* **75**, 594–612 (2004).
29. Swoboda, M. et al. Enzymatic oxygen scavenging for photostability without pH drop in single-molecule experiments. *ACS Nano* **6**, 6364–6369 (2012).
30. Petrosyan, R. Improved approximations for some polymer extension models. *Rheol. Acta* **56**, 21–26 (2017).
31. Odijk, T. Stiff chains and filaments under tension. *Macromolecules* **28**, 7016–7018 (1995).
32. Savitzky, A. & Golay, M. J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639 (1964).
33. Forns, N. et al. Improving signal/noise resolution in single-molecule experiments using molecular constructs with short handles. *Biophys. J.* **100**, 1765–1774 (2011).

34. Welch, P. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **15**, 70–73 (1967).

Acknowledgements We thank E. Koers for help with the substrate constructs and M. E. Aubin-Tam, K. Ganzinger, S. Werner and F. Wruck for comments and critical reading of the manuscript. This study was supported by the Netherlands Organization for Scientific Research (NWO) and by grants of the Deutsche Forschungsgemeinschaft (BB617/17-2, MO970/4-2 and MO 970/4-3) to B.B. and A.M., and the AmPro program of the Helmholtz Society to B.B.

Author contributions M.J.A., B.B., A.M. and S.J.T. conceived the research. K.B.F. and A.M. purified all chaperone variants and performed the biochemical assays. M.J.A. and S.J.T. designed the single-molecule experiments. M.J.A. and V.S. generated the substrate constructs. M.J.A. carried out the single-molecule experiments, developed the Python software and performed the data analysis. M.J.A. and S.J.T. wrote the manuscript with the input and discussion of all authors.

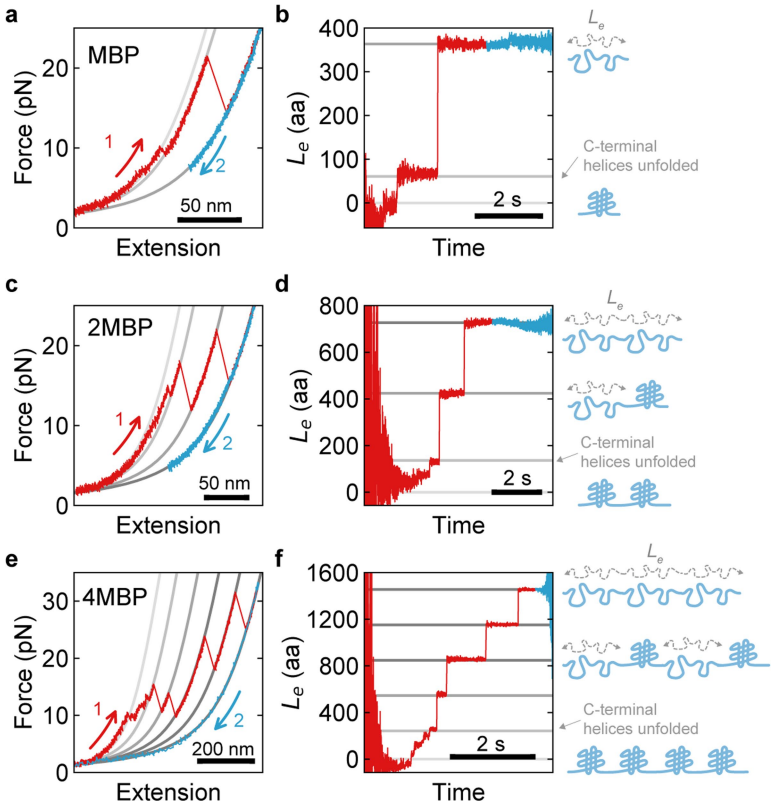
Competing interests The authors declare no competing interests.

Additional information

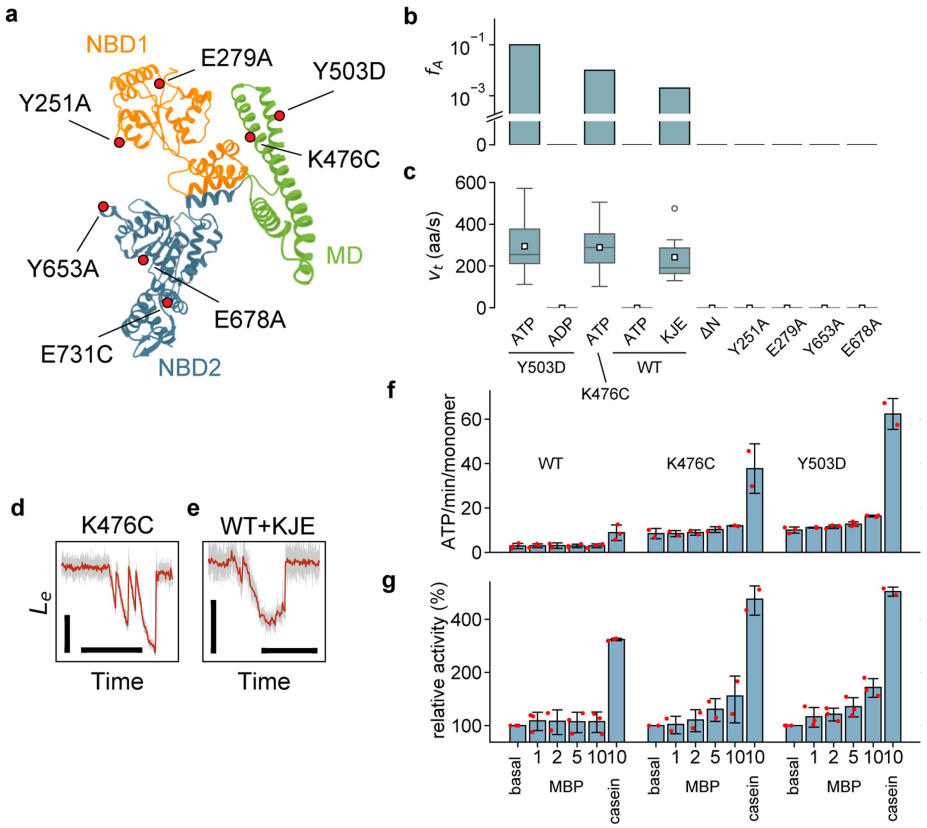
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-1964-y.

Correspondence and requests for materials should be addressed to S.J.T.

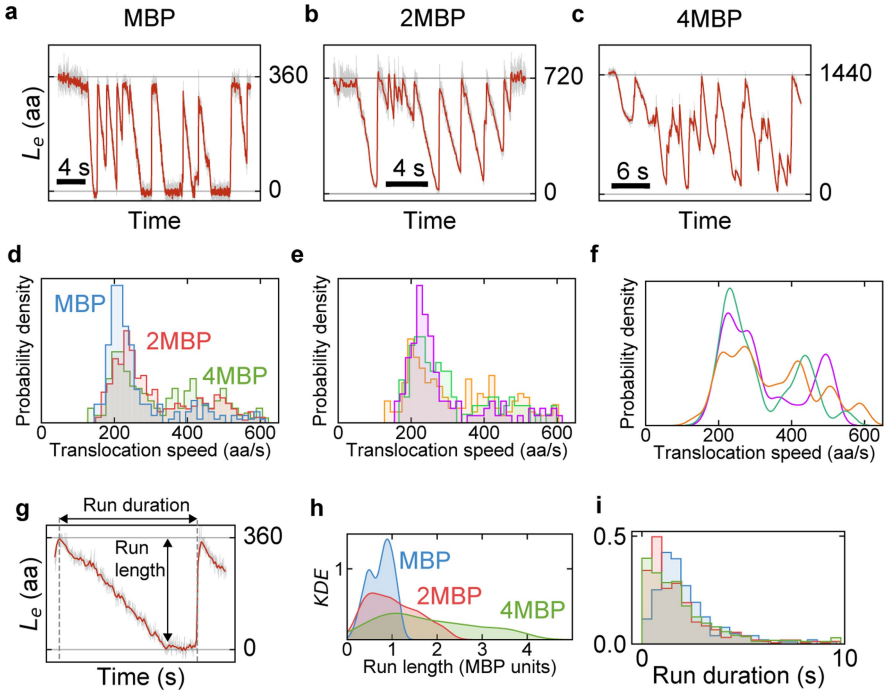
Peer review information *Nature* thanks Toshio Ando, Eilika Weber-Ban, Gijs Wuite and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



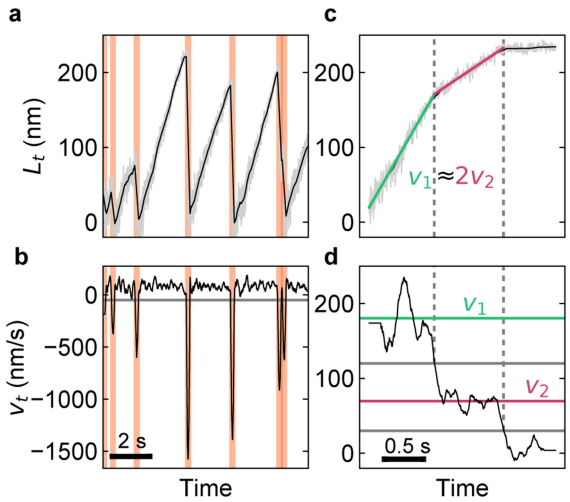
Extended Data Fig. 1 | Mechanical unfolding of substrates and extended length description. a, **c**, **e**, Force-extension curves showing the characteristic unfolding pattern: MBP (**a**), the 2MBP (**c**) and the 4MBP construct (**e**), with an initial gradual and discrete unfolding of C-terminal α-helices (Extended Data Fig. 8a) followed by a sharp unfolding of the cores. Grey lines show WLC fits to the data. Red indicates pulling and blue indicates relaxing of the protein chain. **b**, **d**, **f**, The corresponding extended length L_e of MBP (**b**), the 2MBP (**d**) and the 4MBP construct (**f**). L_e reflects the contour length along the polypeptide backbone, but only of the unfolded part of the protein that is compliant (that is, unfolded and at the *cis*-side of ClpB). L_e is determined from the measured force and extension (distance between beads), and using the WLC model of a non-interacting chain. Grey lines, contour length values obtained from the WLC fits. At low forces, the WLC curves of different contour lengths converge, yielding noisy data.



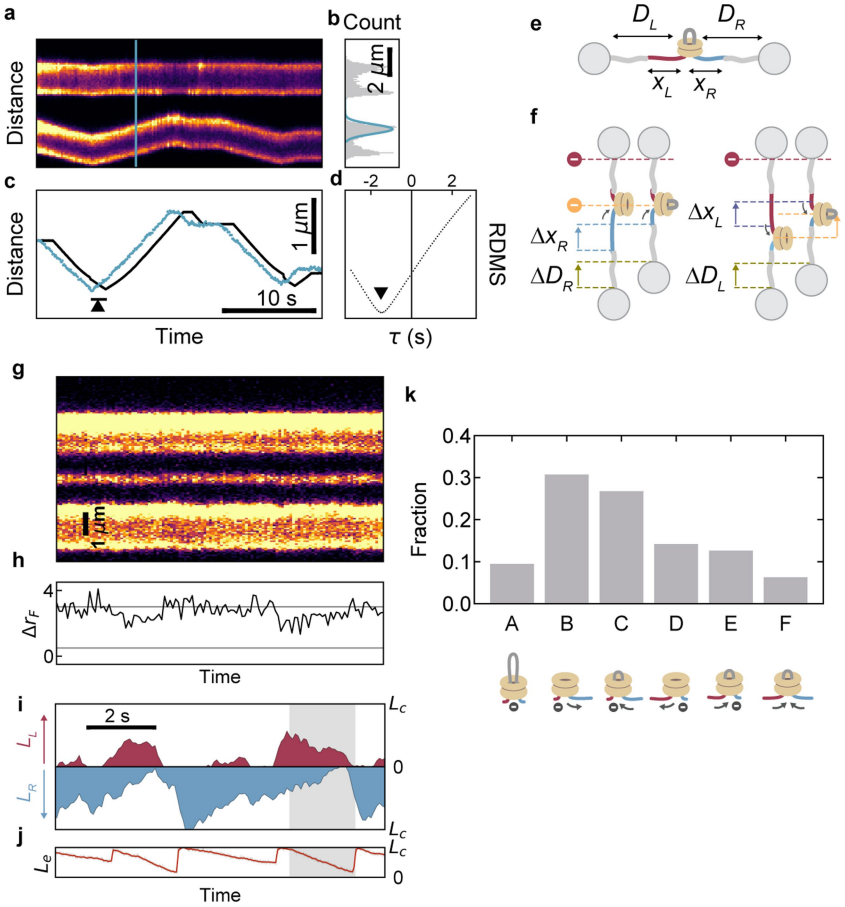
Extended Data Fig. 2 | Translocation by ClpB variants. a, ClpB monomer structure indicating all tested variants. These variants (except K467C) were generated in the constitutively active Y503D background. Variants E279A and E678A are Walker B mutants in the nucleotide-binding domains NBD1 and NBD2, respectively. These mutations abolish ATP hydrolysis at NBD1 or NBD2. Variants Y251A and Y653A are pore-loop mutants in NBD1 and NBD2, respectively. These mutations affect substrate interaction in the ClpB pore at either NBD1 or NBD2. The K476C variant undocks the middle domain (MD), mimicking the effect of Hsp70 (DnaK) activation. MD undocking in the Y503D variant is more pronounced, and therefore activation is more robust. An additional construct (ClpB(ΔN)) lacked the N-terminal domain (NTD), hindering initial substrate binding. Finally, the variant E731C harbours a cysteine at the bottom of NBD2 for fluorophore labelling. **b**, Fraction of time showing activity (f_A) for different mutants (in Y503D background, except K476C and wild type (WT)). **c**, Average translocation speed for all ClpB variants tested. KJE is the DnaK system (DnaK, DnaJ and GrpE). The median is displayed as a horizontal line within the box, and the mean as a white square. Whiskers indicate the lowest datum still within 1.5 interquartile range (IQR) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile. Sample sizes: $n = 1{,}139$ (Y503D), $n = 24$ (K476C), $n = 7$ (wild type) runs. **d**, Translocation example for ClpB(K476C). Scale bars correspond to 200 aa and 10 s. **e**, Translocation example for wild-type ClpB with the DnaK system (DnaK, DnaJ and GrpE). Scale bars correspond to 200 aa and 5 s. **f**, **g**, Absolute ATPase rate (**f**) and ATPase substrate-stimulation (**g**) for the three ClpB variants and different substrate conditions (mean ± s.d.). ATPase activity is higher and more strongly stimulated for Y503D, followed by K476C and wild type. The lower activities observed in the presence of denatured MBP–DM with respect to casein may reflect lower affinity and lower concentrations due to aggregation. The ATPase activity assay was repeated three times for all conditions in **f** and **g**, except for K476C, WT + MBP_2 and Y503D + casein, for which it was repeated two times.



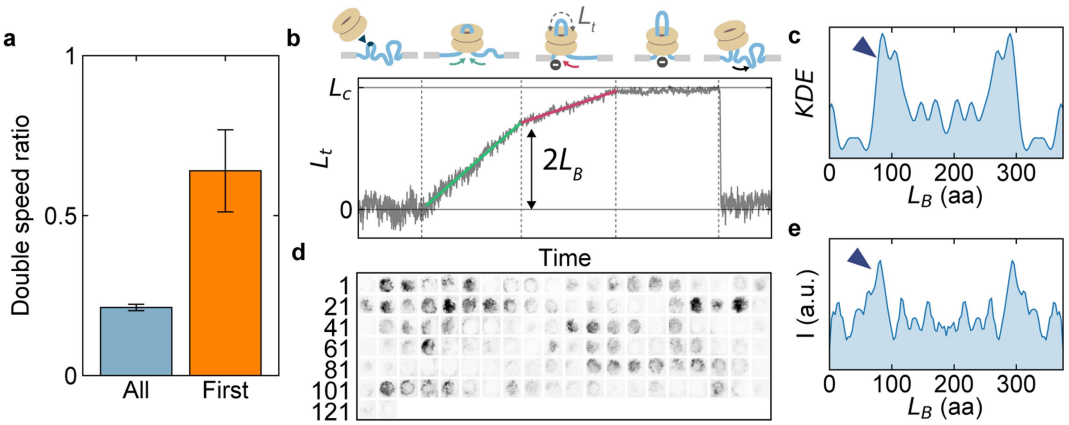
Extended Data Fig. 3 | Translocation runs for different constructs and molecules. Traces of protein extended length contractions in the presence of ClpB(Y503D) and ATP. **a**, MBP (L_c = 360 aa). **b**, 2MBP (L_c = 720 aa). **c**, 4MBP (L_c = 1440 aa). **d**, Speed distribution of translocation runs for the three different constructs (number of runs: n = 213 (MBP), n = 287 (2MBP), n = 306(4MBP)). All show a similar range of speeds, as expected, with one main peak (at $v \approx 240$ aa s^{-1}) and a second peak or shoulder at twice the magnitude ($2v$). A slight change in the ratio is observed between the two peak heights, with $2v$ becoming more pronounced in the longer constructs. This difference could reflect that distances between the initial ClpB binding site and the arresting DNA handles is then larger, and hence double-arm translocation more likely (see also Extended Data Fig. 4). **e**, Translocation speed distributions from three different substrate molecules (number of runs: n_1 = 218 (purple), n_2 = 102 (orange) and n_3 = 114 (green)), which show no significant variability between individual substrates. **f**, Translocation speed distributions for three different translocation bursts, which show continuous run–slip–run activity, and are thus surmised to reflect the action of individual ClpB hexamers (number runs: n_1 = 25 (purple), n_2 = 26 (orange) and n_3 = 49 (green)). Distributions are for ClpB(Y503D) and ATP, at approximately 8 pN. The data indicate no significant variability in the translocation activity between ClpB hexamers. The burst duration varied between 5 and 80 s, whereas the time between bursts varied between 5 and 150 s, for the 2MBP construct. **g**, Example translocation run of MBP showing the definitions of run length and run duration. Run duration is calculated as the time from the start of a run until the next back-slipping event, including the pause after translocation and before the next back-slip. Run length is calculated as the length difference between the start of a run and the next back-slipping event. **h**, **i**, Run length (**h**) and run duration (**i**) (see **g**) distribution for constructs of different lengths. Notably, the run duration distributions are similar for the constructs of different length, which suggests that the moment ClpB loses grip on one of the arms and causes the back-slip is determined by events that are intrinsic to the ClpB hexamer, and do not depend on the substrate nor the encounter with blockades (such as the DNA tether). This would make functional sense in the physiological context, as ClpB can then continue to push in an attempt to disrupt aggregated structures. By contrast, the switch between double- and single-arm translocation is directly triggered by such blockades, though without losing grip on either of the two arms.



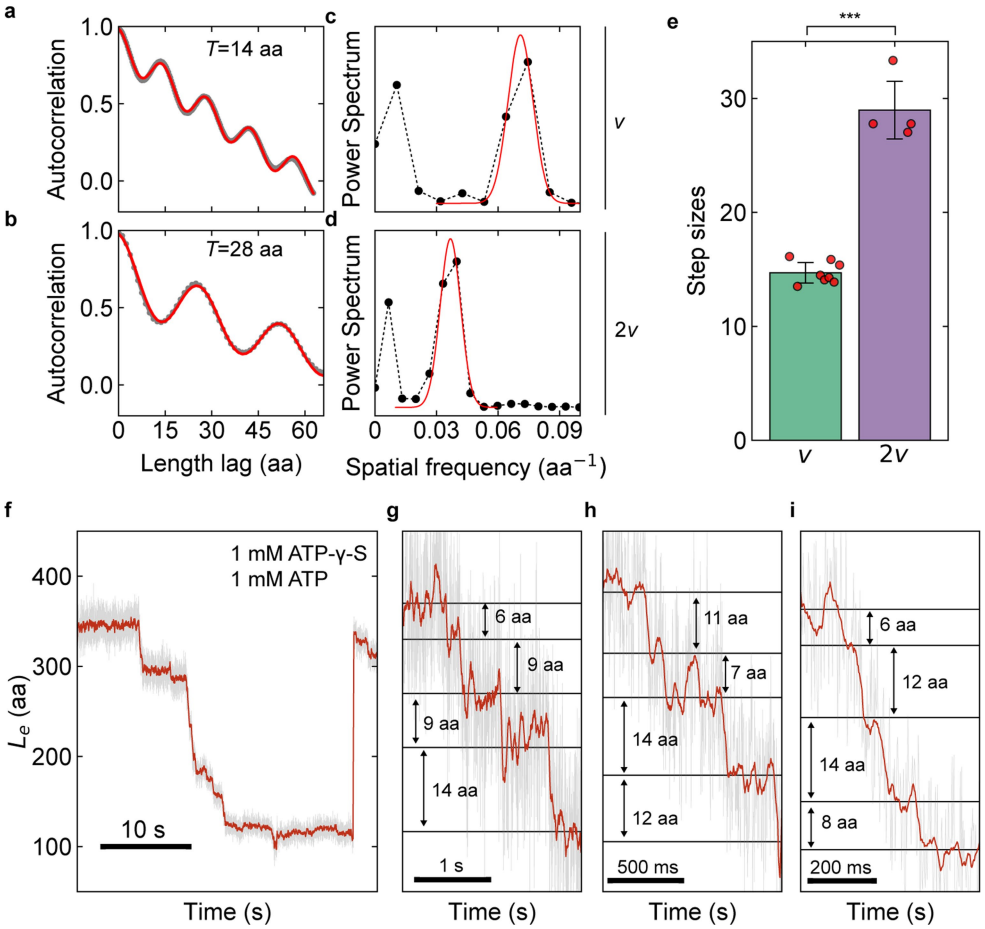
Extended Data Fig. 4 | Speed characterization of translocation runs. a, Translocated length (L_t) during threading of 2MBP by ClpB(Y503D). Raw data (light grey, 500 Hz) is filtered using a Savitzky–Golay filter (black line). **b**, Local translocation speed calculated as the time derivative of the translocated length after Savitzky–Golay filtering. Negative slopes below −50 nm s^{-1} (horizontal line) are considered back-slipping events (orange areas, also in **a**) and help in determining isolated translocation runs. **c**, Identification of different speeds within a single translocation run. Linear fits are used to calculate the speed of the run (green and magenta lines), most times revealing two main velocities, one double that of the other. **d**, Time derivative of the filtered translocated length for a single run, with solid black lines indicating the threshold speeds to distinguish no translocation from single- and double-translocation velocities and green and magenta indicating the fitted velocity values (also shown in **c**).



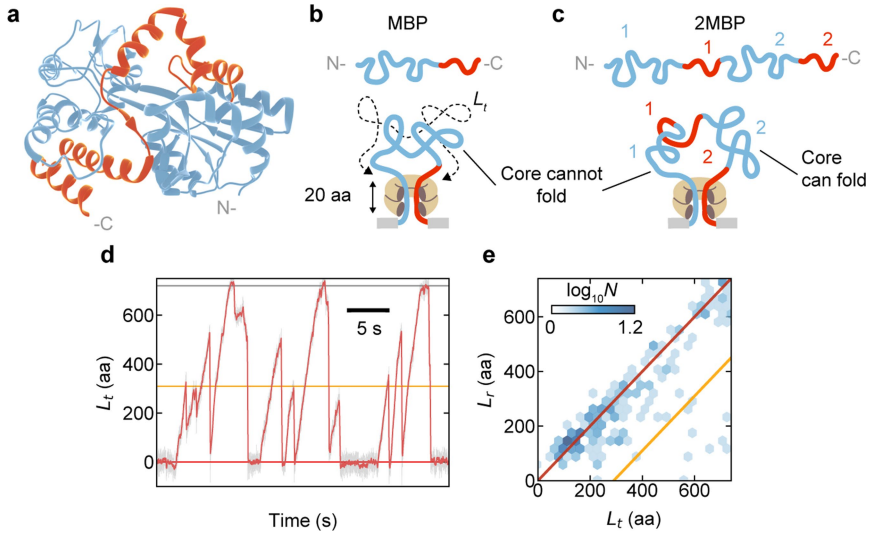
Extended Data Fig. 5 | Integrated tweezers and fluorescence particle tracking method. a–**d**, Synchronization of fluorescence and tweezers signals. **a**, Confocal scanning kymograph of two trapped beads. **b**, Intensity profile of a scanning line (blue in **a**), with a Gaussian fit of the edge of the moving bead (bottom) in blue. **c**, Offset between the fluorescence detection of bead movement as shown in **b** (blue dots), and high-resolution tweezers signal of trap and bead movement (black line) signals. **d**, Root mean square deviation (r.m.s.d.) between the signals for different time shifts τ. The minimum is marked with a triangle and represents the best estimation of the offset between the signals. **e**, **f**, Force clamp and computation of the two length components. **e**, Scheme of the lengths involved. D_L and D_R, distances between beads and ClpB; x_L and x_R, distances between protein termini and ClpB. Note that these distances are not contour lengths. **f**, Bead and ClpB position changes for left-arm (left) and right-arm (right) translocation. **g**, Kymograph underlying data in Fig. 2i. **h**, Corresponding tracked position of ClpB. Horizontal lines indicate extreme ClpB positions. Top line, ClpB is positioned at the left-hand terminus (see **e** and **f**). Bottom line, ClpB at the right-hand terminus; no polypeptide is translocated (the complete polypeptide is thus on the *cis* side of ClpB). Deviations from the top line consistently occur at back-slip moments detected by the tweezers (**j**; see the two shorter back-slips), which shows that the left arm (red) back-slips. Some back-slips detected by the tweezers do not show a corresponding ClpB movement, which is expected because right-arm back-slips should not change the ClpB position. **i**, Corresponding lengths of left arm (red) and right arm (blue) against time, as determined from fluorescence tracking (**g**, **h**) and tweezers (**j**) data. **j**, Corresponding tweezers data showing the distance between termini (contour length of *cis* segments). **k**, Distribution of the different translocation and back-slipping events observed in the fluorescence experiments (number of events $n = 127$, 5 molecules).



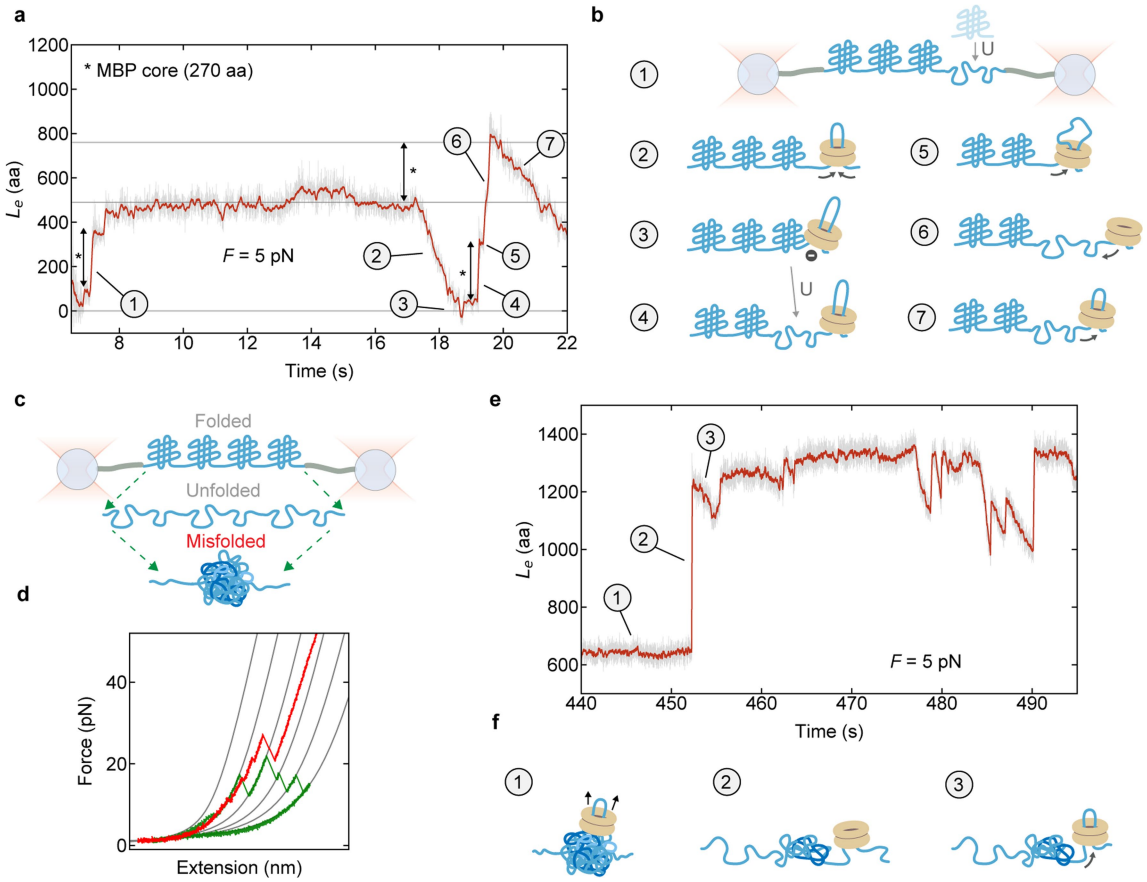
Extended Data Fig. 6 | Initial ClpB binding site estimation. a, Fraction of runs showing double speed when considering all runs ($n = 1{,}704$) and first runs only ($n = 30$). Data are mean ± standard error of a binomial distribution (see Methods). **b**, Example of first translocation run. ClpB binds at a certain location on the polypeptide, starts translocating both strands yielding the double speed (green) until it encounters the closest terminus, when it switches to single strand translocation (red). At the switch, the length translocated thus equals the distance between the initial binding site and the closest terminus (L_B), but times two because ClpB also translocated the other arm. Afterwards, the second terminus is also reached, and translocation stalls and a back-slip occurs, although this is not relevant here. **c**, Kernel density estimation (KDE) distribution of the inferred binding locations based on first runs, as described in **b** ($n = 30$). The distribution is symmetric because N and C termini cannot be distinguished. **d**, Peptide library data indicating regions of MBP that are bound by ClpB(NTD). **e**, Spot intensities were quantified using a custom script in Python. For direct comparison with **c**, the spot intensity distribution was also mirrored.



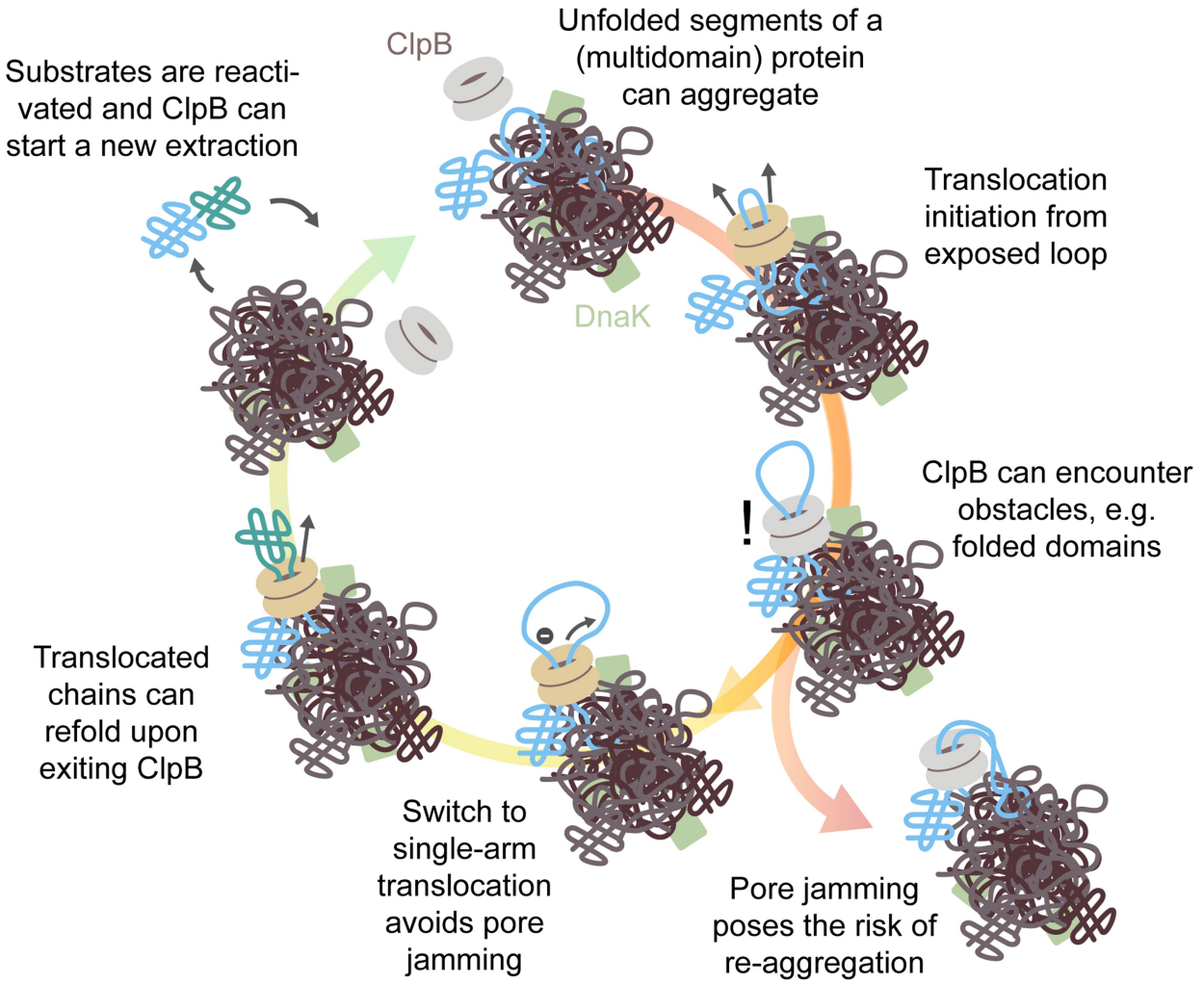
Extended Data Fig. 7 | Single translocation steps by ClpB. a–**d**, Analysis of step periodicity, related to Fig. 3. **a**, **b**, Autocorrelation of the pairwise length distribution for single-speed (**a**) and double-speed (**b**) runs from Fig. 3 (black dots). The red line is a fit, yielding period values of 14 and 28 aa, respectively. **c**, **d**, Power spectrum analysis of the pairwise length distribution for **a** (**c**) and for **b** (**d**), showing a peak that fitted to a Gaussian distribution (red) yields 0.071 and 0.037 aa^{-1}, respectively. This translates to 14 and 27 aa steps, in excellent agreement with the values obtained from the autocorrelation. **e**, The average step size is 14.6 ± 0.9 aa for single-speed translocation and 29 ± 3 aa for double-speed translocation (mean ± s.e.m., n_s = 8 and n_d = 4, 4 molecules), and statistically different ($P = 10^{-7}$; two-sided t-test). **f**, Example run in the presence of ATP–ATPγS mixture (1 mM each). Longer pauses are observed during translocation because ATPγS is hydrolysed much more slowly than ATP, and therefore can result in stalling. The prolonged stalling seen here is in line with a sequential ATP-hydrolysis along ClpB subunits. **g**–**i**, Notably, in these conditions, step-sizes smaller than 14 aa are now observed. These findings provide further support for the 14-aa steps being produced by the rapid consecutive action of multiple or all 6 ClpB subunits, whose individual 2-aa sub-steps would remain unresolved. After starting, a hydrolysis sequence moving along the ClpB hexamer would then arrest prematurely when encountering a ATPγS-bound subunit, and hence yield a smaller step size.



Extended Data Fig. 8 | Trans-refolding does not occur in single MBP and a mutant 2MBP construct. a, Structure of MBP (PDB ID: 2MV0), showing the C-terminal helices (red; around 90 residues) not required for core folding[28] (blue). **b**, Cartoon representation of the extended MBP chain showing the C terminal domain in red. After translocation arrest at the termini, segments at the N- and C termini (approximately 20 aa each) remain stuck inside the ClpB pore, and are thus not available for folding. Whereas the C-terminal segment (red) is not crucial for core formation, the N-terminal segment (blue) is. Thus, trans-refolding of single-MBP is not expected and indeed not observed. **c**, Cartoon representation of the extended 2MBP. The second MBP core (blue, 2) can fold in *trans*, since it now can translocate fully. **d**, Translocation run-and-slip activity for a tandem repeat of double mutant MBP (2MBP(DM)), which is compromised in refolding. Grey line indicates 720 aa, red line corresponds to 0 aa and the orange line corresponds to 310 aa, the length of one MBP core plus the two approximately 20-aa segments inside the pore. Back-slipping arrests at the orange line, as seen for 2MBP (Fig. 4), are no longer observed. **e**, Corresponding length distribution. Upon slipping, the released length (L_r) is now typically equal to the previously translocated length (blue data follows red line, $n = 203$ runs, 6 molecules).



Extended Data Fig. 9 | Disruption of folded and aggregated structures by ClpB. a, Extension length (L_e) of the 4MBP construct plotted against time in the presence of ClpB(Y503D) and ATP. **b**, Cartoons of event sequence suggested in **a**. (1) One MBP core is unfolded by increasing the force, immediately followed by relaxation to 5 pN to avoid unfolding other MBP cores. Some C-terminal helices also unfolded in this process. (2) After a waiting period, ClpB binds the unfolded part and translocates it completely. (3) ClpB reaches the neighbouring folded MBP domain (and the DNA tether), and hence no longer changes L_e. (4) After a short pause, L_e increases in a discrete step of 270 aa, indicating the unfolding of one MBP core, which has precisely that length. (5) ClpB(Y503D) translocates briefly immediately afterwards, further indicating the bound ClpB, and (6) back-slipping occurs. Note that the length of the unfolded chain has increased by 270 aa, the length of one MBP core, as expected (star). (7) Translocation continues. **c**, **d**, To create a misfolded or aggregated state, the 4MBP construct was unfolded and rapidly relaxed (green trace). This sometimes produced non-native structures characterized by being compact and highly resistant to force (red trace). The tether was then relaxed to low force. **e**, Subsequent measurement of extension length against time. **f**, Cartoons of event sequence suggested in **e**: (1) the length remains unchanged, for example, owing to waiting for ClpB binding. (2) The length increases abruptly by about 600 aa, which is more than one MBP core (270 aa), suggesting that ClpB disrupted a non-native (aggregated) structure that contained more than one MBP repeat. (3) ClpB translocation is observed immediately afterwards. This is consistent with the model, because one-step disruption of structures by ClpB (pushing) action can yield unfolded polypeptide segments directly on the *cis* side of ClpB that are then available for translocation. Note that polypeptide may also be liberated on the other side of the misfolded structure, which is not immediately available for translocation. Subsequently, further translocation and slipping behaviour is observed. Note that the structure becomes almost fully disrupted, as it nears the maximum length of 1,440 aa.



Extended Data Fig. 10 | Loop extrusion as a disaggregation principle. Insertion and translocation of loops promotes efficient disaggregation, because aggregates may display few accessible polypeptide termini at the surface. Translocation by Hsp100s of polypeptides entangled in aggregates generates pulling forces that promote their dissociation, cooperative disruption of larger structures, and extraction. The ability of Hsp100s to switch between translocation modes is relevant to prevent pore jamming when encountering structures that resist immediate disruption. To dissolve such resistive structures and larger aggregates, many translocation actions are probably required, involving multiple Hsp100 hexamers and other chaperones such as Hsp70, acting at different moments in time and at different locations within the aggregate. The random non-processive action of Hsp70s probably inherently requires multiple Hsp70s working together, in a manner that does not generate large pulling forces, while exploiting rapid binding and unbinding. In contrast, the processive nature of ClpB translocation enables fast, deterministic, and forced dissociation, which further limits re-aggregation and degradation when in combination with rapid refolding.

nature research

Corresponding author(s): Sander J. Tans

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Data from the Lumicks C-trap set up was acquired using the software provided by Lumicks: Tweez-O-Matic version 36.0, running under LabView 11.0. Fluorescence kymographs were acquired with Lumicks software Scanary 3.4
Data analysis	All analysis was performed using a custom made package in Python 3.5, available online upon request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. Sample sizes were chosen based on previous experience and published studies to assess reproducibility. Experiments were repeated multiple times on multiple substrate molecules, which were sufficient to obtain the described error margins.
Data exclusions	Translocation runs whose linear fits yielded r values below 0.8 were not used for translocation speed determination unless otherwise stated. This criteria was established after observing that some data contained more noise and hence were not suitable for such analysis.
Replication	All experiments were replicated multiple times, using different bead pairs and substrate molecules. All attempts at replication were successful
Randomization	Samples were not randomized in the experiments. Randomization was not applicable as samples were allocated according to different conditions such as buffer conditions.
Blinding	Experiments were not blinded as the data acquisition and analysis were done in different conditions.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Roche Diagnostics Germany, Anti-digoxigenin bodies, Cat. No. 11333089001, polyclonal antibody from sheep
Validation	The polyclonal antibody from sheep is specific to digoxigenin and digoxin and shows no cross-reactivity with other steroids, such as human estrogens and androgens .