Cotranslocational processing of the protein substrate calmodulin by an AAA+ unfoldase occurs via unfolding and refolding intermediates

Rafal Augustyniak[a,b,c,1] and Lewis E. Kay[a,b,c,d,1]

[a]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada; [b]Department of Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada; [c]Department of Chemistry, University of Toronto, Toronto, ON M5S 1A8, Canada; and [d]Program in Molecular Medicine, Hospital for Sick Children, Toronto, ON M5G 1X8, Canada

Edited by David Baker, University of Washington, Seattle, WA, and approved April 10, 2018 (received for review December 14, 2017)

Protein remodeling by AAA+ enzymes is central for maintaining proteostasis in a living cell. However, a detailed structural description of how this is accomplished at the level of the substrate molecules that are acted upon is lacking. Here, we combine chemical cross-linking and methyl transverse relaxation-optimized NMR spectroscopy to study, at atomic resolution, the stepwise unfolding and subsequent refolding of the two-domain substrate calmodulin by the VAT AAA+ unfoldase from *Thermoplasma acidophilum*. By engineering intermolecular disulphide bridges between the substrate and VAT we trap the substrate at different stages of translocation, allowing structural studies throughout the translocation process. Our results show that VAT initiates substrate translocation by pulling on intrinsically unstructured N or C termini of substrate molecules without showing specificity for a particular amino acid sequence. Although the B1 domain of protein G is shown to unfold cooperatively, translocation of calmodulin leads to the formation of intermediates, and these differ on an individual domain level in a manner that depends on whether pulling is from the N or C terminus. The approach presented generates an atomic resolution picture of substrate unfolding and subsequent refolding by unfoldases that can be quite different from results obtained via in vitro denaturation experiments.

VAT | protein unfolding | protein translocation | AAA+ ATPase | methyl-TROSY NMR

Protein unfolding is critically important for cellular proteostasis and is a requisite step in a wide variety of different processes that include mitochondrial import, disaggregation of misfolded species, and proteasomal degradation (1). Its importance is underscored by the fact that cells have dedicated ring-like oligomeric enzymes, members of the AAA+ family (ATPases associated with a variety of cellular activities), to carry out this function (2). Unfolding is achieved using energy from ATP hydrolysis whereby unfoldases apply a mechanical force that threads the substrate through a nanometer-sized central pore in the core of the molecule (3–5). Some AAA+ chaperones, such as ClpA, ClpX, or HslU, collaborate with barrel-shaped proteases so that the central pores of the two molecules align in a manner whereby substrates are directly shuttled into the protease lumen for degradation as they are unfolded (6–8). The most prominent example is provided by the 26S proteasome, a 2.5-MDa complex that degrades ubiquitinated substrates in eukaryotic cells (9). Here a hexameric ring formed by the Rpt1–6 subunits of the 19S regulatory particle, with both ATPase and unfoldase activity, binds to the 20S core particle (CP), a threonine protease (10, 11), and translocates the unfolded substrate polypeptide chain into the proteasome where it is degraded by active-site residues that would otherwise be sequestered from substrates. A similar unfoldase–20S CP interaction has also been characterized in the thermophilic archaeon *Thermoplasma acidophilum*, where the unfoldase, VAT, is a 500-kDa AAA+ chaperone (12–14) and a close homolog of the eukaryotic protein VCP/Cdc48/p97, a multifunctional enzyme involved in cellular processes that include apoptosis, membrane fusion, DNA repair, and the disassembly of protein complexes (15). In contrast to the unfoldases mentioned above whose function is coupled to the binding of proteases, the AAA+ ClpB (in bacteria) or Hsp100 (in eukaryotes) enzymes act in the absence of a peptidase partner to disassemble protein aggregates whose presence would otherwise lead to cellular misfunction (16, 17).

In an effort to understand how the chemical energy of ATP is harnessed into the mechanical force needed to unfold substrates a number of detailed structural studies of unfoldases have been undertaken (18–21). Our laboratory has focused on VAT, due to its stability, high thermal tolerance, and ease of expression that makes it an ideal target for a range of biophysical studies, including NMR analyses (13, 22). VAT forms a homohexameric complex where each protomer is composed of a 745-residue polypeptide chain that, in turn, is divided into three domains. These include an N-terminal domain (NTD) that is not well conserved between different AAA+ proteins (23), followed by two homologous nucleotide binding domains, denoted D1 and D2 (Fig. 1*A*). The biological function of the NTD is currently unknown, although a truncated mutant of VAT (ΔN VAT) lacking the first 182 residues has higher than WT ATPase and unfoldase activities, suggesting a possible partial inhibitory role for this domain (12, 24). The highly conserved D1 and D2 domains play key roles in the coordination and hydrolysis of ATP and ultimately in substrate unfolding. Like in many other AAA+ unfoldases (25),

Significance

In the cell, proteins are continuously synthesized and degraded. Degradation as well as unraveling of aggregated proteins depends on the activity of a family of ATP-driven ring-shaped unfoldases that catalyze unfolding of protein substrates by threading them through central ring channels. Here we follow the unfolding, translocation, and refolding of a substrate as it passes through the central pore of VAT, an archetypal unfoldase, using a method based on chemical cross-linking and methyl transverse relaxation-optimized NMR spectroscopy. The approach yields insights about how molecular machines are able to unfold protein targets, showing that for multidomain proteins unfolding can proceed in a series of steps that vary depending on whether the substrate is unfolded from its N or C terminus.

Author contributions: R.A. and L.E.K. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

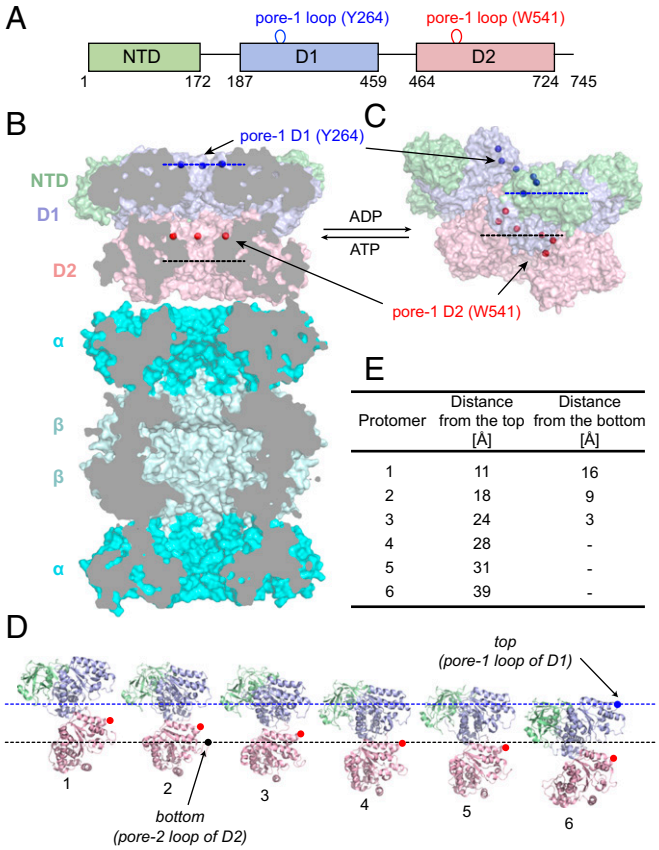
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: rafal@pound.med.utoronto.ca or kay@pound.med.utoronto.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1721811115/-/DCSupplemental.

Published online May 7, 2018.



E

Protomer	Distance from the top [Å]	Distance from the bottom [Å]
1	11	16
2	18	9
3	24	3
4	28	-
5	31	-
6	39	-

Fig. 1. The molecular architecture of the VAT unfoldase. (*A*) General domain organization and domain boundaries of VAT: NTD (green) followed by two AAA-type nucleotide-binding domains (D1 and D2, blue and pink). Locations of the flexible pore-1 loops as well as key residues involved in substrate processing are shown. (*B*) Cross-section of the ATP-bound state of VAT in complex with the 20S proteasome (cyan; arranged as α–β–β–α heptameric rings). Front subunits are removed from both VAT and the 20S CP to reveal internal chambers in the molecules. Blue and red dots are positioned at the Cβ atoms of Y264 and W541 (pore-1 loops); each set of dots can be connected by a plane. Dashed lines indicate the openings of the central channel (top and bottom) that are defined by the positions of pore-1 of D1 (Y264, top) and pore-2 of D2 (residue R576, bottom). The model of the VAT–proteasome complex has been obtained by aligning the two molecules [Protein Data Bank (PDB) ID codes 5g4g (22) and 1pma (70)] using PyMOL (71) and is not based on experimental data. (*C*) Surface representation of the ADP-bound state of VAT adopting the split-ring conformation [PDB ID code 5g4f (22)]. Pore-1 loop residues are organized into a helix and for some protomers become exposed outside the channel's lumen. The top and bottom of the VAT central channel are defined by dashed blue and black lines (discussed above), respectively. (*D*) Unwinding the ADP-bound state of VAT, with each protomer shown in the same orientation, highlighting W541 of the pore-1 loops of D2 (red). (*E*) Distances of W541 Cβ from the top and bottom for all of the protomers are tabulated.

flexible loops lining the interior of the central channel of VAT interact with substrates as they are threaded through the narrow protein core and point mutations in these loops can significantly affect unfoldase activity. Two sets of hydrophobic pore loops have been identified that are involved in polypeptide translocation (12), including pore-1 in D1 with a conserved aromatic residue at position 264 (Tyr264) and pore-1 in D2 with two neighboring residues, Trp541 and Val542. Interestingly, replacement of Tyr264 by Ala results in a complete loss of the unfoldase activity of VAT, whereas both residues of D2 pore-1 have to be mutated to achieve the same result (12).

A recently derived electron cryomicroscopy (cryo-EM)-based structure of VAT saturated with a slowly hydrolyzing variant of ATP, ATP-γS, and lacking NTD domains, provided a unique glimpse of how unfolding might occur (20). Images were obtained of VAT pulling the long unstructured C-terminal tail of a neighboring VAT molecule, showing that during the ATPase cycle the hexameric ring undergoes dramatic conformational rearrangements. Notably, density is observed for substrate in an extended conformation as it runs through the central pore of the protein, with 12–14 residues in contact with pore loops of five of the six VAT subunits. It is noteworthy that detailed structural features of the portion of substrate inside VAT could not be discerned because of its relatively low resolution, and little information was obtained for substrate regions outside of the unfoldase, either before entering or after emerging from the unfoldase pore, as these were at even lower resolution. Thus, although detailed pictures of unfoldases are emerging, providing important insights into how these molecules function, an atomic-level understanding of how protein substrates unfold during the course of the pulling reaction remains to be established.

In principle, NMR spectroscopy can be used as a tool to obtain a detailed picture of how substrate unfolding by unfoldases such as VAT occurs, as it provides information about both structure and dynamics in the solution state, often under close to physiological conditions. Although initially focused on small protein systems, the development of new hardware, and, importantly, transverse relaxation-optimized spectroscopy (TROSY)-based NMR experiments (26, 27) together with isotopic labeling strategies (28–30) have made it possible to study macromolecular machines (31, 32). Here, we use methyl-TROSY-based NMR methods to structurally characterize unfolding of single- and double-domain molecules, including the B1 domain of protein G (referred to as GB1 in what follows) and calmodulin (CaM), respectively, as they are threaded through the central channel of VAT and, in the case of CaM, subsequent refolding as the substrate emerges from the lumen of the unfoldase. This was achieved through the design of intermolecular disulphide bridges between pore-1 loops of the D2 domain and Cys residues inserted at different positions on the substrate so as to generate a series of NMR snapshots that, taken together, provide a detailed picture of unfolding, translocation, and refolding of substrate. Our cross-linking experiments clearly show that the channel of VAT can accommodate up to three copies of a substrate simultaneously and that unfolding can be initiated from either terminus of the target. GB1 unfolds cooperatively during the initial stages of translocation. However, the multidomain nature of CaM results in a more complex behavior with formation of distinct intermediates that depend on the pulling direction.

Results

The Molecular Architecture of VAT. A recent combined cryo-EM and NMR study of VAT in ATP and ADP loaded states highlights the importance of structural plasticity in the function of this enzyme (22). In the ATP state VAT exists in a stacked ring conformation whereby each hexameric ring consists of D1 or D2 subunits, as illustrated in Fig. 1*B*. Here a schematic of VAT interacting with the proteasome is shown, highlighting the central pores of both of these molecules that form the conduit for substrate transport, leading to the central cavity between the β rings of the proteasome where degradation occurs. In the ADP state the enzyme can assume an alternate, split-ring helical conformation (Fig. 1*C*), such that the D1 or D2 pore-1 loops are no longer planar (blue and red spheres in Fig. 1 *B* and *C*). This is illustrated in Fig. 1*D* , where the structure in Fig. 1*C* is unwound and the individual protomers displayed from 1 to 6 as shown. The blue and black dashed lines (also in Fig. 1*C*) denote the top and bottom of the split-ring structure, respectively, as defined by the positions of the Cβ atoms of Y264 of pore-1 from the D1 domain of subunit 6 (top) and of R576 of pore-2 from the D2 domain of subunit 2 (bottom). Beyond the top and bottom,

the channel widens considerably so that substrate extending either above or below would not be expected to be significantly constrained by VAT. Distances from the Cβ of W541 of each protomer to the top and bottom positions as defined above are tabulated in Fig. 1*E*. These distances are of relevance since, as described below, substrate molecules will be attached to position 541 via a disulphide linkage involving a select Cys residue on the substrate and W541C.

VAT Unfolds Substrates with Disordered N or C Termini. Before performing NMR studies of substrate–VAT complexes we initially focused on the substrate specificity of VAT. Protein unfolding is a tightly regulated process and unfoldases can identify their substrates by the presence of specific signals. In the case of the ClpXP and ClpAP protease complexes an 11-residue peptide, encoded by RNA and referred to as an ssrA tag, is appended to the C terminus of proteins designated for degradation (33). Although this tagging system is restricted to eubacteria, both VAT and a second unfoldase, PAN, also from archaea, were found to recognize this degron (12, 34). VAT was also shown to unfold proteins with unstructured C termini (12, 20). To understand better what the requirements for substrate recognition by VAT are, we designed several constructs of GB1 with either N- or C-terminal disordered tails and performed degradation experiments using the 20S proteasome CP that requires globular proteins to be unfolded for proteolysis (9). As expected, in the presence of VAT and the 20S CP, GB1-ssrA was efficiently degraded when ATP was included in a reaction buffer, but not when ATP was replaced by ADP, and proteolysis was much faster when unfolding was catalyzed by ΔN VAT compared with the full-length VAT (fl VAT) (Fig. 2*A* and Fig. S1). Notably, GB1 with a 24-residue N-terminal hexa-His tag followed by the tobacco etch virus protease cleavage site (referred to as His-tag in what follows) was also processed by both variants of VAT (Fig. 2*A*) and the same result was obtained for a C-terminal His-tag attached to GB1 or CaM (Fig. S1). When all six His residues were replaced by Ile or Glu (see Fig. S1 for the exact sequences), so that the tag was predominantly hydrophobic or charged, degradation rates estimated from SDS/PAGE gels decreased by ∼40% but were still substantial. The reaction slows significantly when a shorter, eight-residue tag was used and could not be detected for tags less than 6 aa (Fig. S1). Our results suggest that lack of any structure at one terminus is the critical requirement for VAT-induced unfolding. As a consequence, all of the substrates used in this study retained either an N- or C-terminal His-tag (24 residues) that was used to control the unfolding reaction (i.e., from either N to C or C to N, for tags on the N or C terminus, respectively). Notably, when either of the substrates was not tagged there was no unfolding by VAT as assayed by the approach in Fig. 2*A* (Fig. S1).

An Experimental Strategy for Studying Substrate Unfolding by VAT. It is widely accepted that ATP-driven conformational changes in unfoldases lead to substrate engagement, induce unfolding, and mediate translocation through the central channels of these enzymes (25, 35). Our goal in this set of studies is to use solution NMR spectroscopy to understand at an atomic level how substrates are unfolded during the pulling reaction and subsequently how they refold upon exiting from the lumen of the unfoldase. Unfortunately, it is not possible to follow the pulling reaction in real time and an alternative strategy must be found. Here we exploit the fact that flexible pore loops play a key role in substrate processing for many AAA+ enzymes, including for VAT studied here, with mutations in these loops almost always leading to deleterious effects on unfolding and translocation rates (12, 23). Sauer and coworkers (35) have used a chemical cross-linking strategy to create mixed disulphides between ssrA peptides and different loops of ClpX to establish that the pore loops physically interact with substrates during translocation. Here, we employ a

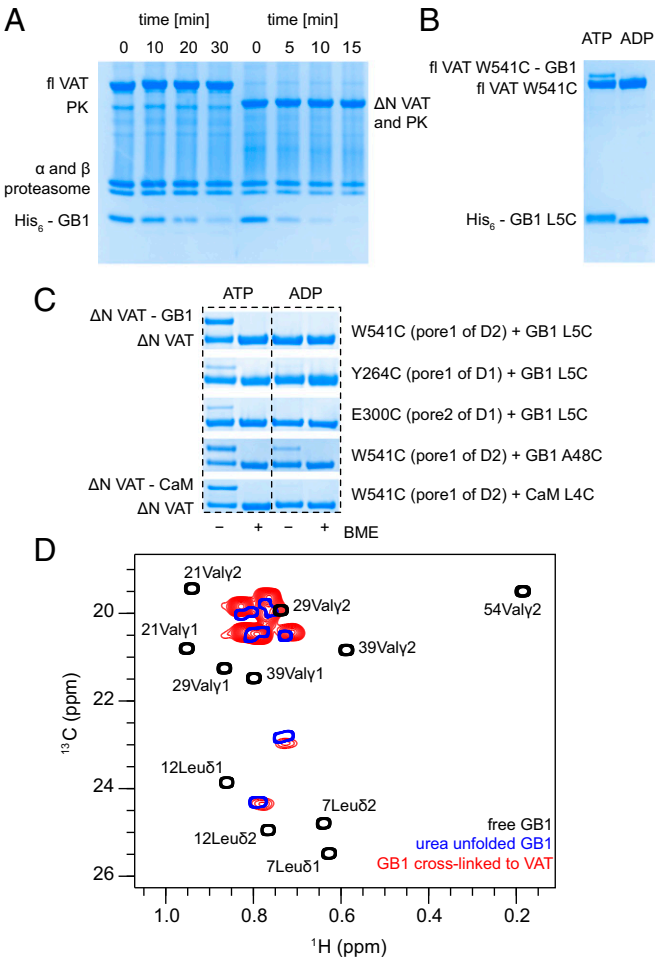


Fig. 2. His-tagged proteins are unfolded by VAT and degraded by the 20S CP. (*A*) SDS/PAGE gels showing degradation of the His-tagged GB1 domain (bottom) by the *T. acidophilum* 20S CP (α,β domains) in collaboration with fl VAT (left) or a truncated version of VAT lacking the NTD (ΔN VAT, right). An ATP regeneration system (involving pyruvate kinase, PK, that comigrates with ΔN VAT) was used. ΔN VAT-20S CP degrades GB1 faster than WT VAT-20S CP. (*B*) Cross-linking reaction between W541C mutant of fl VAT and L5C GB1 monitored by nonreducing SDS/PAGE. In the presence of ATP one can observe an additional band that corresponds to GB1 covalently attached to VAT via a disulphide bond (fl VAT W541C-GB1), as the cross-linked product is 9.4 kDa heavier than a VAT protomer (83 kDa). (*C*) Gel strips showing cross-linking reactions between various Cys mutants of ΔN VAT and His-tagged GB1 and CaM substrates. (*D*) Superposition of Leu/Val regions of ^{13}C-^{1}H HMQC spectra of ILVM-labeled His-tagged GB1 L5C in Hepes buffer, pH 7.4 (black), GB1 denatured with 8 M urea, pH 2 (blue), and ILVM-labeled His-tagged GB1 L5C cross-linked to perdeuterated ΔN VAT W541C (red). Assignments of the methyl groups of GB1 have been obtained from the Biological Magnetic Resonance Bank (BMRB) database (BMRB entry 7280).

similar strategy using folded substrates and VAT through the design of VAT mutants in which pore-loop residues were replaced by Cys. In considering the various pore loops for placement of single Cys residues it was clear that the pore-1 loop of D2 offers a number of advantages. First, a pair of neighboring residues must be replaced simultaneously to achieve significant inhibition of substrate translocation, while single mutations have little effect, unlike the case for the pore-1 loop of D1 where a substitution at position 264 is sufficient to eliminate function (12). Moreover, the pore-1 loop section of D2 is buried inside the channel of the enzyme so as to prevent nonspecific cross-linking in the absence of unfolding by VAT. This is in contrast to the situation for the pore-1 loop of D1 that is located in an

accessible region of structure on top of the hexameric ring of VAT (Fig. 1*B*). Hence, a W541C mutant of VAT was created after removing the two Cys residues present in the WT sequence (C77A/C679A) and used in all experiments. To prevent formation of disulphide bridges between protomers of the same hexameric ring, and to facilitate the reaction with substrates, pore-loop Cys residues were activated by formation of mixed disulphides with Ellman's reagent under denaturing conditions (*Materials and Methods* and Fig. S2). In this way, we have obtained hexameric rings of VAT with six activated cysteine residues within a central channel whose unfoldase activities were found to be similar to those of WT protein (Fig. S2).

The efficacy of cross-linking was established through a series of experiments in which activated VAT W541C was incubated with His-tagged GB1 containing the L5C mutation and the reaction followed by nonreducing SDS/PAGE. Initially, a full-length (fl) VAT construct was used (Fig. 2*B*). When ATP was included in the reaction buffer (Fig. 2*B*, left lane) an additional band was observed that corresponds to a single VAT protomer cross-linked to GB1. Importantly, cross-linking did not occur when ADP was substituted for ATP (Fig. 2*B*, right lane), confirming that cross-linking to VAT is coupled with substrate unfolding and translocation as would be expected since the side chain of residue 5 of GB1 is oriented toward the core of the protein (36), making it unavailable for cross-linking without unfolding. When the experiment was repeated using ΔN VAT W541C, instead of the full-length enzyme, the cross-linking efficiency was significantly higher, consistent with the higher ATPase and unfoldase activities of this truncated mutant (12) (Fig. 2*C*, top). Here, the two SDS/PAGE bands corresponding to cross-linked and substrate-free VAT have almost equal intensities, suggesting that within a hexameric ring on average three protomers can be attached to substrate; a similar situation is observed for the CaM substrate (Fig. 2*C*). We have tried to optimize reaction conditions (temperature, incubation time, and the amount of ATP) to check if more than three substrate molecules could bind but our attempts were unsuccessful. When Cys was inserted within the pore-1 loop of D1 (Y264C) or in the middle of the less-conserved pore-2 loop of D1 (E300C), the efficiency of substrate cross-linking was very poor.

Interestingly, GB1 A48C was able to cross-link to ΔN VAT W541C when either ADP or ATP was included in the buffer, although ATP-dependent cross-linking was more successful (Fig. 2*C*). Cross-linking when ADP is the nucleotide can be explained by the split-ring conformation of VAT (22) that renders the pore-1 D2 loops of protomers 4–6 accessible (Fig. 1 *C*–*E*) and by the fact that residue 48 is located within a flexible loop with a solvent-exposed side chain in GB1. This emphasizes the importance of selecting a sequestered position in the substrates for placement of a Cys residue for cross-linking; all mutants considered in this study were evaluated by ensuring that cross-linking did not occur without the addition of ATP, ensuring that the cross-linked product was always formed by passage of substrate through the VAT pore. In the studies that follow, as indicated above, ΔN VAT W541C was used throughout. In this manner, the NMR-derived snapshots (discussed below) provide pictures of substrate structure as the position of the mutated Cys residue (for cross-linking) passes the pore-1 D2 loop of one of the VAT protomers. It is important to emphasize that the ΔN VAT in the ADP state studied here has a split-ring structure (Fig. 1*C*) so that the pore-1 D2 loops of each protomer are not coplanar but rather trace a helical structure (red dots in Fig. 1*C*). As a result, the region of substrate that resides within the lumen of the barrel, as well as the sequence of substrate above and below the entry and exit points, respectively, varies depending on which protomer the substrate attaches. In principle, therefore, the snapshots obtained in the NMR studies described below are necessarily averages over substrates attached to different protomers; however, relatively small variations are expected for the different conjugation points (*SI Text*).

GB1 Unfolds Cooperatively at the Start of Translocation. GB1 is a small, highly stable, β–β–α–β–β domain comprising 56 residues (36). We have followed the above-described protocol to obtain an NMR sample of N-terminal His-tagged GB1 L5C cross-linked to ΔN VAT W541C. In what follows, all substrates are assumed to be N-terminal His-tagged, unless specifically indicated otherwise. Because of the size of the complex we have focused on methyl group probes using a labeling scheme in which VAT is perdeuterated, while the Ile (δ1) and Met methyl groups of GB1 are labeled as $^{13}CH_3$ and the Leu, Val methyls as $^{13}CH_3$, $^{12}CD_3$ (referred to as ILVM labeling in what follows) in an otherwise deuterated background (28, 37). The ^{13}C-^{1}H heteronuclear multiple-quantum coherence (HMQC) spectra that exploit a methyl-TROSY effect described previously (27) have been recorded to evaluate structural changes of substrate upon cross-linking to VAT. In this way, we were able to obtain high-quality NMR spectra of substrate using concentrations of the complex on the order of 30 μM, corresponding to 100 μM in total substrate (recall that there are three substrate molecules per complex, on average).

In addition to controlling for sample quality via the SDS/PAGE analysis described above (Fig. 2*C*, top), both with and without ATP, we have measured the translational diffusion of substrate-VAT using pulsed-field gradient (PFG) NMR (38) so as to ensure the integrity of the complex. Using a ^{13}C-edited experiment described previously (39), samples of ILVM-labeled ΔN VAT (380 kDa, no substrate), ILVM-labeled GB1 L5C cross-linked to deuterated ΔN VAT W541C (400 kDa), and free GB1 L5C (8 kDa) were examined. The diffusion constants obtained for ΔN VAT and the cross-linked complex were very similar ($2.42 \pm 0.15 \times 10^{-7}$ and $2.15 \pm 0.18 \times 10^{-7}$ cm^2/s, respectively, at 25 °C), indicating that the cross-linked sample is intact and forms hexamers. For the sake of comparison, free GB1 diffuses more rapidly ($14.21 \pm 0.07 \times 10^{-7}$ cm^2/s under the same conditions), as expected for a smaller protein. Theoretical values for ΔN VAT and free GB1 obtained with HYDROpro software (40) are 3.44×10^{-7} cm^2/s and 14.17×10^{-7} cm^2/s, respectively. We have performed similar PFG NMR experiments to monitor the stabilities of all of the cross-linked NMR samples (discussed below).

The Leu-Val region of the ^{13}C-^{1}H HMQC correlation map of an ILVM-labeled GB1 L5C cross-linked sample is shown in red (Fig. 2*D*). By means of reference, HMQC spectra of free GB1 L5C in the same buffer and of free GB1 L5C unfolded in 8 M urea, pH 2, are superimposed on the dataset (black and blue, respectively). Notably, the spectra of urea unfolded GB1 and GB1 cross-linked to VAT are very similar. Because chemical shifts are exquisitely sensitive reporters of protein structure (41, 42) it is clear that upon cross-linking the substrate loses both secondary and tertiary interactions, and its conformation resembles that of a random coil induced by chemical denaturation. To reach the activated cysteine located on the pore-1 loop of D2 a minimum of seven to eight residues from the GB1 sequence (i.e., after the His tag) must enter the cavity for cross-linking to protomer 1 and more for the other protomers (Fig. 1*E*), leading to the global unfolding of the protein.

Cooperative Unfolding of the N-Lobe Is the First Step in the Translocation of Calmodulin from the N Terminus. Although the initial experiments with GB1 were important for establishing the cross-linking protocol, the architecture of the stable GB1 domain almost guarantees that unfolding will proceed via a highly cooperative mechanism. We were interested in examining a more complex, multidomain protein in an effort to establish how unfolding by VAT and subsequent refolding would proceed in this case. We have chosen to use the calcium binding protein CaM from *Xenopus laevis* as a substrate. In its apo state CaM consists of two homologous and

highly independent domains, comprising the N- and C-lobes of the molecule that are separated by a flexible linker (43). Both lobes contain a pair of EF-hand motifs that are connected by a short loop (Fig. 3*A*). The choice of CaM was motivated, in addition, by the fact that it has been studied extensively by NMR so that chemical shift assignments of the protein are available (44, 45), by the fact that the primary sequence of the WT protein does not contain Cys residues, and by the fact that it can spontaneously refold from an unfolded state (46), allowing both unfolding and refolding to be studied, the latter as substrate exits the central pore of the unfoldase.

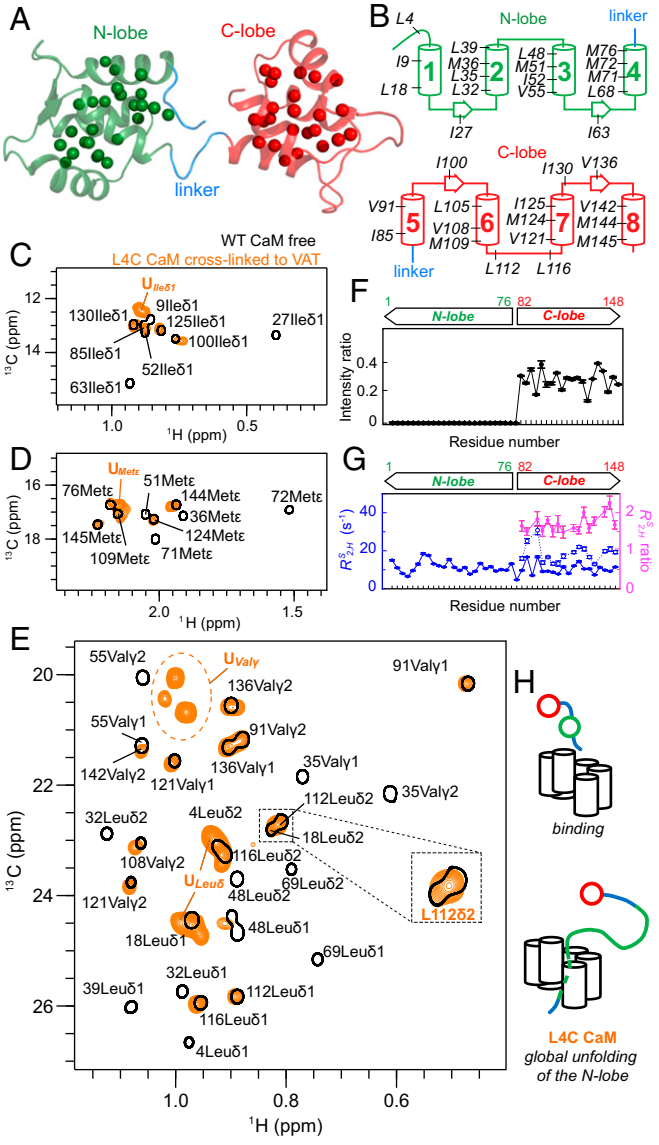


Fig. 3. Pulling by VAT at the N terminus of CaM results in global unfolding of the N-lobe. (*A* and *B*) Cartoon representation of the apo state of CaM [PDB ID code 1cfd (43)] with ILVM methyl groups shown as spheres (*A*) and highlighted on secondary structure elements (*B*). (*C*–*E*) Superposition of selected regions of ^{13}C-^{1}H HMQC datasets of ILVM-labeled WT apo-CaM bearing an N-terminal His-tag (black) and His-tagged L4C ILVM-CaM cross-linked to ^{2}H ΔN VAT W541C (orange). (*F*) NMR signal intensity ratios, I_b/I_f, where I_f and I_b are intensities of methyl cross-peaks of free and cross-linked CaM, respectively, showing that all of the signals from the N-lobe disappear. (*G*) Slow component of methyl ^{1}H transverse relaxation rates ($R^s_{2,H}$) measured on a sample of unbound (filled blue circles) and cross-linked L4C CaM (empty blue circles). $R^s_{2,H}$ ratios (cross-linked vs. free CaM) are shown in magenta. N-lobe methyl groups are not visible for the cross-linked sample so there are no relaxation data for this region. (*H*) Schematic illustration of the first step in the translocation of CaM that results in the cooperative unfolding of the complete N-domain.

Initial studies focused on an N-terminal His-tagged L4C mutant of apo-CaM, with the Cys inserted two residues before the beginning of the first EF-hand motif (Fig. 3*B*). This construct, as for all other cysteine mutants of CaM used in this study, is properly folded with similar thermal stability as the WT protein (Fig. S3). The cross-linking of L4C CaM and VAT was shown to be ATP-dependent (Fig. 2*C*) and the measured diffusion constant of 2.25 ± 0.16 × 10^{-7} cm^2/s was consistent with a tethered substrate. Fig. 3 *C*–*E* shows selected regions of a superposition of a pair of 2D ^{13}C-^{1}H HMQC spectra of ILVM-labeled L4C CaM cross-linked to VAT (orange) and free L4C apo-CaM (black). Notably, all of the peaks derived from the N-lobe in the free protein have disappeared in the complex (residues 1–76), while the peak positions of methyl groups from the C-lobe (residues 82–148) in the complex remained mostly unaffected, although their intensities were significantly attenuated, to 30–40% (to 60–70%, when peak volumes were used; Fig. 3*F*). Concomitantly several new peaks appeared in the spectrum of the complex with chemical shifts in the random coil region for methyl groups that superimpose well with peaks in a spectrum recorded on an 8 M urea unfolded N-domain construct (*SI Text* and Fig. S3). These peaks disappear upon addition of 1 mM TCEP, a compound that reduces disulphide bridges and signals from the N-lobe reappear (Fig. S3). Fig. 3*G* compares methyl transverse relaxation rates of slowly relaxing ^{1}H transitions, R^s_{2H} (47), of CaM in the free (filled circles) and bound (open) states, showing roughly a factor of two increase for the C-lobe of the complex (N-terminal peaks are not present in the complex, discussed above). Taken together, our results show that upon cross-linking at position 4 the entire N-lobe unfolds to form a long unstructured linker that connects the folded C-lobe to VAT (Fig. 3*H*).

The Core of the N-Lobe Begins to Refold Before It Completely Exits the Channel of VAT. To explore unfolding (and refolding) as different positions of CaM pass through the central channel of VAT we have generated a number of Cys mutants of N-terminal His-tagged ILVM-CaM, with the locations of the inserted Cys residues shown in Fig. 4*A*, and prepared complexes with perdeuterated ΔN VAT W541C. Note, as described above, that the presence of an unfolded N-terminal tag, coupled with a folded C terminus, ensures entry into VAT via the N-lobe. ^{13}C-^{1}H HMQC spectra were recorded for each complex and peak intensities quantified. As the linkage point on the substrate was moved from residue 4 to 15, 52, 71, and finally to 73, intensities of peaks from the C-lobe residues became progressively weaker (Fig. 4*B*), although all remained at positions corresponding to those for the folded C-domain and could thus be easily identified. The decrease in peak intensity is a direct result of increased penetration of the N-domain inside VAT so that the flexible unstructured polypeptide connecting VAT with the intact C-lobe becomes shorter in length (Fig. 3*H*), leading to its slower tumbling. More significant changes were noted for the methyl-containing residues of the N-lobe. The strong signals from the (only) two Val residues in the N-lobe (Val35 and Val55) in the unfolded region of the ^{13}C-^{1}H HMQC spectrum for the L4C construct significantly decrease in intensity for A15C-CaM and I52C-CaM and then regain intensity for the M71C-CaM complex (Fig. 4*C*), while no intensity for these peaks, or any others in the N-domain, is observed at the positions expected for the folded domain (Fig. 4*B*). The decrease in unfolded peak intensities for Val35 and Val55 in the A15C construct, in particular, is interesting as it is expected that both of these residues would be well outside the lumen of VAT (i.e., have not yet penetrated the unfoldase; *SI Text*). Signal intensities in NMR spectra can be very sensitive to conformational exchange processes that can lead to extensive

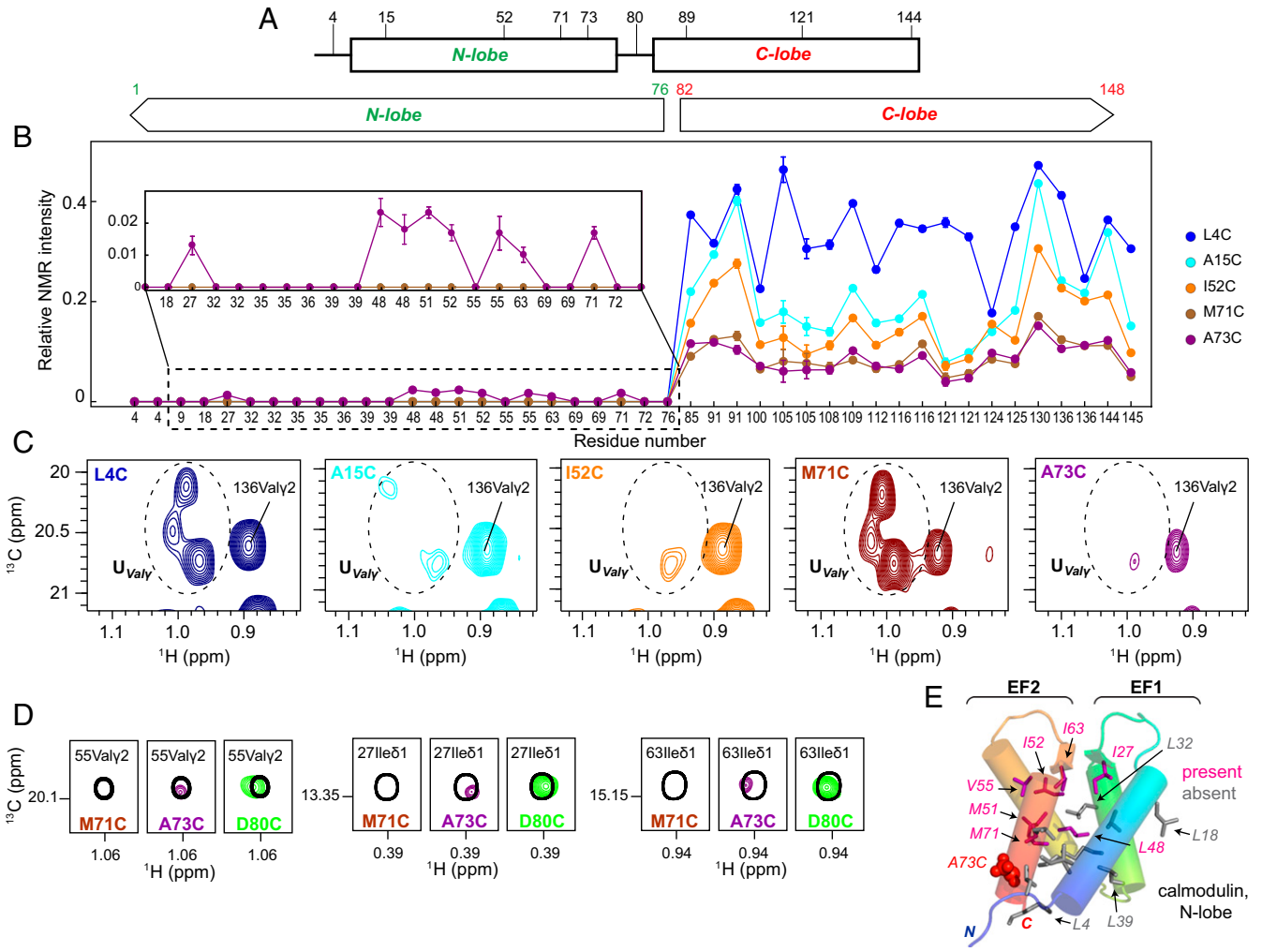


Fig. 4. The N-lobe of CaM traverses the central channel of VAT in an unfolded state and begins to acquire native-like structure before the domain is released. (*A*) Positions of single Cys point mutations introduced into CaM in this study. (*B*) Relative methyl intensities for different CaM Cys mutants cross-linked to VAT as a function of residue number. A sample of unbound WT CaM was used as a reference. (*C*) Intensity changes of Val residues from the unfolded N-lobe ($U_{Valγ}$) as a function of cross-linking position, indicated in the upper left-hand corner. (*D*) Partial refolding of the N-lobe observed for the cross-linked sample of CaM A73C. The reference spectrum of unbound WT CaM is shown in black. (*E*) Cartoon representation of the N-lobe of CaM with methyl-containing residues shown as sticks. ILVM residues of the N-lobe for which cross-peaks are observed at natively folded positions for CaM A73C-VAT are colored in pink; methyls from residues that are not yet visible in HMQC spectra are in gray. A73, that is replaced by Cys, is highlighted in red.

peak broadening. Our data suggest, therefore, that as the N-lobe increasingly penetrates the lumen of the unfoldase (i.e., as the point of contact between pore loop-1 of D2 and substrate increases from substrate residue 4 to 15 to 52) regions of the N terminus that are outside of VAT, and reported on by Val35 and 55, may interconvert between different conformers on microsecond–millisecond timescales. Once the majority, but not all, of the N-lobe has traversed the pore (point of attachment at position 71) the unfolded peaks for Val reappear, consistent with increased picosecond–nanosecond rapid dynamics and less sampling of conformations on a slower timescale.

Interestingly, when the attachment is at position 73 many of the peaks for the folded N-lobe begin to appear (Fig. 4*B*, *Inset* and 4*D*), albeit with low intensity, consistent with refolding of the N-domain, even while a portion of the fourth and final helix of the N-lobe resides within the VAT pore and must therefore not be natively formed. The low intensities of the folded peaks are consistent with an N-domain in contact with VAT, limiting its flexibility, and also likely indicate that this intermediate is only partially formed and thus structurally heterogeneous, leading to further decreases in peak intensities. Nevertheless, the fact that peaks corresponding to the folded state are observed for methyl groups from many of the residues that comprise the hydrophobic core of the N-lobe, including Ile27, Leu48δ1,δ2, Met51, Ile52, Val55γ2, Ile63, and Met71 (Fig. 4 *D* and *E*), where Ile27 and Ile63 form backbone hydrogen bonds that stabilize EF1 and EF2 of the domain (48), argues that the structure of this partially folded intermediate must, at least somewhat, resemble the N-lobe in its native state, albeit heterogeneous and lacking a native C-terminal helix (Table S1). Of note, probes on helix 1, Ile9 and Leu18, are not observed in spectra (Fig. 4*B*, *Inset*), perhaps reflecting the fact that the side chains of these residues point toward helix 4 in the native state (Ile9), with helix 1 not yet properly formed due to the absence of interactions with helix 4 that is at least partly unfolded. Attachment of substrate at position 80 leads to a spectrum where the N-domain intensities have increased significantly, suggesting that once position 80 passes the pore-1 loop of D2 a stable N-lobe is formed.

The Stepwise Unfolding of the C-Lobe. As the position of substrate attachment increases further (residues 80, 89, 121, and 144) the NMR signals derived from residues 1–76 become sharper and

more intense (Fig. 5 *A* and *B* and Fig. S4), reflecting the increased flexibility of the N-domain due to the growing linker. Upon attachment at position 80 the C-lobe begins to partially unfold. C-lobe methyl groups with little or no intensities at their folded state positions derive from residues on helix 5 (the four C-lobe helices are numbered 5–8, extending from the N terminus of this domain) or at other positions in the domain that interact with helix 5 when it is folded. These include Ile85 on helix 5, Leu105δ2 on helix 6 whose shift is affected by the position of Phe89 of helix 5, Val108γ1,γ2 on helix 6 that contacts Phe89 of helix 5, Leu112δ1 in contact with N terminus of helix 5 and residing in a loop connecting helices 6 and 7, and M145 of helix 8 that makes contacts with Ile85 of helix 5. In contrast, many other methyl resonances in the C-lobe remain at their folded-state positions, including Ile100 and Val136 (Fig. 5 *C* and *D*) that form a critical hydrogen bond that stabilizes the orientation of EF3 and EF4 (48), in much the same way as Ile27 and Ile63 of the N-lobe. Other residues reporting on native-like structure include Met124 that is a core residue from helix 7, as well as Met144 that points into the core of the domain (see Table S2 for a more detailed list). Taken together, this suggests that a C-domain intermediate is formed with unfolding of the first α-helix (helix 5) but with significant residual structure in other areas. It is noteworthy that homology modeling shows that position 80 of the C-lobe corresponds to position 4 of the N-domain and that cross-linking with L4C CaM led to the complete unfolding of the N-lobe (discussed above). By contrast, only the first helix of the C-lobe unfolds when substrate is tethered at position 80.

Moving the cross-link point closer toward the C terminus of CaM (positions 89, 121, and 144) resulted in a complete loss of signals from their positions in the folded C-lobe including Ile100 and Val136 (Fig. 5*C*) and the concomitant appearance of intense peaks in random-coil regions of the spectrum (Fig. 5 *A* and *B*), which suggests a transition toward the unfolded state. This is conformed in a comparison with a spectrum recorded of an 8 M urea unfolded CaM C-domain construct (*SI Text* and Fig. S3). Thus, while the N-domain is able to acquire native structure before it exits the lumen of the unfoldase (Fig. 4*E*), the C-domain cannot (Fig. 5*E*).

We have also studied unfolding/refolding starting from the C terminus, by preparing constructs of CaM with a C-terminal His-tag (Fig. S5). The cross-linking experiments in this case are in good agreement with expectations based on pulling at the N terminus, due to the inherent symmetry of the structure of CaM extending from either the N to C or C to N termini. A picture is thus obtained whereby unfolding/refolding in this case is essentially the same as that described in Figs. 3*H* and 5*E* with the N and C termini reversed. As discussed below, this results in an unfolding mechanism that on an individual domain level depends on the pulling direction.

Discussion

Protein unfolding by unfoldases is highly regulated, involving specialized ring-shaped enzymes that transform chemical energy

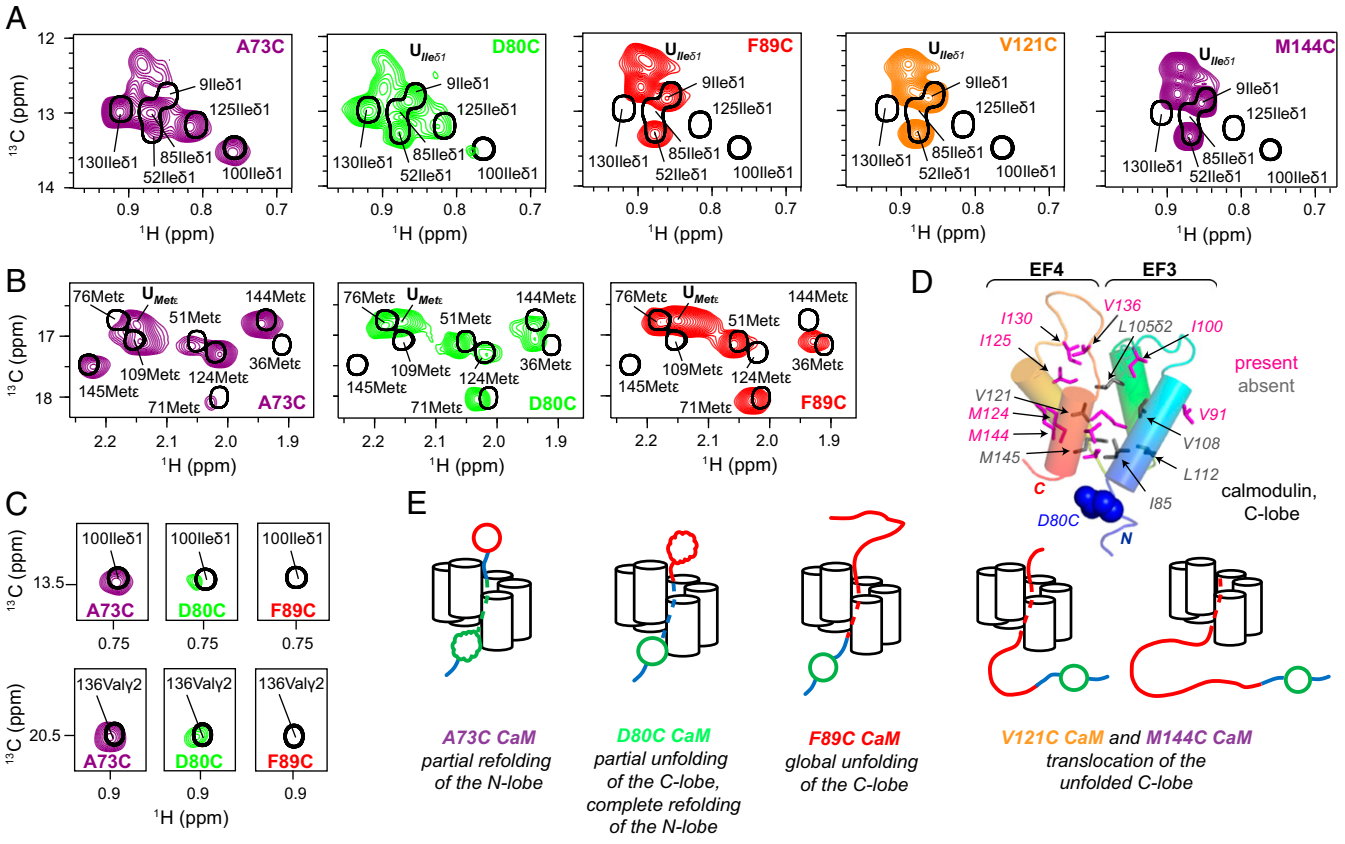


Fig. 5. The C-lobe of CaM unfolds sequentially and does not refold until substrate is released from VAT. Ile (*A*) and Met (*B*) regions of ^{13}C-^{1}H HMQC spectra of cross-linked Cys mutants of CaM (reference spectrum of unbound WT CaM is shown in black). Only some of the C-lobe "folded" peaks disappear for CaM D80C-VAT, while all of the peaks are gone for CaM F89C-VAT, with a concomitant increase in intensities of peaks located in the random-coil regions of the spectra, $U_{Ile\delta1}$ and $U_{Met\varepsilon}$. (*C*) Intensities of Ile100δ1 and Val136γ2 are highlighted, as these residues make critical interactions in the folded state. (*D*) Cartoon representation of the C-lobe of CaM with ILVM methyl-containing residues shown as sticks. Residues for which cross-peaks are observed at natively folded positions for CaM D80C-VAT are colored in pink; methyls from residues that are not visible are in gray. Blue spheres highlight D80. (*E*) Proposed model for unfolding and translocation of CaM by VAT.

into a series of power strokes that force substrates into their narrow central pores (3). Proteins designated for unfolding are recognized through the attachment of various tags such as the ssrA sequence in bacteria (33), small archaeal modifier proteins that are structural homologs of ubiquitin in archaea (49), ubiquitin in the case of coupled unfolding/degradation by the proteasome in eukaryotic cells (50), and small presequences of ∼10–30 aa or internal sequences in the case of transport into mitochondria (51). We show here that the AAA+ unfoldase VAT, which collaborates with an archaeal proteasome (24, 52), is able to process protein substrates so long as they contain a highly flexible and unstructured N- or C-terminal tail in a manner that appears to depend only slightly on its primary sequence (Fig. S1).

Here we have exploited the fact that substrates interact closely with pore loops in the lumen of unfoldases (25, 35), including VAT (20), to generate a series of substrate–VAT complexes with a disulphide-based cross-link connecting substrate molecules to W541C of the D2 pore-1 loop of VAT. Cross-linking depends on ATP hydrolysis so that substrates must translocate the central pore of the unfoldase until the single Cys of the target protein becomes proximal to position 541 of VAT, with formation of the disulphide and concomitant termination of the translocation processes. As many as three substrate molecules can associate with VAT at once (Fig. 2), similar to results from studies showing that multiple copies of a polypeptide chain can be degraded simultaneously by 26S rabbit reticulocyte lysate (53), from studies with ClpX showing that a pair of polypeptide chains can be unfolded simultaneously (54), and from recent degradation experiments using the ClpXP system and knotted proteins as substrates (55) establishing that tightened polypeptide knots can pass through the unfoldase.

A series of 2D ^{13}C-^{1}H HMQC spectra have been recorded of ILVM substrates attached at various positions to deuterated VAT, focusing on methyl-group probes of substrate structure and dynamics and taking advantage of a methyl-TROSY effect (27) that results in high spectral quality even in the case of applications to large protein complexes such as those studied here (31, 32). These spectra provide detailed insights into unfolding/refolding processes because the positions of methyl cross-peaks are sensitive indicators of structure that can be used to ascertain whether specific regions of sequence are unfolded or folded at various stages of the translocation process. As described above, all of the NMR experiments were carried out on samples with three substrates/VAT on average (Fig. 2*C*). Controls where only one substrate is attached on average lead to similar conclusions (Fig. S6). A first substrate, the small 56-residue GB1 domain, was used to establish the methodology and to ensure that attachment could be achieved in an ATP-dependent manner. Notably, HMQC spectra indicate that GB1 unfolds cooperatively when the attachment point is placed at the N terminus of the domain (position 5), resembling the globally unfolded state obtained by chemical denaturation.

The second substrate, apo-CaM, presents an interesting case because it is a two-domain protein with each domain composed of a pair of EF hands that are connected by a linker (43). A series of nine Cys mutants have been made to an N-terminal His-tagged CaM construct that span the entire molecule so as to generate snapshots of how unfolding occurs during translocation, followed by refolding when the substrate emerges from the other end of VAT. Formation of a substrate Cys4–VAT W541C disulphide link leads to complete unfolding of the N-domain (Fig. 3), as observed for GB1. At this point, ∼3–11 additional residues C-terminal to the Cys are sequestered inside the channel (depending on the attached protomer and assuming ∼4 Å between residues, Fig. 1*E*; see *SI Text*), unraveling the domain. When polypeptide comprising the first three helices in the folded protein emerges from the central pore of the unfoldase, corresponding to attachment at position 73, weak cross-peaks appear at some of the positions expected for the folded domain. This strongly supports the formation of an intermediate with some degree of native-like structure despite the fact that the C-terminal portion of helix 4 remains in the pore and hence is not natively formed. The weak intensity of the observed cross-peaks also suggests that the domain is conformationally heterogeneous; however, the signal-to-noise ratio in HMQC datasets (1–2% of free CaM) precludes recording relaxation experiments that could be used to verify this. Continued threading of the substrate polypeptide leads to proper refolding of the N-domain and formation of a second intermediate (attachment at position 80) whereby the first helix of the C-domain is sequestered in the channel, yet native-like structure persists, at least in part, for the remaining helices that have not yet traversed the pore. Attachment at position 89, at the C-terminal end of first helix of the C-lobe (helix 5), results in complete unfolding of the C-lobe likely by distorting the loop between EF3 and EF4 and thus interfering with a key interaction between Ile100 and Val136. Refolding of the C-lobe is only possible once substrate is released from the unfoldase. Fig. 6*A* summarizes the unfolding/refolding trajectory when substrate is translocated from the N terminus, focusing on key cross-linking positions.

It is interesting to note that, as discussed above, attachment of substrate at position 4, close to the N terminus, unfolds the N-domain completely, while cross-linking at the homologous position before the C-lobe (position 80) only unfolds the first helix of the C-terminal domain. The N-terminal helix of the N-lobe is both longer and richer in aromatic residues than the corresponding helix in the C-lobe and appears to play a more important role in domain stability than its C-lobe counterpart.

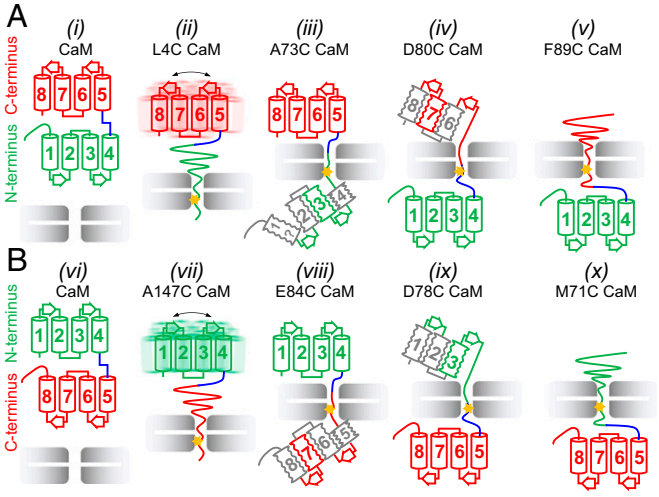


Fig. 6. Cooperativity of domain unfolding depends on pulling direction. (*A*) Schematic of key unfolding steps for a pulling reaction starting from the N terminus of CaM. The N-lobe unfolds cooperatively (*ii*, attachment at position 4 indicated by star) and begins to refold before helix 4 is completely released from the unfoldase (*iii*, attachment at 73). Unfolding of helix 5 (*iv*, attachment at 80) precedes complete unfolding of the C-lobe (*v*, attachment at 89) that only refolds upon release of substrate. (*B*) Schematic of unfolding starting from the C terminus of CaM. The C-lobe unfolds cooperatively (*vii*, attachment at 147), with formation of a native-like intermediate before helix 5 is completely released (*viii*, attachment at 84). Attachment at position 78 (*ix*) leads to unfolding of helix 4 before cooperative unfolding of the remaining helices (*x*). Note that in *A* the N-lobe unfolds cooperatively, with the C-lobe unfolding in steps and the reverse occurs in *B*. Helices are colored in green or red when there are many probes indicating a native-like element, in gray when the evidence is indirect, and in gray with the symbol "?" when there is no evidence. Squiggles indicate the likely presence of conformational heterogeneity. In *ii* and *vii* the overall dynamics of C- and N-domains are only moderately affected by attachment to VAT.

Molecular dynamics simulations of apo-CaM unfolding show that the initial helix of the C-lobe loses structure early in the unfolding trajectory while the first N-lobe helix persists much longer (56). Unfolding studies with a C-terminal His-tag CaM substrate that enters VAT in a C- to N-terminus direction show that, overall, unfolding and subsequent refolding proceed in an analogous fashion to what is observed when pulling is from the N-terminal end of the substrate (compare Fig. 6 *A* and *B*). Thus, analogous trajectories occur when translocation is from N to C or C to N, a consequence of the inherent symmetry of this substrate. This results in an unfolding mechanism on a domain level that clearly depends on the pulling direction: When a mechanical force is applied to the N terminus, the N-lobe of CaM unfolds cooperatively while the C-lobe unfolds in steps, whereas pulling from the C terminus results in a single-step unfolding transition for the C-domain with stepwise unfolding of the N-domain (Fig. 6 *A* and *B*). Similar conclusions about the dependence of unfolding on pulling direction have been made for thioredoxin threaded through the narrow pore of α-haemolysin (57). Interestingly, thioredoxin and both domains of CaM were shown to unfold according to a two-state mechanism with chemical denaturants or increased temperature (58, 59). The unique influence of unfoldases on substrate unfolding/refolding has parallels with what has been observed for nascent chain folding upon exiting the ribosome. In this case folding of polypeptides occurred only well beyond the exit tunnel (60), although the corresponding isolated proteins could fold spontaneously even when truncated. This suggests that ribosomes, as for the VAT unfoldase, modulate the folding process.

Our study of substrate translocation was based on equilibrium measurements using cross-linking that generates stable samples for NMR analysis. In contrast, protein unfolding by unfoldases is not an equilibrium process and the snapshots obtained here may, thus, only approximate the VAT unfolding process. It is noteworthy in this regard that studies of translocation by the ClpX and ClpA unfoldases show translocation rates of up to 70 aa per s for ClpX (61), which is a factor of two faster than for ClpA (62). In contrast, the folding rate constants for free apo-CaM were measured to be ∼13,000 s^{-1} and 1,400 s^{-1} for the N- and C-lobes, respectively (63). How CaM refolding rates are affected by proximity to VAT for regions outside of the central pore is not clear, but assuming that folding rates exceed rates of translocation it is likely that intermediate structures would form during translocation that are similar to those observed in our equilibrium studies.

A wide range of experiments has established that cotranslocational protein unfolding is not related to the global thermodynamic or kinetic stability of the substrate but rather depends on structural elements proximal to the pulling site (54, 57, 64, 65). For example, barnase is more resistant to unfolding than dihydrofolate reductase (DHFR) as assayed by traditional in vitro approaches, yet barnase is efficiently unfolded by the mitochondrial unfolding machinery while DHFR is not (66). An explanation is provided by the differences in N-terminal structures of these two proteins, where for barnase an N-terminal surface helix can be relatively easily unwound by pulling, while the N terminus for DHFR is more sequestered, forming a β-strand that is sandwiched between a pair of α-helices. Further, in the case of barnase, spontaneous global unfolding in vitro proceeds via a pathway in which central regions of structure unfold initially, in contrast to the case for unfolding via mechanical force applied at one of the termini (66).

Previous studies of CaM in combination with the results reported here reinforce the notion that protein unfolding via unfoldases can proceed via a different mechanism than spontaneous unfolding of isolated molecules. In contrast to conclusions based on experiments involving isolated apo-CaM in which individual domains unfolded in a two-state manner (67), our methyl-TROSY NMR studies of CaM–VAT complexes establish that unfolding and subsequent refolding proceed through intermediates. Symmetric pathways for pulling from the N or C termini are observed (Fig. 6), leading to distinct differences in unfolding properties of individual domains. An advantage of the NMR approach is that many probes are available for analysis, unlike other techniques that are limited to fluorescent substrates (68) or that lack atomic resolution (57, 69), facilitating a more detailed picture of protein translocation. In the past year cryo-EM applications involving a number of different AAA+ proteins have led to a significant increase in the understanding of how these molecular machines function (18–21). The smaller size of the substrate and its inherent dynamics during unfolding make it more difficult to obtain a detailed description of how it evolves during the translocation process. In this regard, a combination of cryo-EM and NMR spectroscopy offers a unique opportunity to generate detailed insights into both unfoldases and their substrate targets as translocation occurs.

Materials and Methods

Nearly all NMR experiments were performed on a Varian Inova 18.8 T spectrometer equipped with a pulse-field gradient triple-resonance room-temperature probe, 20 °C. Sample concentrations (ILVM-labeled substrate) were between 100–130 μM (except for VAT conjugated with a single substrate on average, rather than approximately three, where the concentration was 50 μM) corresponding to VAT concentrations of ∼40 μM dissolved in buffer containing 25 mM Hepes (pH 7.5 uncorrected), 2 mM ADP, and 50 mM NaCl in 99.9% D_2O for samples of free and cross-linked GB1 and 25 mM imidazole (pH 6.8 uncorrected), 2 mM ADP, 100 mM KCl, 1.5 mM EGTA, and 1.5 mM EDTA in 99.9% D_2O for samples of free and cross-linked CaM. Further details are given in *SI Materials and Methods*.

ACKNOWLEDGMENTS. This work was supported by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada. L.E.K. holds a Canada Research Chair in Biochemistry.

1. Matouschek A (2003) Protein unfolding–An important process in vivo? *Curr Opin Struct Biol* 13:98–109.
2. Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* 475:324–332.
3. Hanson PI, Whiteheart SW (2005) AAA+ proteins: Have engine, will work. *Nat Rev Mol Cell Biol* 6:519–529.
4. Lee C, Schwartz MP, Prakash S, Iwakura M, Matouschek A (2001) ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol Cell* 7:627–637.
5. Zolkiewski M (2006) A camel passes through the eye of a needle: Protein unfolding activity of Clp ATPases. *Mol Microbiol* 61:1094–1100.
6. Olivares AO, Nager AR, Iosefson O, Sauer RT, Baker TA (2014) Mechanochemical basis of protein degradation by a double-ring AAA+ machine. *Nat Struct Mol Biol* 21: 871–875.
7. Baker TA, Sauer RT (2012) ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim Biophys Acta* 1823:15–28.
8. Ramachandran R, Hartmann C, Song HK, Huber R, Bochtler M (2002) Functional interactions of HslV (ClpQ) with the ATPase HslU (ClpY). *Proc Natl Acad Sci USA* 99: 7396–7401.
9. Pickart CM, Cohen RE (2004) Proteasomes and their kin: Proteases in the machine age. *Nat Rev Mol Cell Biol* 5:177–187.
10. Nickell S, et al. (2009) Insights into the molecular architecture of the 26S proteasome. *Proc Natl Acad Sci USA* 106:11943–11947.
11. Kish-Trier E, Hill CP (2013) Structural biology of the proteasome. *Annu Rev Biophys* 42: 29–49.
12. Gerega A, et al. (2005) VAT, the thermoplasma homolog of mammalian p97/VCP, is an N domain-regulated protein unfoldase. *J Biol Chem* 280:42856–42862.
13. Golbik R, Lupas AN, Koretke KK, Baumeister W, Peters J (1999) The Janus face of the archaeal Cdc48/p97 homologue VAT: Protein folding versus unfolding. *Biol Chem* 380: 1049–1062.
14. Pamnani V, et al. (1997) Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon Thermoplasma acidophilum. *FEBS Lett* 404: 263–268.
15. Baek GH, et al. (2013) Cdc48: A swiss army knife of cell biology. *J Amino Acids* 2013: 183421.
16. Bukau B, Weissman J, Horwich A (2006) Molecular chaperones and protein quality control. *Cell* 125:443–451.

17. Parsell DA, Kowal AS, Singer MA, Lindquist S (1994) Protein disaggregation mediated by heat-shock protein Hsp104. *Nature* 372:475–478.
18. Monroe N, Han H, Shen PS, Sundquist WI, Hill CP (2017) Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. *eLife* 6:e24487.
19. Puchades C, et al. (2017) Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. *Science* 358:eaao0464.
20. Ripstein ZA, Huang R, Augustyniak R, Kay LE, Rubinstein JL (2017) Structure of a AAA+ unfoldase in the process of unfolding substrate. *eLife* 6:e25754.
21. Deville C, et al. (2017) Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. *Sci Adv* 3:e1701726.
22. Huang R, et al. (2016) Unfolding the mechanism of the AAA+ unfoldase VAT by a combined cryo-EM, solution NMR study. *Proc Natl Acad Sci USA* 113:E4190–E4199.
23. White SR, Lauring B (2007) AAA+ ATPases: Achieving diversity of function with conserved machinery. *Traffic* 8:1657–1667.
24. Barthelme D, Sauer RT (2012) Identification of the Cdc48•20S proteasome as an ancient AAA+ proteolytic machine. *Science* 337:843–846.
25. Martin A, Baker TA, Sauer RT (2008) Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat Struct Mol Biol* 15:1147–1151.
26. Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94:12366–12371.
27. Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE (2003) Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc* 125:10420–10428.
28. Tugarinov V, Kay LE (2004) An isotope labeling strategy for methyl TROSY spectroscopy. *J Biomol NMR* 28:165–172.
29. Kainosho M, et al. (2006) Optimal isotope labelling for NMR protein structure determinations. *Nature* 440:52–57.
30. Gans P, et al. (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chem Int Ed Engl* 49: 1958–1962.
31. Jiang Y, Kalodimos CG (2017) NMR studies of large proteins. *J Mol Biol* 429: 2667–2676.
32. Rosenzweig R, Kay LE (2014) Bringing dynamic molecular machines into focus by methyl-TROSY NMR. *Annu Rev Biochem* 83:291–315.
33. Gottesman S, Roche E, Zhou Y, Sauer RT (1998) The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* 12:1338–1347.
34. Benaroudj N, Zwickl P, Seemüller E, Baumeister W, Goldberg AL (2003) ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. *Mol Cell* 11:69–78.
35. Martin A, Baker TA, Sauer RT (2008) Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol Cell* 29:441–450.
36. Gronenborn AM, et al. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* 253:657–661.
37. Gelis I, et al. (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131:756–769.
38. Augustyniak R, Ferrage F, Paquin R, Lequin O, Bodenhausen G (2011) Methods to determine slow diffusion coefficients of biomolecules: Applications to Engrailed 2, a partially disordered protein. *J Biomol NMR* 50:209–218.
39. Choy W-Y, et al. (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J Mol Biol* 316:101–112.
40. Ortega A, Amorós D, García de la Torre J (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys J* 101:892–898.
41. Spera S, Bax A (1991) Empirical correlation between protein backbone conformation and C. alpha and C. beta 13C nuclear magnetic resonance chemical shifts. *J Am Chem Soc* 113:5490–5492.
42. Wishart DS, Sykes BD (1994) Chemical shifts as a tool for structure determination. *Methods Enzymol* 239:363–392.
43. Kuboniwa H, et al. (1995) Solution structure of calcium-free calmodulin. *Nat Struct Biol* 2:768–776.
44. Finn BE, Drakenberg T, Forsén S (1993) The structure of apo-calmodulin. A 1H NMR examination of the carboxy-terminal domain. *FEBS Lett* 336:368–374.
45. Latham MP, Kay LE (2012) Is buffer a good proxy for a crowded cell-like environment? A comparative NMR study of calmodulin side-chain dynamics in buffer and E. coli lysate. *PLoS One* 7:e48226.
46. Rabl C-R, Martin SR, Neumann E, Bayley PM (2002) Temperature jump kinetic study of the stability of apo-calmodulin. *Biophys Chem* 101–102:553–564, and erratum (2003) 104:697.
47. Tugarinov V, Kay LE (2006) Relaxation rates of degenerate 1H transitions in methyl groups of proteins as reporters of side-chain dynamics. *J Am Chem Soc* 128: 7299–7308.
48. Browne JP, Strom M, Martin SR, Bayley PM (1997) The role of β-sheet interactions in domain stability, folding, and target recognition reactions of calmodulin. *Biochemistry* 36:9550–9561.
49. Maupin-Furlow JA (2014) Prokaryotic ubiquitin-like protein modification. *Annu Rev Microbiol* 68:155–175.
50. Nandi D, Tahiliani P, Kumar A, Chandu D (2006) The ubiquitin-proteasome system. *J Biosci* 31:137–155.
51. Omura T (1998) Mitochondria-targeting sequence, a multi-role sorting sequence recognized at all steps of protein import into mitochondria. *J Biochem* 123: 1010–1016.
52. Forouzan D, et al. (2012) The archaeal proteasome is regulated by a network of AAA ATPases. *J Biol Chem* 287:39254–39262.
53. Lee C, Prakash S, Matouschek A (2002) Concurrent translocation of multiple polypeptide chains through the proteasomal degradation channel. *J Biol Chem* 277: 34760–34765.
54. Kenniston JA, Burton RE, Siddiqui SM, Baker TA, Sauer RT (2004) Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation. *J Struct Biol* 146:130–140.
55. San Martín Á, et al. (2017) Knots can impair protein degradation by ATP-dependent proteases. *Proc Natl Acad Sci USA* 114:9864–9869.
56. Kleinjung J, Fraternali F, Martin SR, Bayley PM (2003) Thermal unfolding simulations of apo-calmodulin using leap-dynamics. *Proteins* 50:648–656.
57. Rodriguez-Larrea D, Bayley H (2014) Protein co-translocational unfolding depends on the direction of pulling. *Nat Commun* 5:4841.
58. Rodriguez-Larrea D, Bayley H (2013) Multistep protein unfolding during nanopore translocation. *Nat Nanotechnol* 8:288–295.
59. Yokouchi T, Izumi Y, Matsufuji T, Jinbo Y, Yoshino H (2003) Unfolding intermediate of a multidomain protein, calmodulin, in urea as revealed by small-angle X-ray scattering. *FEBS Lett* 551:119–122.
60. Cabrita LD, et al. (2016) A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding. *Nat Struct Mol Biol* 23:278–285.
61. Aubin-Tam M-E, Olivares AO, Sauer RT, Baker TA, Lang MJ (2011) Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. *Cell* 145: 257–267.
62. Miller JM, Lin J, Li T, Lucius AL (2013) E. coli ClpA catalyzed polypeptide translocation is allosterically controlled by the protease ClpP. *J Mol Biol* 425:2795–2812.
63. Stigler J, Rief M (2012) Calcium-dependent folding of single calmodulin molecules. *Proc Natl Acad Sci USA* 109:17814–17819.
64. Sato T, Esaki M, Fernández JM, Endo T (2005) Comparison of the protein-unfolding pathways between mitochondrial protein import and atomic-force microscopy measurements. *Proc Natl Acad Sci USA* 102:17999–18004.
65. Matouschek A, et al. (1997) Active unfolding of precursor proteins during mitochondrial protein import. *EMBO J* 16:6727–6736.
66. Huang S, Ratliff KS, Schwartz MP, Spenner JM, Matouschek A (1999) Mitochondria unfold precursor proteins by unraveling them from their N-termini. *Nat Struct Biol* 6: 1132–1138.
67. Masino L, Martin SR, Bayley PM (2000) Ligand binding and thermodynamic stability of a multidomain protein, calmodulin. *Protein Sci* 9:1519–1529.
68. Nager AR, Baker TA, Sauer RT (2011) Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. *J Mol Biol* 413:4–16.
69. Ibrahim Z, et al. (2017) Time-resolved neutron scattering provides new insight into protein substrate processing by a AAA+ unfoldase. *Sci Rep* 7:40948.
70. Löwe J, et al. (1995) Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. *Science* 268:533–539.
71. Schrödinger, LLC (2015) *The PyMOL Molecular Graphics System. Version 2.0* (Schrödinger, LLC, New York).