# Probing allosteric interactions in homo-oligomeric molecular machines using solution NMR spectroscopy

Yuki Toyama[a,b,c] 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 3H6, Canada; and [d]Program in Molecular Medicine, The Hospital for Sick Children Research Institute, Toronto, ON, M5G 0A4, Canada

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2020.

Edited by G. Marius Clore, Laboratory of Chemical Physics, NIH, Bethesda, MD; received September 3, 2021; accepted October 27, 2021

**Developments in solution NMR spectroscopy have significantly impacted the biological questions that can now be addressed by this methodology. By means of illustration, we present here a perspective focusing on studies of a number of molecular machines that are critical for cellular homeostasis. The role of NMR in elucidating the structural dynamics of these important molecules is emphasized, focusing specifically on intersubunit allosteric communication in homo-oligomers. In many biophysical studies of oligomers, allostery is inferred by showing that models specifically including intersubunit communication best fit the data of interest. Ideally, however, experimental studies focusing on one subunit of a multisubunit system would be performed as an important complement to the more traditional bulk measurements in which signals from all components are measured simultaneously. Using an approach whereby asymmetric molecules are prepared in concert with NMR experiments focusing on the structural dynamics of individual protomers, we present examples of how intersubunit allostery can be directly observed in high-molecular-weight protein systems. These examples highlight some of the unique roles of solution NMR spectroscopy in studies of complex biomolecules and emphasize the important synergy between NMR and other atomic resolution biophysical methods.**

methyl-TROSY | human high-temperature requirement A2 protein | proteasome | ClpP1P2 | p97/VCP

Structural biology, as the name implies, is heavily biased toward the generation of high-resolution biomolecular structures and their interpretation to understand function. Detailed three-dimensional images of proteins have been forthcoming from X-ray crystallography since the late 1950s with the publication of the structure of sperm whale myoglobin by Kendrew et al. (1) and followed subsequently by the structure of human hemoglobin by Muirhead and Perutz in 1963 (2). Since these early seminal achievements, the number of available protein structures has grown exponentially, providing detailed insights into the nature and types of secondary, tertiary, and quaternary structural elements, the mechanisms by which these are stabilized in proteins, and, more generally, how a large variety of different molecular systems might function. A second high-resolution structural method, solution NMR spectroscopy, emerged in the 1980s, largely fueled by a collaboration between Kurt Wüthrich and Richard Ernst. This led, in 1985, to the first three-dimensional solution structure of a protein, the proteinase inhibitor IIA from bull seminal plasma (3). More recently, electron cryomicroscopy (cryo-EM) has emerged as a forefront technology for studies of biomolecular structure (4), and while the focus has been on large molecular complexes, there is great optimism that structures of proteins less than 100 kDa may also become routinely available using this approach (5).

With the emergence of cryo-EM, the continued development and application of X-ray methods, and significantly improved predictive algorithms of structure based on artificial intelligence (6, 7), the role of solution NMR spectroscopy as a tool for elucidating detailed structures of proteins has become more ambiguous than in the past. Most certainly, there will still be a need for the experimental determination of atomic resolution structures of relatively small proteins and their complexes in cases in which other structural modalities fail, and the unique ability to investigate molecules in their more natural physiological solution state remains an attractive feature of NMR. Yet both X-ray and cryo-EM methods provide, in general, a more facile and rapid approach for determining structure than NMR spectroscopy. Furthermore, while the motivation behind any X-ray or cryo-EM study is based on the determination of a high-resolution structure (or structures) to inform on some aspect of function as a de facto initial goal, NMR applications can often be, at least initially, less well focused. This has led, over the past several years, to a reevaluation of the role of biomolecular solution NMR spectroscopy in structural biology and biophysics, as its impact as a tool for solving structures has diminished. Yet advances in solution-state NMR studies have been remarkable over the past decades (8–15), and, now much more than ever before, it is possible to address a wide range of questions focusing on the kinetics, thermodynamics, and dynamics of interactions, both intra- and intermolecular, involving a wide range of biomolecular conformers from folded to intrinsically disordered (16). Insights provided by such studies, pertaining to both biological function and misfunction, are most often simply not available using other technologies.

## Significance

**Much like the macroscopic machines that carry out important functions in everyday life, microscopic molecular machines play critical roles in cellular activities. Often, these are regulated by communication between different regions of the molecule. Here, several examples are presented focusing on solution NMR spectroscopy studies of machines that are important for protein quality control in the cell. In particular, the unique role of NMR in such studies is emphasized, highlighting that applications to systems with aggregate molecular weights of several hundreds of kilodaltons are now possible and that insights from NMR studies are highly complementary to, yet distinct from, those obtained via X-ray and electron cryomicroscopy approaches.**

Author contributions: Y.T. and L.E.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.



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

Published December 10, 2021.

In the perspective that follows, we have chosen to focus on studies of protomer cooperativity and allostery in molecular machines, one area of research in the laboratory, so as to provide examples of how NMR can uniquely shape our understanding of protein science in a manner that is complementary to other biophysical approaches. We have particularly chosen this topic because it is generally not well appreciated that solution NMR spectroscopy can be used to provide quantitative insights into the structural dynamics of large protein complexes with aggregate molecular weights in the many hundreds of kilodaltons. In what follows, the goal is not to be inclusive but to initially focus briefly on the methodology that has led to these applications and then to choose a few examples from systems that we have worked on for the past several years, highlighting unique features that can be obtained through NMR studies.

## Results and Discussion

**Methyl Transverse Relaxation Optimized Spectroscopy as a Tool to Explore Structural Dynamics in Molecular Machines.** A contributing factor to the success of solution NMR for studies of high-molecular-weight proteins has been the development of transverse relaxation optimized spectroscopy (TROSY) as a tool to improve spectral sensitivity and resolution (8, 17). The basic principle behind this technique involves the net suppression of fluctuating magnetic fields whose time dependence derives from molecular tumbling and internal motions. Although these local fields are typically more than four orders of magnitude smaller than the static magnetic field of the spectrometer, their effects on NMR spectra can be profound, as their time dependence leads to a rapid decay of NMR signals in the case of large molecules (18). The reduction of these local fields, therefore, is highly desirable. The first application of TROSY involved amide $^{1}H$-$^{15}N$ spin pairs in which one-half of the inherent signal in two-dimensional amide spectra derives from slowly relaxing components that "feel" reduced effective local fields (8). So called amide-TROSY, this innovation has had a major impact on increasing the size of protein molecules amenable to routine study to ~50 to 100 kDa (19), with the exact molecular weight limit very much dependent on the overall shape of the protein under investigation as well as on the internal dynamics of the molecule. In cases in which these dynamics are on the micro- to millisecond timescale, they can give rise to a "blurring" of the signals, thus decreasing the utility of this approach.

A long-standing goal of one of us (L.E.K.) has been to develop methodology to study the structural dynamics of much larger systems than 50 to 100 kDa in an effort to complement the beautiful and highly informative, albeit static, high-resolution structures of molecular machines that have been and continue to be produced by crystallographic studies and, more recently, by cryo-EM. What emerged, building upon techniques for labeling methyl groups in proteins as $^{13}CH_3$ in an otherwise deuterated background (20, 21), was what is now referred to as methyl-TROSY (17). The underlying physics behind the methyl-TROSY approach is based on quantum mechanics and is considerably more involved than for its amide-TROSY counterpart, as there are four NMR spins ($^{13}C$-$H_3$) rather than two ($^{15}N$-$^{1}H$), giving rise to a much more complex spin energy-level diagram that has to be navigated in the construction of experiments. Nevertheless, an intuitive feel for the methyl-TROSY effect can be partially (and simply) obtained by consideration of some concepts in NMR that are covered in any freshman-level course in organic chemistry. The basic ideas are discussed here without any attempt at rigor so as to illustrate qualitatively the power of spin alchemy that is exploited in one form or another in all NMR experiments (22). Fig. 1 *A*, *Left* shows a one-dimensional $^{13}C$ NMR spectrum of an isolated methylene group ($^{13}CH_2$) in a small molecule that tumbles rapidly in solution. Attached to the $^{13}C$ nucleus is a pair of hydrogens (referred to as protons in what follows) that affect the energy of the carbon nucleus through an interaction that is mediated via the shared electrons in each C-H bond. Each $^{13}C$-$^{1}H$ interaction leads to a pair of lines in a $^{13}C$ NMR spectrum, separated by a fixed distance called the coupling constant and denoted by $J_{HC}$. Because there are two $^{1}Hs$ in a $^{13}CH_2$ spin system, the $^{13}C$ spectrum of a methylene group consists of three lines (Fig. 1 *A*, *Right*) with the ratio of intensities of the lines given by 1:2:1. As a small molecule is being considered, it is assumed that the three lines decay at very similar rates. In the case of an application involving a protein for which the tumbling is much slower and when the methylene group is rigidly affixed to the molecule, there are still three components to the $^{13}C$ spectrum, but the central component relaxes approximately twofold faster than the outer lines so that a triplet with an approximate 1:1:1 intensity is observed (Fig. 1*B*). In contrast, if the $^{13}CH_2$ group is attached to a chain that rotates rapidly about an averaging axis along which both C-H vectors have equivalent projections, then the spectrum of Fig. 1*C* is obtained, for which the central line is narrowed. Such a situation arises for a methylene of a fatty acid chain in a biological membrane in which there is rapid averaging about the long axis of the chain. For the case of

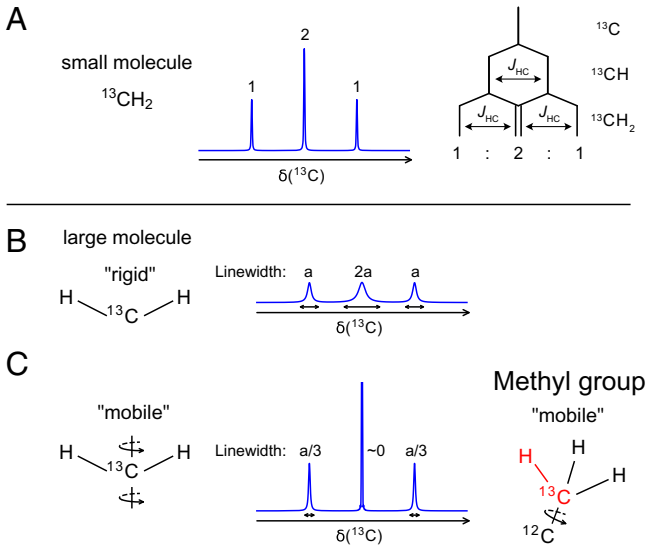


**Fig. 1.** Methyl-TROSY NMR. (*A*, *Left*) $^{13}C$ NMR spectrum of an isolated $^{13}CH_2$ fragment attached to a small molecule. A triplet is observed with each line separated from the next by $J_{HC}$ ~130 Hz, where $J_{HC}$ is the one-bond $^{1}H$-$^{13}C$ scalar coupling constant. (*Right*) Schematic illustrating why a 1:2:1 triplet structure is observed for a $^{13}CH_2$ moiety. An isolated $^{13}C$ nucleus gives rise to a single line, which in turn is split into two lines in the case of a methine group. Each line is subsequently spilt into an additional pair of lines in the case of a methylene. Assuming that all $^{1}H$-$^{13}C$ $J_{HC}$ values are the same, a triplet results with the middle line twice more intense than the outer lines. (*B*) When the methylene group is rigidly affixed within a large molecule, such as a protein, a triplet is still observed, but the linewidth of the middle line (denoted by 2a) is approximately twofold larger than the outer lines so that the intensities of each component are similar. (*C*, *Left*) As in *B* except for a mobile methylene in which there is free rotation about an averaging axis, as indicated. This scenario occurs for fatty acids in a bilayer in which rapid rotation about the long axis of the fatty acid (bilayer normal) occurs. In this case, the motion leads to a reduction of all three linewidths, but particularly for the middle line, where local magnetic fields generated by the geminal $^{1}H$ spins do not lead to line broadening (23). (*Right*) Methyl groups experience rapid rotation about a threefold symmetry axis formed by the bond connecting the methyl carbon and the adjacent atom. Creation of a "super spin" comprising the $^{13}C$ and one of the $^{1}H$ nuclei of the methyl (red) results in a slowly relaxing middle line that can be exploited in studies of high-molecular-weight biomolecules.

an isolated methylene group that averages in this way, the time-dependent magnetic fields from the pair of proton spins directly bonded to the $^{13}C$ nucleus that would normally dictate the linewidth of the central component (middle line) cancel for large molecules (23) and thus do not contribute to its signal decay. Clearly, NMR applications that could exploit such a narrow line would benefit significantly from improvements in resolution and sensitivity. But how to generate this situation in proteins, as methylene groups are not typically rapidly rotating about an averaging axis to create the spectrum shown? The secret lies in 1) using methyl groups, as these do rotate rapidly around a three-fold averaging axis, typically with rates of $10^{10}$ to $10^{12}$ $s^{-1}$ (24, 25) and 2) in convincing a methyl group to behave like a methylene during certain periods of the experiment. It turns out that the latter requirement can be achieved by creating a "super spin" comprising the $^{13}C$ and one of the $^{1}H$s (shown in red in Fig. 1*C*). In the methyl-TROSY effect, the super spin is analogous to the $^{13}C$ spin of a methylene in that there is a pair of bound protons (black in Fig. 1*C*) so that an analogous spectrum to that obtained for a methylene can now be obtained for a methyl. To understand how the super spin is created requires a bit of quantum mechanics, but the effect can be readily achieved using a standard NMR experiment termed HMQC (heteronuclear multiple quantum coherence) (26, 27) that is very easily implemented. This discussion has only focused on one aspect of methyl-TROSY, namely, line narrowing in the $^{13}C$ dimension of the experiment, and is superficial in many ways, yet it hopefully illustrates how very simple pictures can be used to direct the development of experiments that are specifically tailored for a desired application. The interested reader is referred to a number of papers that discuss the methyl-TROSY effect in great detail (17, 28, 29).

In order to fully exploit the methyl-TROSY effect, as for amide-TROSY, it is necessary to produce molecules that are highly deuterated, with the exception of the methyl groups, of course. This ensures that local magnetic fields derived from hydrogens proximal to the methyl probes are minimized (the size of the deuteron bar magnet is ∼1/7th that of the proton and the rate of signal decay scales as the square of the local magnetic field strength). Robust methods for producing such proteins have been described in the literature (21, 30), and typically highly deuterated molecules labeled as Ile $C^{\delta}H_3$, Leu $C^{\delta}H_3$, Val $C^{\gamma}H_3$ are produced [in general, only one of the prochiral $C^{\delta}/C^{\gamma}$ methyls of Leu/Val are labeled (20, 31) and often supplemented by $^{13}C^{\varepsilon}H_3$ labeling of Met (13)]. In what follows, we refer to these proteins as U-$^{2}H$, ILVM-$^{13}CH_3$–labeled molecules. Precursors are now commercially available for the production of highly deuterated proteins with $^{13}CH_3$ labeling at any of the methyl positions of the six methyl-containing amino acids.

Having provided a brief description of the methodology employed, we now focus on examples from four protein complexes so as to emphasize the functional insights that can be obtained on a variety of different molecular systems.

**Interprotomer Allostery in the Proteostasis HtrA2 Enzyme Controls Function.** The human high-temperature requirement A2 (HtrA2, 105 kDa) mitochondrial enzyme plays essential roles in cellular activity, functioning as a stress-protective protease (32–35), with mutations in the protein giving rise to a series of neurological disorders including Parkinson's disease and essential tremor (36, 37). X-ray studies reveal a trimeric architecture in which individual protomers, each comprising one PDZ and one protease domain, are arranged in a pyramid-like structure (38) (Fig. 2*A*). The PDZ domain is juxtaposed on the protease domain such that the catalytic triad is sequestered, as is a binding region on the PDZ domain that interacts with carboxyl-terminal residues of substrates. As the only available structure is of an inactive, substrate-free closed form of the enzyme (38), it remains unknown how substrate binding leads to HtrA2 activation or whether protomers communicate with each other during the activation process. Some evidence that intersubunit allostery is prevalent is available from a recent methyl-TROSY–based NMR study showing that ligands successively bind to PDZ domains with increasingly higher affinities, resulting in a change of subunit structure from a closed to an open state (Fig. 2*A*) (39). In what follows, we will refer to these ligands, mimicking carboxyl-terminal sites of protein substrates, as PDZ-peptides. Additionally, the NMR study revealed that trimers can form hexamers, either in apo- or PDZ-peptide bound states, but that the hexamer affinity for PDZ-peptides is markedly reduced relative to the trimer so that ligand binding to the hexamer can be ignored. The best fit models to peak intensities in NMR spectra recorded as a function of added PDZ-peptide were those that included intersubunit allostery. Allostery is thus inferred exclusively from an analysis of bulk signals (i.e., all protomers contribute) and goodness of model fitting, as is most often the case in biophysical studies, rather than from direct evidence as to how an individual protomer responds as neighboring subunits are perturbed through ligand binding.

In an effort to control the peptide-bound state of each protomer, a cysteine-based strategy has been used in which subunits are prepared that are either disulfide locked, mimicking the closed state, or stapled with PDZ-peptide (i.e., PDZ-peptide covalently linked to the PDZ domain at E425 via a Cys-based chemistry approach), serving as a mimic of a stable open structure (40) (Fig. 2*B*). Symmetric trimers prepared in this manner (NMR invisible) are then mixed in a prescribed ratio with NMR active trimers (i.e., $^{13}CH_3$ methyl–labeled) without the Cys modifications (in the example of Fig. 2*B* 9:1), and the proteins are unfolded and subsequently refolded to generate a distribution of configurations A through D as highlighted in Fig. 2*B*, in which only configurations B through D are NMR visible. By skewing the population of NMR-inactive (gray) to NMR-active (green outline, protease and PDZ domains denoted in pink and blue, respectively) protomers in a 9:1 ratio, it is possible to engineer a dominant HtrA2 configuration (configuration "B") in which the NMR-visible subunit has either all closed or all open NMR-invisible neighbors. NMR titrations of this set of molecules (Fig. 2 *C*, *Top* and *Bottom*) allow a direct measurement of the first (two neighboring closed protomers) and third (two neighboring ligated, open protomers) PDZ-peptide binding constants that are in excellent agreement with the values fit using bulk signal measurements (Fig. 2*A*).

In order to understand the origin of the differences in affinity associated with successive ligand-binding events, the NMR spectra of the samples shown in Fig. 2 *D*, *Top Left* are compared (*Bottom Left*). Although the spectra are very similar, indicating that the NMR-visible protomer remains in the closed state irrespective of the open/closed status of neighboring subunits, there are important differences. A marked reduction in some of the peak intensities is observed for the asymmetric sample, indicating increased millisecond–microsecond timescale dynamics, while small, but clear, chemical shift differences are also noted. These changes in dynamics and chemical shifts map to the intrasubunit interface between protease and PDZ domains as well as to intersubunit regions that make protease–protease domain contacts (Fig. 2 *D*, *Right*). A picture thus emerges whereby the changes to intersubunit protease contacts leads to conformational heterogeneity at the intradomain interface, priming the closed subunit to open and subsequently bind substrate (40).

The mixing experiments used in the NMR studies are also valuable in establishing how the activity of a given subunit depends on neighboring protomers. For example, peptidase assays, performed on samples prepared by mixing disulfide

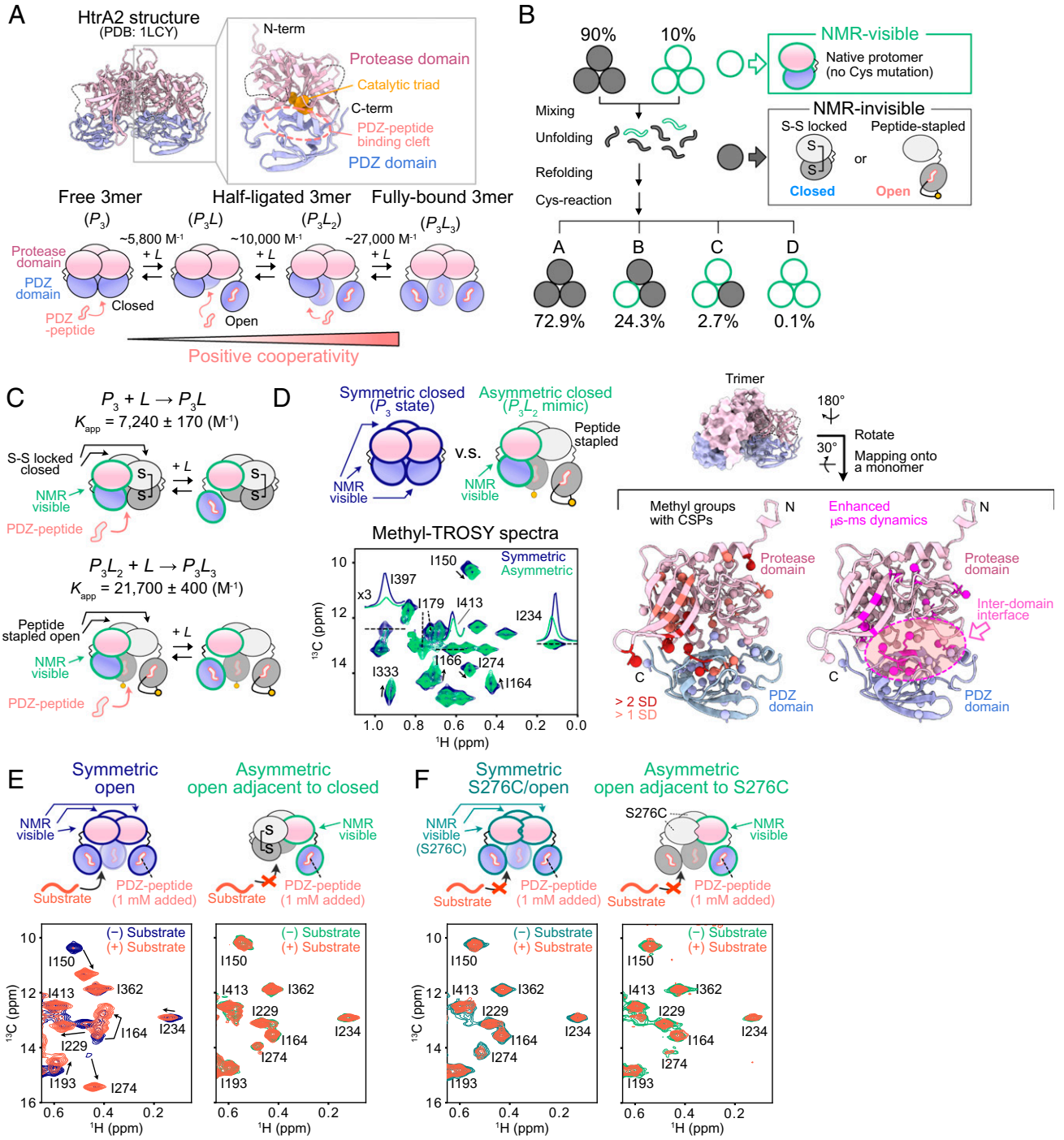


**Fig. 2.** Intersubunit cooperativity modulates HtrA2 function (40). (*A*) Schematic of PDZ-peptide binding to the HtrA2 trimer (39) along with the crystal structure of trimeric HtrA2 [Protein Data Bank (PDB) identification (ID): 1LCY (38)]; values of association constants for each binding step are indicated. The protease and PDZ domains of a subunit are in the closed configuration prior to binding PDZ-peptide, after which an open structure is formed (39). (*B*) Summary of the strategy used to produce mixed trimers. Trimers are shown by three circles, where the gray circles indicate NMR-invisible, disulfide-locked or PDZ-peptide–stapled (attached) subunits, while the green circles correspond to NMR-visible, no-Cys, protomers. Configurations B and C are NMR active (signal is not observed from D as the population is too low), with the majority of the signal derived from B. (*C*) Schematic of configurations reporting on binding of PDZ-peptide to (*Top*) $P_3$ (apo HtrA2) and (*Bottom*) $P_3L_2$. The resulting affinity constants agree with those obtained from fits of titration data recorded on symmetric samples in *A*. (*D*) Selected region of $^{13}C$-$^{1}H$ HMQC spectra recorded on symmetric and asymmetric samples, as illustrated (*Top*), highlighting small chemical shift perturbations (CSPs) and significant intensity differences between spectra. *proR* ILVM methyl groups with significant CSPs (*Bottom Left*, light orange > 1 SD and deep red > 2 SD spheres) and differences in motion (*Bottom Right*, magenta spheres) mapped onto the HtrA2 protomer (PDB ID: 1LCY). Other methyl groups are shown in light pink (protease) and blue (PDZ). (*E*, *Left*) Binding of substrate peptides to symmetric, open (PDZ-peptide–bound) HtrA2 results in large CSPs that likely reflect a concerted structural change that leads to an active conformation. (*Right*) In contract, no CSPs are observed when substrate is added to an asymmetric sample with one or two locked, closed subunits, indicating that the concerted conformational transition to the active state cannot occur. (*F*) The S276C disease mutant prevents a concerted conformational change upon addition of substrate, for both symmetric (all protomers are S276C) and asymmetric (one or two S276C protomers) samples. For asymmetric samples in *E* and *F*, only protomers outlined in green (with pink and blue protease and PDZ domains, respectively) are NMR active. Adapted from ref. 40.

locked (closed) protomers with wild-type (WT) subunits, showed that only the symmetric trimers in which all protomers are WT remain active (40). Consistent with the activity assays, binding of active site substrates to a symmetric open HtrA2 state (saturated with PDZ-peptide) led to significant changes in chemical shifts (Fig. 2 *E*, *Left*), indicating a conformational change involving protease domains. Yet no such changes were observed when substrate was added to asymmetric samples in which at least one of the subunits was locked in the closed conformation (Fig. 2 *E*, *Right*), presumably because a concerted structural change that is required for a fully active enzyme was prevented. In this manner, only when a sufficiently high concentration of substrate is achieved such that all PDZ domains are ligated and in the open conformation does the enzyme become active, preventing unwanted degradation of nonsubstrate or sparsely populated substrate molecules.

Dysregulation of the allosteric network that connects adjacent protomers in HtrA2 plays an important role in the loss of function of the S276C disease mutant that is causative for a neurodegenerative condition in mice (41). The X-ray–derived structure of S276C HtrA2 is very similar to that of the WT enzyme (42), suggesting that a gross structural change is not the root cause of disease. Spectra recorded of symmetric (all three protomers are S276C mutants; Fig. 2 *F*, *Left*) or asymmetric (one or two S276 mutant protomers; Fig. 2 *F*, *Right*) HtrA2 samples saturated with PDZ-peptide and either with or without substrate are superimposable, indicating that addition of substrate does not result in the concerted conformational change that is required for HtrA2 activity, much like for the case involving trimers harboring at least one disulfide locked subunit (Fig. 2 *E*, *Right*).

**The Dynamic Archaeal Proteasome Gates—Cooperative or Noncooperative Motions?** The 20S core particle (CP) proteasome (670 kDa) plays an integral role in proteostasis by degrading misfolded or mutated proteins and is involved in the regulation of many biological processes by controlling the concentrations of key cellular proteins in a timely manner (43–45). Additionally, proteasome activity is coupled to the immune system, as degraded protein fragments can be used as antigenic peptides (44). Detailed X-ray studies revealed that the proteasome adopts a cylindrical structure comprised of four rings that are arranged as shown in Fig. 3 *A*, *Top Left* (46, 47). In the case of the 20S CP from *Thermoplasma acidophilum*, a pair of protomers, termed α- and β-, form $\alpha_7$ and $\beta_7$ heptameric rings, respectively, that then stack concentrically to generate an $\alpha_7\beta_7\beta_7\alpha_7$ structure. A pore runs through the 20S CP, connecting the two antechambers with a central catalytic chamber that contains 14-proteolytic centers in which substrate degradation occurs. Proteasomes from eukaryotes have a similar architecture as the archaeal version, but each of the seven α- and β-protomers are distinct, with the β-subunits having different catalytic activities (47). Although the 20S CP is often in complex with regulators that control entry of substrates (44, 48–50), preventing unwanted proteolysis, there is an increasing realization that it can also exist in an apo-state (51), in which the first 10 to 15 N-terminal residues of each $\alpha_7$-protomer function as a gate that "guards" the α-annulus at the top and bottom of the 20S CP (Fig. 3 *A*, *Bottom Left*) and hence restricts access to the proteolytic chamber. In the case of the *T. acidophilum* CP, these gating residues are unfolded, yet they exist predominantly in two orientations with respect to the 20S CP, either outside (Fig. 3 *A*, *Top Right*, blue) or inside (Fig. 3 *A*, *Top Right*, pink) the lumen of the barrel (52). Fig. 3 *A*, *Bottom Right* shows a region of a $^{13}C$-$^{1}H$ HMQC spectrum of an $\alpha_7$-ring focusing on cross peaks from methionine residues, including methinione-1 (M-1, an additional methionine introduced at the N terminus of the construct, in addition to M1 of the natural sequence) and M6 that are part of the gates. Notably, the multiple peaks for M-1 can be assigned individually to the in and out states of the gates, Fig. 3 *A*, *Bottom Right*, as described previously (52), and, additionally, NMR experiments show that each of the gates exchanges its position between out and in conformations with lifetimes on the seconds timescale (52). From the intensities of the in and out peaks, an equilibrium fractional population for the in conformation of 29% is obtained, corresponding to two of the seven gates of an $\alpha_7$-ring occupying the in state.

These bulk NMR measurements that include signal contributions from all 14 gates of the CP provide insight into how unfolded polypeptide chains can regulate proteasome function. Yet the intrinsic probability of a single gate entering the proteasome lumen cannot be obtained from the simple NMR spectrum of Fig. 3*A*, and whether the in/out motion of each gate is cooperative, that is, whether the probability of a gate entering into the proteasome lumen depends on the positions of neighboring gating residues, can also not be ascertained.

In order to address these biophysical questions, a set of molecules was prepared by mixing M1I U-$^{2}H$, M-$^{13}CH_3$ α-protomers (referred to as pseudo WT or pWT) with gateless U-$^{2}H$ α-subunits (Δ-12) in a ∼1:20 ratio, using an unfolding/refolding procedure in which $\alpha_7$ rings are produced with the two types of protomers distributed in a random manner in each $\alpha_7$ particle (53). In this way, the relative populations of the different rings (i.e., one pWT, six gateless: 25.7%; two pWT, five gateless: 4.1%) can be calculated using simple binomial statistics. The M1I mutant was used for spectral simplification and did not influence the results (53). As only the pWT protomers were $^{13}CH_3$ labeled at Met positions (M-$^{13}CH_3$), all the NMR signals are derived from them. Roughly 75% of the NMR signal derives from $\alpha_7$ rings containing only a single gate, illustrated in the inset of Fig. 3 *B*, *Top* (red versus blue protomers). Notably, the resulting spectrum shows that the intrinsic probability of a gate entering the lumen of the proteasome, $P^{in}$, is not 2/7, as might have been inferred from bulk measurements (Fig. 3 *A*, *Bottom Right*), but ∼95% based on the relative intensities of the M-$1_{in}$ and M-$1_{out}$ peaks. In order to verify that $P^{in} = 0.95$ is not an artifact of using $\alpha_7$ rings that were largely gateless, a second hybrid $\alpha_7$ particle was generated by mixing pWT α-subunits with protomers containing a Gly-rich gate (∼1:20 ratio, with only pWT subunits NMR labeled). Notably, these Gly-rich gates do not enter the proteasomal lumen ($P^{in}$ ∼0) (53), so that the $P^{in}$ value obtained in this system provides a second measure of the probability of an initial gate entering an empty barrel, unencumbered by the presence of other gating molecules. Again, $P^{in} = 0.95$ was obtained (Fig. 3 *B*, *Bottom*). Taken together, these experiments establish a high probability for gate entry and illustrate how conclusions based on bulk measurements, rather than on measurements focusing on the microscopic properties of the system, can be misleading.

In order to address whether the in/out motions of gates from separate subunits of an $\alpha_7$ particle are cooperative, a set of NMR samples consisting of different mixtures of U-$^{2}H$, M-$^{13}CH_3$ pWT and U-$^{2}H$ gateless protomers was prepared and the fraction of the in conformation of the gates determined from measuring peak intensities derived from the in and out states of the gates (Fig. 3*C*) (53). The resulting data were analyzed in terms of a model in which the probability of a gate entering the lumen of the barrel depends only on the number of gates already in the in state (Model 1) and a second model in which interactions between immediately neighboring gates in the out state are explicitly taken into account (Model 2). These models are, in a sense, orthogonal, as they involve different assumptions, with interactions either in the in (Model 1) or out (Model 2) states explicitly accounted for. Thus, similar probability values obtained from an analysis using both models would strengthen conclusions and establish that the result is not

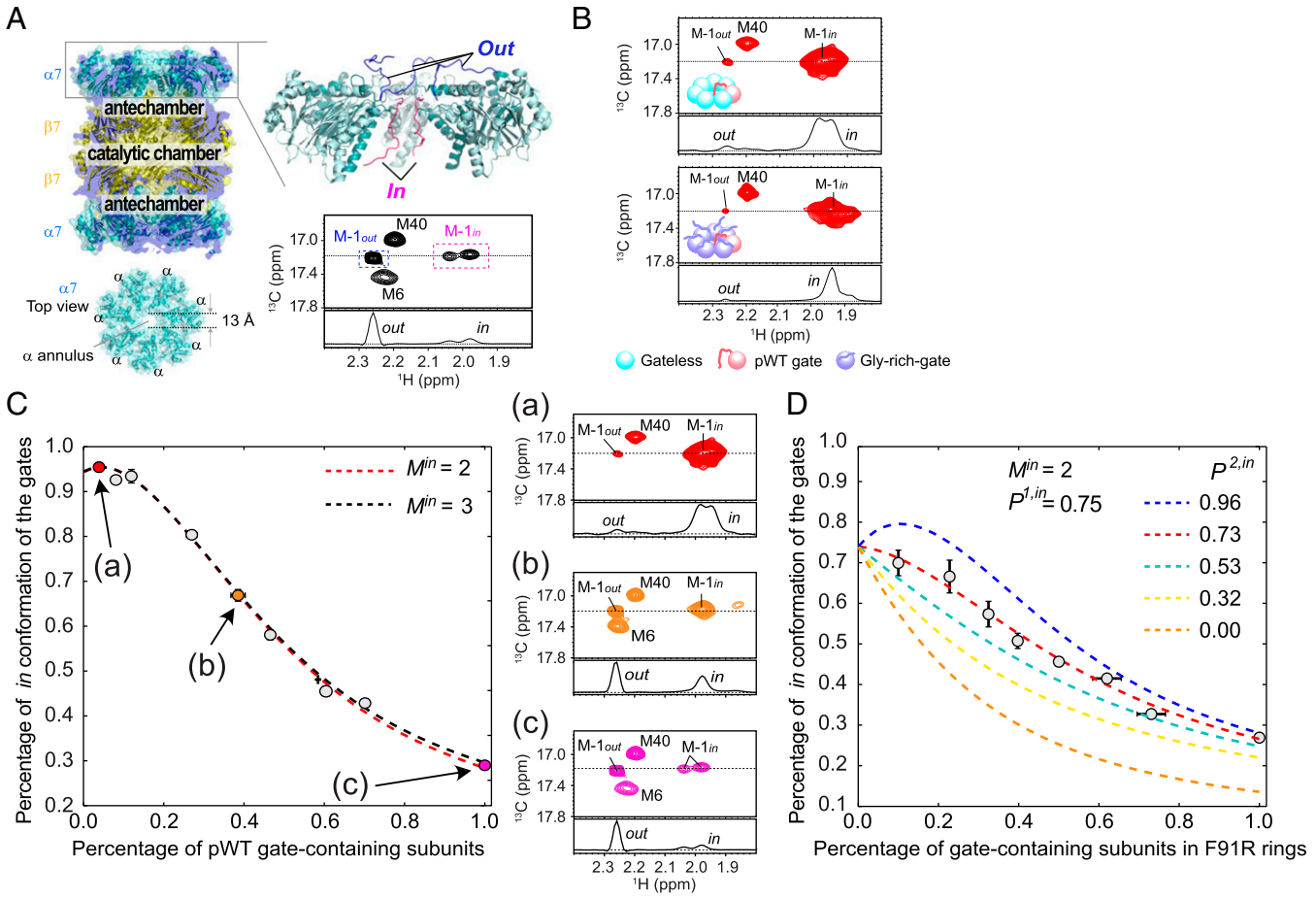


**Fig. 3.** *T. acidophilum* proteasome gating is not cooperative (53). (*A*) Side (*Top Left*) and top (*Bottom Left*) views of 20S CP [PDB ID: 1YA7 (80)]. Two subunits have been removed from the side view to highlight the interior of the barrel and the ante- and catalytic chambers. (*Top Right*) Enlarged $\alpha_7$ ring illustrating out (blue) and in (pink) gates [PDB ID: 2KU1 (52)]. (*Bottom Right*) Selected region of a $^{13}$C-$^{1}$H HMQC spectrum of pWT $\alpha_7$, showing three peaks for M-1, corresponding to out (blue) and in (pink) gate conformations. A fractional population of in gates of 28.9 ± 0.8% is calculated from peak volumes (53). (*B*) Selected region of $^{13}$C-$^{1}$H HMQC spectra of $\alpha_7$ prepared by mixing 5% pWT (red, NMR visible) with either (*Top*) 95% gateless (blue; NMR invisible) or (*Bottom*) 95% Gly-rich gates (purple; NMR invisible). In both cases, the fractional population of in gates is ~95%. (*C*) Fractional population of in gates measured from NMR cross peak volumes as a function of pWT subunits in samples generated by mixing pWT (NMR visible) and gateless (NMR invisible) protomers. Dashed red and black lines correspond to fits of Model 1 with $M^{in}$ = 2 and 3, respectively, to the titration data to obtain $P^{1,in}$ = 0.95 ± 0.01, $P^{2,in}$ = 1.0 ± 0.03 for $M^{in}$ = 2 and $P^{1,in}$ = 0.95 ± 0.01, $P^{2,in}$ = 1.0 ± 0.02, $P^{3,in}$ = 0.05 ± 0.03 for $M^{in}$ = 3. (*a–c*) $^{13}$C-$^{1}$H HMQC spectra focusing on Met peaks used in the analysis, shown for a number of pWT fractions. (*D*) Fits of titration data as in *C* with the exception that samples were obtained by mixing 5% F91R (NMR visible) with 95%, F91R gateless (NMR invisible) protomers. A best fit with Model 1 was obtained with $P^{1,in} \sim P^{2,in}$ = 0.75 (red dashed lines). Additional curves were generated for different $P^{2,in}$ values that are in poor agreement with the experimental data, indicating that the system is not cooperative. Adapted from ref. 53.

model dependent. The data were fit assuming maximal numbers for in gates ($M^{in}$) of either two or three, as the size of the α-annulus limits the number of chains that can penetrate the lumen. Using Model 1, best fits were obtained with $P^{1,in}$ = 0.95 ± 0.01, $P^{2,in}$ = 1.0 ± 0.03 for $M^{in}$ = 2 and $P^{1,in}$ = 0.95 ± 0.01, $P^{2,in}$ = 1.0 ± 0.02, and $P^{3,in}$ = 0.05 ± 0.03 for $M^{in}$ = 3 (Fig. 3*C*), where $P^{i,in}$ is the probability of the $i$th gate entering the lumen given that $i - 1$ gates are already inside. This analysis is consistent with a maximum of two gates entering the lumen, and since $P^{1,in} \sim P^{2,in} \sim 1$, gating is not negatively cooperative. That is, entry of the second gate into the barrel chamber is not decreased by the presence of the first. However, because the $P^{i,in}$ values are close to 1, it is not possible to establish whether gate entry is positively cooperative from this experiment. Similar conclusions are obtained when using Model 2.

In order to obtain further insight into gating cooperativity, we searched for an $\alpha_7$ mutant that would decrease the $P^{1,in}$ value. This was achieved by measuring contacts between gates localized to the inside of the barrel and residues of the lumen by attaching a paramagnetic spin label at several positions along the gate (one residue at a time) and quantifying changes to spectra of U-$^{2}$H, ILVM-$^{13}$CH$_3$–labeled $\alpha_7$ rings (53). As the spin label broadens peaks according to the inverse sixth power of the distance between the label and the methyl group probed, it was possible to highlight key regions of the lumen that stabilize the in-gate conformation and to identify possible sites of mutation to change $P^{in}$; an F91R mutation was chosen in this way, as position 91 is proximal to the region most affected by the spin label. Fig. 3*D* shows a profile of the percent in conformation versus the fraction of gate-containing α-subunits, obtained from an experiment similar to that described in the context of Fig. 3*C*, but where all of the $\alpha_7$ rings contained the F91R mutation that significantly decreases $P^{in}$. A best fit was obtained for $P^{1,in}$ = 0.75 ± 0.02, $P^{2,in}$ = 0.73 ± 0.03 (Model 1), so that $P^{1,in} \sim P^{2,in}$, as before, but it is now clear that gating is neither negatively ($P^{1,in} > P^{2,in}$) nor positively ($P^{1,in} < P^{2,in}$) cooperative. A similar analysis of the data using Model 2 further supported these conclusions.

**Subunit Allostery Regulates *Mycobacterium tuberculosis* ClpP1P2 (MtClpP1P2) Function.** The 300-kDa ClpP1P2 protease from *M. tuberculosis* is essential for bacterial growth and virulence and, consequently, has been a target for the development of small molecules that dysregulate its function (54). The enzyme consists of two distinct heptameric rings, MtClpP1 and MtClpP2, each comprised of seven identical subunits that are arranged in a concentric manner to form a tetradecameric cylindrical structure in which the 14 active sites of the complex are sequestered from solvent (55). Unlike many other ClpPs that are functional in the apo-state and that are comprised of pairs of identical heptamers, MtClpP1P2 cannot cleave substrates without activator peptides binding to catalytic sites in the enzyme (56). These activators compete with substrate for binding so that an enzyme exhibiting maximal turnover would have a number of its sites "reserved" for substrate. In an effort to understand the role of allostery in this unusual, yet medically relevant, ClpP system, a set of molecules has been generated by mixing NMR-visible WT and NMR-invisible modified subunits, similar to what was described for HtrA2 and $\alpha_7$ in the previous two sections, for study using methyl-TROSY–based NMR methods (57).

Fig. 4 *A*, *Left* highlights Ile and Met regions of a $^{13}C$-$^{1}H$ HMQC spectrum of MtClpP1P2 in which the MtClpP1 and MtClpP2 rings were U-$^{2}H$, ILVM-$^{13}CH_3$– (NMR active) and U-$^{2}H$– (NMR inactive) labeled, respectively. All methyl groups from the 17 Ile, 8 Met, and 7 Val residues are assigned, as well as methyls from 14 of the 18 Leu residues, providing a large number of probes to study allosteric effects in this complex (57). Upon titration of MtClpP1P2 with an activating dipeptide, benzoyl-Leu-Leu (Bz-LL), a new set of peaks appeared (orange), and, concomitantly, the old set disappeared (purple), as illustrated in Fig. 4 *A*, *Top Right* for a number of methyl groups in the MtClpP1 ring of MtClpP1P2. Plots of normalized peak intensities from the set of disappearing/reappearing correlations are shown in Fig. 4 *A*, *Bottom Right* for MtClpP1 (average over 52 peaks) that were subsequently fit to a simple Hill model to yield an effective dissociation constant of 3.3 ± 0.2 mM and a Hill coefficient of 7.4 ± 0.9. The weak affinity suggests that binding of Bz-LL should be in the fast exchange regime, whereby a single "average" peak for each methyl probe titrates from the unbound to the bound position with increasing concentration of Bz-LL. Yet separate sets of peaks are observed in this case, suggesting that binding of Bz-LL is accompanied by a much slower transition in which an inactive conformation is converted to one that is active (recall that Bz-LL is an activator). The large Hill coefficient for the association of Bz-LL with MtClpP1 (a high value of 7.9 ± 0.7 is also obtained for MtClpP2) provides strong evidence for a cooperative structural transition to the active state. To obtain further insight, leading to a more comprehensive model of allostery in this system, activity profiles of MtClpP1P2 were obtained as a function of concentrations of Bz-LL and a small fluorogenic substrate peptide. Notably, the initial rise in activity observed in fluorescence assays closely parallels the increase in the fraction of Bz-LL–bound MtClpP1P2 established by the NMR titration (Fig. 4 *A*, *Bottom Right*, green band). Combined fits of the NMR titration data and the activity assays could be accomplished using a modified Monod-Wyman-Changeux (MWC) model (57) from which binding affinities for both substrate and activator peptides were obtained, as well as an equilibrium constant, L~$2 \times 10^5$, given by the ratio of concentrations of inactive to active MtClpP1P2 in the absence of substrate or activator.

In order to understand the structural basis for the catalytically dead and active states (so-called T and R forms in the MWC parlance, respectively), the conformations of apo-(T) (57) and fully Bz-LL bound (R) (55) states are compared in Fig. 4*B*, highlighting a number of significant differences that are responsible for the lack of enzymatic activity in the absence of activator peptide. Notably, the disordered handle regions in the compressed apo-state result in misalignment of catalytic residues and in disengagement of residues that form "oligomerization sensors" linking protomers from opposite rings that are required for stabilizing the active conformation of the enzyme (57). In contrast, binding of Bz-LL leads to the adoption of an extended structure, similar to other active ClpPs, and formation of a well-defined handle region that is required for catalysis (55).

Insight into allostery on a per-subunit level can be obtained by NMR studies on complexes prepared by mixing NMR-active hybrid MtClpP1 and WT U-$^{2}H$ MtClpP2 rings. In Fig. 4*C*, the MtClpP1 heptamers used in this study are indicated (*Left*; NMR active protomers are denoted by filled green circles), along with schematics highlighting the most populated, NMR-observable MtClpP1 heptameric rings that are produced by the mixing procedure (*Right*). Several sets of MtClpP1P2 particles were produced (57), including those in which the NMR-active, unbound MtClpP1 subunit is surrounded by NMR-invisible protomers that are a) unbound or b) covalently attached with GLF-CMK (Z-Gly-Leu-Phe-chloromethyl ketone), in which GLF-CMK is a substrate mimic. In this way, the effect of neighboring "substrate-loaded" subunits on a substrate-unbound protomer can be evaluated. Activity assays show that GLF-CMK binding to ~50% of the subunits provides optimal activity, although lower than what was observed for Bz-LL stimulation (57). Additionally, a pair of controls was generated in which all MtClpP1 subunits were c) GLF-CMK bound or d) in which excess Bz-LL was added. NMR spectra recorded on MtClpP1P2 samples with MtClpP1 labeled as indicated in Fig. 4*C* are shown for selected methyl groups in Fig. 4 *D*–*G*. As the apo- and Bz-LL–bound conformations correspond to T (inactive) and R (active) states of the enzyme, respectively (Fig. 4 *A*, *Right*, and Fig. 4*B*), separate sets of peaks are observed in each case (Fig. 4 *D*–*G*, contours denoted by "a" and "d" for apo- and Bz-LL–bound molecules). Covalent attachment of GLF-CMK to all seven subunits of MtClpP1 (contours indicated by "c") leads to pairs of peaks in NMR spectra, resulting from a shift in the MtClpP1P2 T<>R equilibrium to approximately equimolar amounts of each state. A similar set of pairs of peaks (T and R states) was observed in spectra reporting on the conformation of an unbound MtClpP1 subunit surrounded by GLF-CMK protomers (contours indicated by "b"). Thus, the bound subunits allosterically regulate the conformation of unbound neighbors, providing support for the MWC mechanism of allostery in which protomers of an oligomer are all in the same state irrespective of whether they are ligated.

The allostery between subunits within a ClpP ring provides a framework for understanding some of the differences in the catalytic properties of diverse ClpPs. Although nonsaturating amounts of an active site inhibitor bortezomib are necessary to activate *Thermus thermophilus* ClpP (TtClpP) (58), a role that Bz-LL plays for MtClpP1P2 (56), *Escherichia coli* ClpP has high activity without binding any activator (54). In the context of the MWC model, ~0.0005% (L=[T]/[R]~$2 \times 10^5$) and 20% (L~4) of MtClpP1P2 and TtClpP apo particles are in the active, R state, respectively, in the absence of activator or substrate, while close to all of *E. coli* ClpP molecules are active ($L \ll 1$). The T<>R equilibrium thus regulates ClpP function by controlling basal enzyme activity and the catalytic response to added substrate.

**p97/VCP Function Is Modulated through Cooperative Subunit Dynamics.** The p97/VCP homeostasis protein (referred to in what follows as p97, 540 kDa) participates in a wide range of cellular processes including proteasomal and lysosomal protein degradation, removal of proteinaceous components from membranes or other molecular complexes, membrane fusion, and

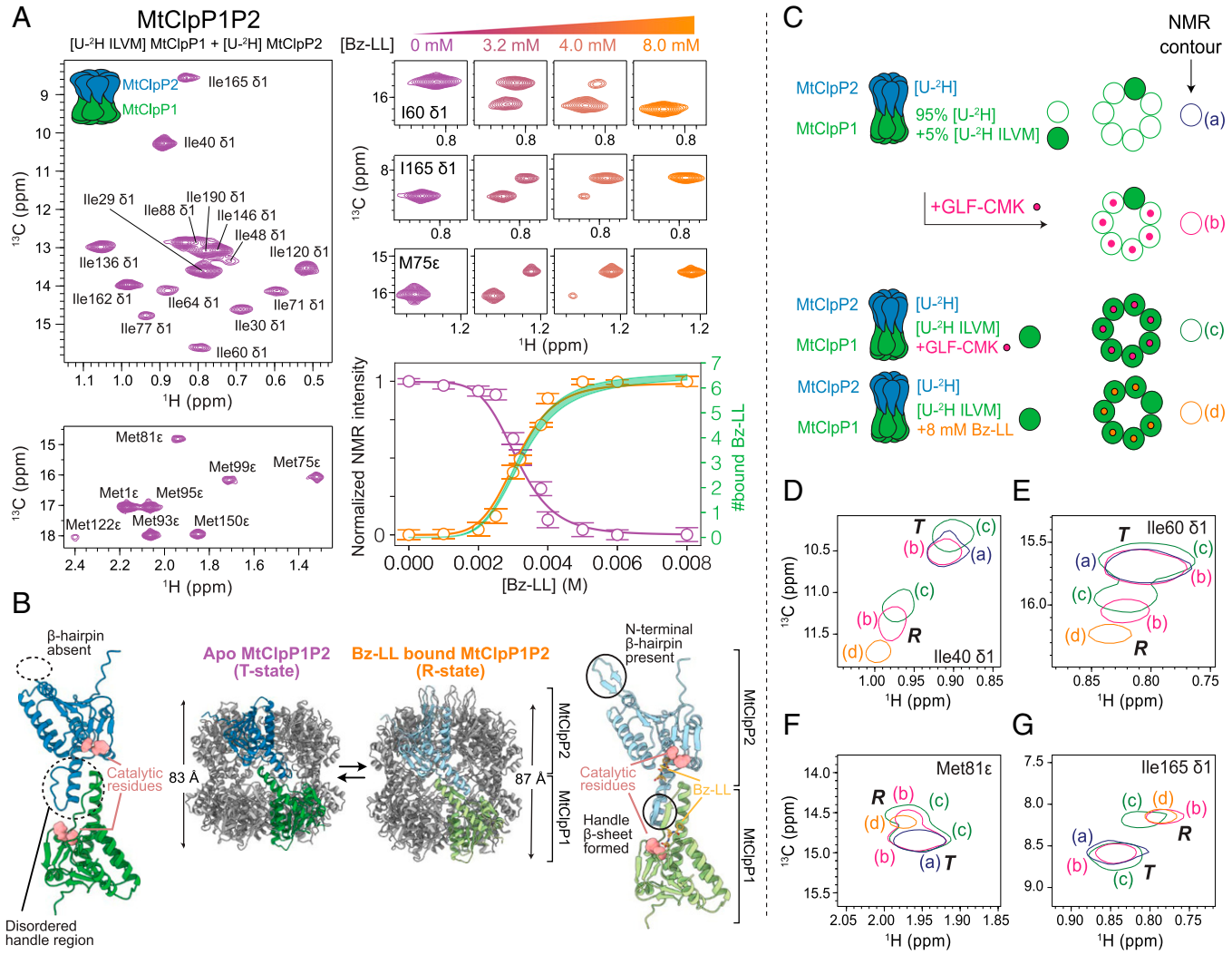


**Fig. 4.** An allosteric switch regulates MtClpP1P2 function (57). (*A*) Selected Ile (*Top Left*) and Met (*Bottom Left*) regions from a $^{13}C$-$^{1}H$ HMQC spectrum recorded of MtClpP1P2 where MtClpP1 and MtClpP2 are U-$^{2}H$, ILVM-$^{13}CH_3$– (NMR-active), and U-$^{2}H$– (NMR-invisible) labeled, respectively. Assignments are indicated. (*Top Right*) Titration of MtClpP1P2 with Bz-LL showing selective correlations from MtClpP1 that disappear (purple) and reappear (orange) with the addition of increasing peptide. (*Bottom Right*) Normalized intensities from a set of apo- and Bz-LL–bound MtClpP1 resonances (circles) obtained from the titration and fit to a Hill model (solid curve); a Hill coefficient of ~7 is obtained. The green curve plots the number of bound Bz-LL molecules to MtClpP1 based on fits using a modified MWC model (57). At [Bz-LL] = 8 mM, approximately six of the seven active sites are Bz-LL bound. Curve thickness indicates error of 2 SD from the mean. (*B*) MtClpP1P2 is in equilibrium between an inactive, apo (T) state [PDB ID: 6VGK (57)] and an active (R) conformation [depicted by the Bz-LL–bound structure; PDB ID: 5DZK (55)] with distinct structural differences (ribbon diagrams to the left and right of the cartoons). (*C*) Schematic of the labeling procedure used to probe intraring allostery in MtClpP1 rings of MtClpP1P2 complexes; details of sample production are provided in Vahidi et al. (57). (*Top Row*) MtClpP1 heptamers were prepared by mixing 95% WT U-$^{2}H$ protomers (NMR invisible; denoted by open circles) with 5% U-$^{2}H$ ILVM-subunits (NMR visible; green filled circles) to produce a series of rings with the major configuration (*a*) highlighted. (*Second Row*) A second set of molecules was produced by reaction with GLF-CMK to generate particles indicated by *b*. The resulting NMR spectra report on how the conformation of a WT particle is affected by GLF-CMK–bound neighbors (mimic of substrate bound subunits). (*Third Row*) WT particles are prepared for which each MtClpP1 subunit is NMR labeled and covalently modified by GLF-CMK. (*Fourth Row*) Particles are modified by addition of excess (8 mM) Bz-LL. Based on the fitted binding parameters using a modified MWC model, ~6 of the 7 subunits are expected to be Bz-LL bound (see *A*, *Bottom Right*, green curve). (*D*–*G*). Selected regions of $^{13}C$-$^{1}H$ HMQC spectra recorded on complexes indicated in *C*; each cross peak is color coded and further denoted by *a*–*d* to indicate the particle from which it derives (see *C*, *Right*). *R* and *T* denote relaxed (active) and tense (inactive) states of MtClpP1P2 that were assigned based on the fact that apo- and Bz-LL–bound states are inactive and active, respectively (56). Adapted from ref. 57.

regulation of cell signaling, proliferation, and survival (59–62). More than 20 different missense mutations of the p97 gene have been linked to a range of neurodegenerative diseases that have been referred to collectively as multisystem proteinopathy type 1 (MSP1) (63). Detailed X-ray (64) and now cryo-EM (65) studies of p97 have established that it forms a double-ring architecture composed of six copies of an 800-residue polypeptide chain. Each chain, in turn, is divided into three domains, including an N-terminal domain (NTD) and a pair of nucleotide binding domains, D1 and D2, with the D1 and D2 structures forming the top and bottom rings of the enzyme, respectively. The NTDs interact with a large number of adaptor proteins so as to regulate p97 function both spatially and temporally, while adenosine triphosphate (ATP) hydrolysis by D1 and D2 provides the energy for various p97 mechanical functions, including its unfoldase activity (66, 67). Structural studies have shown that upon ATP hydrolysis, the NTDs change orientation from a position in which they are aligned above the plane of the D1 ring (ATP bound) to coplanar (ADP bound) with the ring (64, 65). Fig. 5*A* illustrates this clearly with a pair of

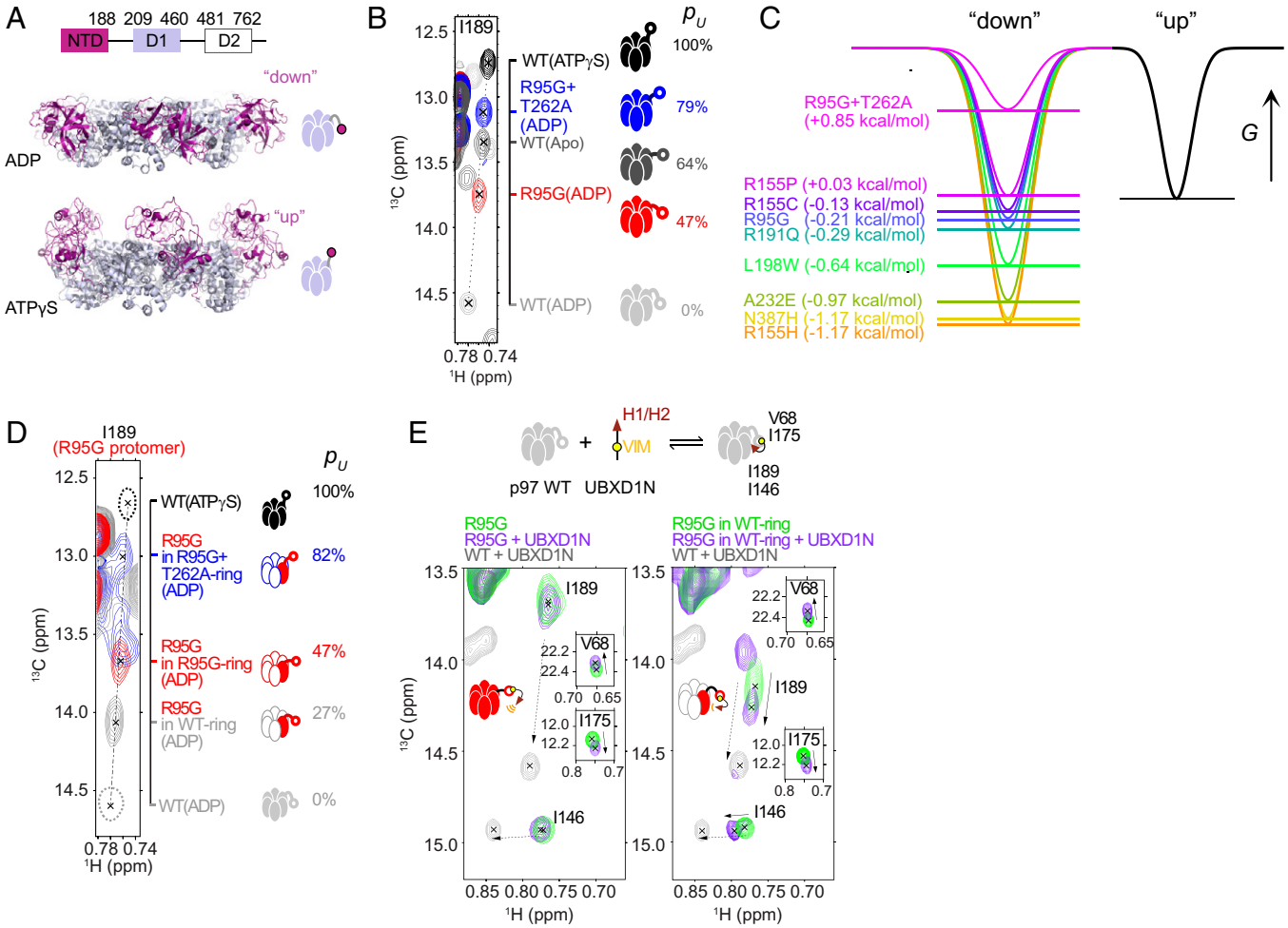


**Fig. 5.** Cooperative p97 subunit dynamics (68). (*A*) Domain organization of each subunit and cryo-EM structures of ADP- [PDB ID: 5FTK (65)] and ATPγS- [a slowly hydrolyzing analog of ATP; PDB ID: 5FTN (65)] bound states of ND1L. Cartoon depictions of ND1L are shown to the right with a single NTD highlighted in either the down (ADP) or up (ATPγS) conformation. (*B*) Superposition of a selected region of $^{13}$C-$^1$H HMQC spectra of ND1L recorded with different bound nucleotides or disease mutants, focusing on I189δ1 that reports on the up/down NTD equilibrium. The up/down NTD equilibrium constant for a particular ND1L particle can be determined from the position of the I189δ1 cross peak (denoted by crosshairs) relative to WT(ADP) and WT(ATPγS); shown on the right are fractional populations of the up NTD conformation, $p_U$, for each particle. Peaks for I189δ1 are color coded to denote the particle from which they derive. (*C*) Two-state free energy diagrams for the NTD up/down equilibrium for a number of disease mutants, as indicated. The mutations are assumed to affect the energy of the down state, exclusively. (*D*) Superposition of a spectral region from $^{13}$C-$^1$H HMQC datasets, focusing on I189δ1, and recorded on a series of particles prepared by mixing NMR visible R95G subunits (15%) with NMR invisible WT, R95G, or R95G + T262A subunits (85%) color-coded according to the particle from which the cross peak derives. Values of $p_U$ are indicated to the right. (*E*, *Top*) Schematic of the two-pronged binding interaction of UBXD1N [N-terminal region of the UBXD1 adaptor, residues 1–133, that regulates lysosomal-based protein degradation (73)] and WT p97. V68 and I175 report on VIM binding, while I146 and I189 are sensitive to binding of the H1/H2 domain, as indicated. (*Bottom Left*) The VIM domain binds to R95G ND1L, as evidenced from V68 and I175 peak titrations upon addition of UBXD1N (green to purple), but neither of the peaks derived from I146δ1 or I189δ1 move, indicating that the H1/H2 prong of the interaction is defective. (*Bottom Right*) In contrast, when the R95G subunit is surrounded by WT neighbors, binding of the H1/H2 UBXD1N domain is partially restored, as evidenced by movement of I146 and I189 peaks upon addition of UBXD1N. Note that only the R95G subunit (red) is NMR active and hence the spectrum reports directly on interactions involving it. Peaks from a complex of UBXD1N and WT ND1L (gray) are shown as a reference. Adapted partly from ref. 68.

cryo-EM structures comprising the NTDs (purple) and D1 rings (gray) of p97 (D2 omitted from diagram). A 320-kDa construct, referred to as ND1L, in which the D2 ring is removed, has been used in a variety of different structural studies, including the NMR work discussed here (68). In what follows, we will denote the down and up NTD conformations (relative to the D1 ring) using schematics that are illustrated to the right of the structures highlighted in Fig. 5*A*, where only one of the six NTDs is shown for simplicity (purple ball). As might be expected on the basis of the large structural changes between the ATP- and adenosine diphosphate (ADP)–bound p97 (Fig. 5*A*), NMR spectra of the two nucleotide-loaded states are distinct, with key residues, such as I189, serving as very sensitive reporters of the up/down NTD equilibrium (69). This is illustrated in Fig. 5*B*, where a superposition of HMQC spectra is shown, focusing on the I189δ1 methyl cross peak. Notably, in the ADP-loaded form of p97, the I189 peak from the severe R95G disease mutant [R95G(ADP), red] is not at the same position as in the WT(ADP) state (gray) but rather approximately equidistant between the WT(ADP) and WT(ATP) I189δ1 peaks. An ND1L construct harboring a pair of disease mutations (R95G + T262A) shows a further shift toward the WT(ATP) position. In an extensive NMR study involving a large series of disease mutants, in which all of the ND1L particles were in the ADP-loaded form, and hence the NTDs would be expected to be in the D1 coplanar state (Fig. 5*A*), it was found that the corresponding peaks for I189 (and many other methyl probes) were positioned linearly between the

WT(ADP)—all NTDs down—and WT(ATP)—all NTDs up—"boundaries" (gray and black peaks in Fig. 5*B*) (69). This implies that, as a function of disease mutation, the NTDs are not all "down" in the ADP state but, rather, progressively move up, as illustrated in Fig. 5 *B*, *Right*, with the fractional population of the up state, $P_U$, calculated simply from the relative position of the peak with respect to the corresponding WT(ADP) and WT(ATP) correlations. It is worth emphasizing that the NTDs are not statically oriented with respect to the D1 ring but rather, in the simplest model consistent with the NMR data, exchange between down and up conformations with a rate in excess of 15,000/s (70). Disease mutations, thus, lead to changes in the ND1L free-energy landscape, as highlighted in Fig. 5*C*, where a set of energy surfaces is shown for some of the disease mutants that have been studied (all ADP bound). Small changes in relative free energies between states can lead to significant population shifts. Interestingly, the up/down NTD equilibrium for less severe disease mutants (ADP bound), such as R155H, is less aberrant (i.e., more down) than for more severe mutants such as R95G (69).

MSP1 patients have one WT and one disease-mutated allele, and it has been shown that p97 hexamers can contain mixtures of WT and diseased protomers (71, 72). As most of the biophysical and biological studies have focused on particles in which all six protomers are homogeneous, it was of interest to see how the aberrant up/down equilibrium associated with disease would be affected in the context of a ring comprising a mixture of disease and WT neighboring subunits. This is a difficult problem to address using X-ray crystallography or cryo-EM because the difference between disease and WT subunits lies in one amino acid out of 800. Yet a mixing strategy whereby WT and mutant subunits are unfolded and subsequently refolded and in which one of the protomer types is U-$^{2}$H, ILVM-$^{13}CH_3$ labeled (i.e., NMR visible) and the other U-$^{2}$H (NMR invisible) provides an avenue for addressing this question (68). A number of different ND1L particles were prepared by mixing NMR-visible R95G subunits (fractional amount of 15%) with invisible WT, R95G, or R95G + T262A subunits, and a superposition of the resulting spectra, again focusing on I189, is shown in Fig. 5*D*. The skewed up/down equilibrium of the R95G protomer is significantly affected by the neighboring subunits; if the neighbors have an intrinsic tendency to be down (for example, neighbors are WT), $P_U$ for the R95G protomer decreases, with $P_U$ increasing if the surrounding protomers have a higher propensity to be up than that of the R95G subunit (for example, R95G + T262A).

In order to address how changes to the up/down NTD equilibrium affect downstream processes, studies of the interactions between the UBXD1 adaptor molecule and p97 were undertaken. UBXD1 primes p97 for its role in lysosomal degradation, a pathway affected by MSP1 disease mutations (73). NMR spectra were, therefore, recorded of p97 in the absence and presence of an N-terminal 133-residue fragment of UBXD1 (UBXD1N) that contains two p97 binding motifs, VIM and H1/H2 (74, 75). Studies of an NTD construct including the linker connecting NTD with D1 (but lacking D1 or D2) indicate that VIM binds to the NTD while H1/H2 contacts the linker (75). Methyl-TROSY–based NMR studies of UBXD1N binding in the context of ND1L confirm this result and, furthermore, show that the H1/H2 domain binds at the interface between NTD and D1 domains for the WT, ADP-loaded form of p97 (Fig. 5 *E*, *Top*) (69). As such, binding of UBXD1 to WT p97-ADP generates a locked-down structure in which all the NTDs are coplanar with the D1 ring. As a large number of methyl groups are available, each step of the two-pronged binding process can be monitored, with V68 and I175 reporting on VIM binding and I189 and I146 probing the up/down status of a given protomer (Fig. 5*E*). Since binding of H1/H2 would be expected to shift the up/down equilibrium to the down conformation, perhaps only partially for disease mutants, the spectral positions of I189 and I146 methyl groups can be used as reporters of the interaction between p97 and H1/H2 (69). In an initial study focusing on ND1L in which all protomers contained the R95G mutation, binding of the VIM domain was confirmed (Fig. 5 *E*, *Left*), as the peak positions for V68 and I175 were shifted upon addition of UBXD1N (green to purple). However, no shifts in the peak positions for I189 and I146 were observed, indicating that contacts with H1/H2 are not formed (68). Thus, the aberrant up/down NTD equilibrium leads to a complex in which only one of the prongs of the WT UBXD1–p97 interaction is engaged. Recall that when an R95G protomer is surrounded by WT neighbors, its up/down NTD equilibrium is partially reverted toward the down conformational state (Fig. 5*D*). A prediction might be, therefore, that the two-pronged UBXD1-binding interaction would be restored to an extent in this case. This is clearly observed in NMR spectra (Fig. 5 *E*, *Right*) in which there is a shifting of the I189 and I146 correlations upon addition of adaptor (green to purple), indicating that contacts involving H1/H2 are at least partially formed. To "guide the eye," spectra of WT ND1L with UBXD1 are also shown (gray), indicating the positions of the I189 and I146 δ1-methyl groups in the fully down conformation. Interestingly, the perturbed equilibrium in the R155C severe disease mutant can be restored through the addition of an N387C mutation, with the position of the NTDs collinear with the D1 ring, as for the WT protein. As might be expected, therefore, the R155C + N387C revertant recovers WT UBXD1N-binding affinity (70). Small-molecule drugs that potentially could push the up/down NTD equilibrium even moderately toward the down NTD state might, therefore, play an important therapeutic role in mitigating the effects of the disease mutations.

## Concluding Remarks

Solution NMR spectroscopy offers unique opportunities to investigate how biomolecules carry out their essential cellular activities. In the context of studies of molecular machines described here, NMR can play an important role in providing insights into functional dynamics (29, 76–79). As such, it is highly complementary to X-ray and cryo-EM approaches that provide detailed yet static structures of these large macromolecules. In this perspective, several homo-oligomeric systems have been discussed, focusing on the role(s) of interprotomer communication in regulating function. In particular, we have described an NMR-based strategy by which the structural dynamics of individual protomers can be assessed in the context of different neighboring subunits. By judicious choice of NMR-labeling schemes and relative amounts of the different protomers that are mixed, direct experimental evidence for the presence or absence of interprotomer allostery can be provided, and atomic resolution insights are often available. Such an approach is complementary and, in many cases, superior to the established way of demonstrating allostery involving fits of models to data for which conclusions are based simply on goodness-of-fit values. In the case of the *T. acidophilum* 20S proteasome CP, the in/out gating equilibrium of a protomer was shown not to depend on neighboring subunits so long as there was space in the proteasome lumen to accommodate the gate (53). Here, the gates are comprised of ~10 to 15 residues and are intrinsically disordered. Their small size and high flexibility challenges studies using other atomic resolution methods. In contrast, NMR studies of p97 clearly establish a coupling between protomers in this hexameric machine (68). Single-point mutants that give rise to a series of neurodegenerative diseases were shown to affect the up/down NTD equilibrium and, as a consequence, interactions with the UBXD1 adaptor molecule (69). Dissecting how the aberrant up/down

equilibrium of one protomer is influenced by neighboring subunits in the context of heterogeneous p97 particles composed of combinations of WT and mutated protomers is complicated by the fact that each subunit differs by, at most, a single amino acid. Yet the NMR approach described here readily identifies intersubunit allostery. Similarly, strong interprotomer allostery was shown to be critical for regulating HtrA2 function, whereby subunits undergo a concerted structural change from an open, inactive state to one which is catalytically active (40). Dysregulation of this network, at least in the context of an S276C disease mutant, results in a neurodegenerative condition in mice that serves as a model for Parkinson's disease (41). The importance of NMR-based experiments focusing on a single subunit in the context of an oligomeric system is made clear by the fact that WT and S276C HtrA2 have very similar structures (42) that, unfortunately, therefore, offer little insight into the mechanism of disease. Similar to p97 and HtrA2, ligated subunits in the MtClpP1P2 protease affect the conformation of unbound protomers, providing strong experimental support of the MWC model that was used to analyze activity profiles (57). The potential of NMR spectroscopy to contribute to studies of molecular machines is made clear by these examples. It is likely that as the number of structures of this class of molecules continues to grow, the importance of NMR will increase in parallel, providing a key link between static structure and dynamic function.

**Data Availability.** All data are included in the article.

**ACKNOWLEDGMENTS.** Y.T. is supported through a Japan Society for the Promotion of Science Overseas Research Fellowship and an Uehara Memorial Foundation postdoctoral fellowship. This research was funded through 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. L.E.K. is grateful to all trainees in his laboratory over the past three decades. In particular, Vitali Tugarinov (methyl-TROSY), Rui Huang (proteasome), Rui Huang and Anne Schuetz (p97/VCP), and Siavash Vahidi and Zev Ripstein (MtClpP1P2) are thanked for their efforts in projects highlighted in this research.

1. J. C. Kendrew *et al.*, A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662–666 (1958).
2. H. Muirhead, M. F. Perutz, Structure of hæemoglobin: A three-dimensional Fourier synthesis of reduced human haemoglobin at 5.5 Å resolution. *Nature* **199**, 633–638 (1963).
3. M. P. Williamson, T. F. Havel, K. Wüthrich, Solution conformation of proteinase inhibitor IIA from bull seminal plasma by $^{1}$H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **182**, 295–315 (1985).
4. E. Nogales, S. H. W. Scheres, Cryo-EM: A unique tool for the visualization of macromolecular complexity. *Mol. Cell* **58**, 677–689 (2015).
5. M. A. Herzik Jr., M. Wu, G. C. Lander, High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nat. Commun.* **10**, 1032 (2019).
6. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
7. M. Baek *et al.*, Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
8. K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated $T_2$ 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. U.S.A.* **94**, 12366–12371 (1997).
9. K. Pervushin, Impact of transverse relaxation optimized spectroscopy (TROSY) on NMR as a technique in structural biology. *Q. Rev. Biophys.* **33**, 161–197 (2000).
10. N. Tjandra, A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278**, 1111–1114 (1997).
11. P. Schanda, E. Kupće, B. Brutscher, SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR* **33**, 199–211 (2005).
12. R. Sprangers, L. E. Kay, Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **445**, 618–622 (2007).
13. I. Gelis *et al.*, Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* **131**, 756–769 (2007).
14. N. J. Anthis, G. M. Clore, Visualizing transient dark states by NMR spectroscopy. *Q. Rev. Biophys.* **48**, 35–116 (2015).
15. P. Vallurupalli, A. Sekhar, T. Yuwen, L. E. Kay, Probing conformational dynamics in biomolecules via chemical exchange saturation transfer: A primer. *J. Biomol. NMR* **67**, 243–271 (2017).
16. T. Mittag, L. E. Kay, J. D. Forman-Kay, Protein dynamics and conformational disorder in molecular recognition. *J. Mol. Recognit.* **23**, 105–116 (2010).
17. V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, L. E. Kay, Cross-correlated relaxation enhanced $^{1}$H-$^{13}$C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125**, 10420–10428 (2003).
18. A. Abragam, *Principles of Nuclear Magnetism* (Clarendon Press, Oxford, UK, 1961).
19. C. Fernández, G. Wider, TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr. Opin. Struct. Biol.* **13**, 570–580 (2003).
20. V. Tugarinov, L. E. Kay, An isotope labeling strategy for methyl TROSY spectroscopy. *J. Biomol. NMR* **28**, 165–172 (2004).
21. R. Kerfah, M. J. Plevin, R. Sounier, P. Gans, J. Boisbouvier, Methyl-specific isotopic labeling: A molecular tool box for solution NMR studies of large proteins. *Curr. Opin. Struct. Biol.* **32**, 113–122 (2015).
22. R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon Press, Oxford, UK, 1987).
23. J. H. Prestegard, D. M. Grant, Characterization of anisotropic motion in fatty acid micelles by analysis of transverse relaxation in an AX2 nuclear spin system. *J. Am. Chem. Soc.* **100**, 4664–4668 (1978).
24. L. K. Nicholson *et al.*, Dynamics of methyl groups in proteins as studied by proton-detected $^{13}$C NMR spectroscopy. Application to the leucine residues of staphylococcal nuclease. *Biochemistry* **31**, 5253–5263 (1992).
25. L. E. Kay, D. R. Muhandiram, N. A. Farrow, Y. Aubin, J. D. Forman-Kay, Correlation between dynamics and high affinity binding in an SH2 domain interaction. *Biochemistry* **35**, 361–368 (1996).
26. A. Bax, R. H. Griffey, B. L. Hawkins, Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J. Magn. Reson.* **55**, 301–315 (1983).
27. L. Mueller, P. Alto, Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *J. Am. Chem. Soc.* **101**, 4481–4484 (1979).
28. J. E. Ollerenshaw, V. Tugarinov, L. E. Kay, Methyl TROSY: Explanation and experimental verification. *Magn. Reson. Chem.* **41**, 843–852 (2003).
29. S. Schütz, R. Sprangers, Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems. *Prog. Nucl. Magn. Reson. Spectrosc.* **116**, 56–84 (2020).
30. A. M. Ruschak, L. E. Kay, Methyl groups as probes of supra-molecular structure, dynamics and function. *J. Biomol. NMR* **46**, 75–87 (2010).
31. P. Gans *et al.*, Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew. Chem. Int. Ed. Engl.* **49**, 1958–1962 (2010).
32. L. Faccio *et al.*, Characterization of a novel human serine protease that has extensive homology to bacterial heat shock endoprotease HtrA and is regulated by kidney ischemia. *J. Biol. Chem.* **275**, 2581–2588 (2000).
33. C. W. Gray *et al.*, Characterization of human HtrA2, a novel serine protease involved in the mammalian cellular stress response. *Eur. J. Biochem.* **267**, 5699–5710 (2000).
34. Y. Suzuki *et al.*, A serine protease, HtrA2, is released from the mitochondria and interacts with XIAP, inducing cell death. *Mol. Cell* **8**, 613–621 (2001).
35. L. Vande Walle, M. Lamkanfi, P. Vandenabeele, The mitochondrial serine protease HtrA2/Omi: An overview. *Cell Death Differ.* **15**, 453–460 (2008).
36. K. M. Strauss *et al.*, Loss of function mutations in the gene encoding Omi/HtrA2 in Parkinson's disease. *Hum. Mol. Genet.* **14**, 2099–2111 (2005).
37. H. Unal Gulsuner *et al.*, Mitochondrial serine protease HTRA2 p.G399S in a kindred with essential tremor and Parkinson disease. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18285–18290 (2014).
38. W. Li *et al.*, Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. *Nat. Struct. Biol.* **9**, 436–441 (2002).
39. Y. Toyama, R. W. Harkness, T. Y. T. Lee, J. T. Maynes, L. E. Kay, Oligomeric assembly regulating mitochondrial HtrA2 function as examined by methyl-TROSY NMR. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025022118 (2021).
40. Y. Toyama, R. W. Harkness, L. E. Kay, Dissecting the role of interprotomer cooperativity in the activation of oligomeric high-temperature requirement A2 protein. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2111257118 (2021).
41. J. M. Jones *et al.*, Loss of Omi mitochondrial protease activity causes the neuromuscular disorder of *mnd2* mutant mice. *Nature* **425**, 721–727 (2003).
42. A. R. Wagh, K. Bose, Structural basis of inactivation of human counterpart of mouse motor neuron degeneration 2 mutant in serine protease HtrA2. *Biosci. Rep.* **38**, BSR20181072 (2018).
43. S. Bhattacharyya, H. Yu, C. Mim, A. Matouschek, Regulated protein turnover: Snapshots of the proteasome in action. *Nat. Rev. Mol. Cell Biol.* **15**, 122–133 (2014).
44. A. J. Marques, R. Palanimurugan, A. C. Matias, P. C. Ramos, R. J. Dohmen, Catalytic mechanism and assembly of the proteasome. *Chem. Rev.* **109**, 1509–1536 (2009).
45. A. L. Goldberg, Functions of the proteasome: From protein degradation and immune surveillance to cancer therapy. *Biochem. Soc. Trans.* **35**, 12–17 (2007).
46. J. Löwe *et al.*, Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 A resolution. *Science* **268**, 533–539 (1995).
47. M. Groll *et al.*, Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature* **386**, 463–471 (1997).

48. D. Forouzan *et al.*, The archaeal proteasome is regulated by a network of AAA ATPases. *J. Biol. Chem.* **287**, 39254–39262 (2012).
49. M. Bolten *et al.*, Structural analysis of the bacterial proteasome activator Bpa in complex with the 20S proteasome. *Structure* **24**, 2138–2151 (2016).
50. M. Schmidt *et al.*, The HEAT repeat protein Blm10 regulates the yeast proteasome by capping the core particle. *Nat. Struct. Mol. Biol.* **12**, 294–303 (2005).
51. G. Ben-Nissan, M. Sharon, Regulating the 20S proteasome ubiquitin-independent degradation pathway. *Biomolecules* **4**, 862–884 (2014).
52. T. L. Religa, R. Sprangers, L. E. Kay, Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* **328**, 98–102 (2010).
53. R. Huang, F. Pérez, L. E. Kay, Probing the cooperativity of *Thermoplasma acidophilum* proteasome core particle gating by NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9846–E9854 (2017).
54. T. Akopian *et al.*, Cleavage specificity of *Mycobacterium tuberculosis* ClpP1P2 protease and identification of novel peptide substrates and boronate inhibitors with antibacterial activity. *J. Biol. Chem.* **290**, 11008–11020 (2015).
55. M. Li *et al.*, Structure and functional properties of the active form of the proteolytic complex, ClpP1P2, from *Mycobacterium tuberculosis. J. Biol. Chem.* **291**, 7465–7476 (2016).
56. T. Akopian *et al.*, The active ClpP protease from *M. tuberculosis* is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO J.* **31**, 1529–1541 (2012).
57. S. Vahidi *et al.*, An allosteric switch regulates *Mycobacterium tuberculosis* ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5895–5906 (2020).
58. J. Felix *et al.*, Mechanism of the allosteric activation of the ClpP protease machinery by substrates and active-site inhibitors. *Sci. Adv.* **5**, eaaw3818 (2019).
59. H. Meyer, M. Bug, S. Bremer, Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat. Cell Biol.* **14**, 117–123 (2012).
60. K. Yamanaka, Y. Sasagawa, T. Ogura, Recent advances in p97/VCP/Cdc48 cellular functions. *Biochim. Biophys. Acta* **1823**, 130–137 (2012).
61. Y. Ye, W. K. Tang, T. Zhang, D. Xia, A mighty "protein extractor" of the cell: Structure and function of the p97/CDC48 ATPase. *Front. Mol. Biosci.* **4**, 39 (2017).
62. A. Stolz, W. Hilt, A. Buchberger, D. H. Wolf, Cdc48: A power machine in protein degradation. *Trends Biochem. Sci.* **36**, 515–523 (2011).
63. G. D. J. Watts *et al.*, Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat. Genet.* **36**, 377–381 (2004).
64. X. Zhang *et al.*, Structure of the AAA ATPase p97. *Mol. Cell* **6**, 1473–1484 (2000).
65. S. Banerjee *et al.*, 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science* **351**, 871–875 (2016).
66. M. M. Olszewski, C. Williams, K. C. Dong, A. Martin, The Cdc48 unfoldase prepares well-folded protein substrates for degradation by the 26S proteasome. *Commun. Biol.* **2**, 29 (2019).
67. N. O. Bodnar, T. A. Rapoport, Molecular mechanism of substrate processing by the Cdc48 ATPase complex. *Cell* **169**, 722–735.e9 (2017).
68. R. Huang, Z. A. Ripstein, J. L. Rubinstein, L. E. Kay, Cooperative subunit dynamics modulate p97 function. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 158–167 (2019).
69. A. K. Schuetz, L. E. Kay, A dynamic molecular basis for malfunction in disease mutants of p97/VCP. *eLife* **5**, e20143 (2016).
70. A. K. Schütz, E. Rennella, L. E. Kay, Exploiting conformational plasticity in the AAA+ protein VCP/p97 to modify function. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6822–E6829 (2017).
71. C. C. Weihl, S. Dalal, A. Pestronk, P. I. Hanson, Inclusion body myopathy-associated mutations in p97/VCP impair endoplasmic reticulum-associated degradation. *Hum. Mol. Genet.* **15**, 189–199 (2006).
72. S. Dalal, M. F. N. Rosser, D. M. Cyr, P. I. Hanson, Distinct roles for the AAA ATPases NSF and p97 in the secretory pathway. *Mol. Biol. Cell* **15**, 637–648 (2004).
73. D. Ritz *et al.*, Endolysosomal sorting of ubiquitylated caveolin-1 is regulated by VCP and UBXD1 and impaired by VCP disease mutations. *Nat. Cell Biol.* **13**, 1116–1123 (2011).
74. C. Stapf, E. Cartwright, M. Bycroft, K. Hofmann, A. Buchberger, The general definition of the p97/valosin-containing protein (VCP)-interacting motif (VIM) delineates a new family of p97 cofactors. *J. Biol. Chem.* **286**, 38670–38678 (2011).
75. F. Trusch *et al.*, The N-terminal region of the ubiquitin regulatory X (UBX) domain-containing protein 1 (UBXD1) modulates interdomain communication within the valosin-containing protein p97. *J. Biol. Chem.* **290**, 29414–29427 (2015).
76. R. Rosenzweig, L. E. Kay, Bringing dynamic molecular machines into focus by methyl-TROSY NMR. *Annu. Rev. Biochem.* **83**, 291–315 (2014).
77. Y. Jiang, C. G. Kalodimos, NMR studies of large proteins. *J. Mol. Biol.* **429**, 2667–2676 (2017).
78. T. R. Alderson, L. E. Kay, NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. *Cell* **184**, 577–595 (2021).
79. G. Mas *et al.*, Structural investigation of a chaperonin in action reveals how nucleotide binding regulates the functional cycle. *Sci. Adv.* **4**, eaau4196 (2018).
80. A. Förster, E. I. Masters, F. G. Whitby, H. Robinson, C. P. Hill, The 1.9 A structure of a proteasome-11S activator complex and implications for proteasome-PAN/PA700 interactions. *Mol. Cell* **18**, 589–599 (2005).