JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Perspective

pubs.acs.org/JACS

Solution NMR Spectroscopy Provides an Avenue for the Study of Functionally Dynamic Molecular Machines: The Example of Protein Disaggregation

Rina Rosenzweig*,† and Lewis E. Kay*,†,‡

†Departments of Molecular Genetics, Biochemistry, and Chemistry, The University of Toronto, Toronto, Ontario, Canada M5S 1A8
‡Program in Molecular Structure and Function, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8

ABSTRACT: Solution-based NMR spectroscopy has been an important tool for studying the structure and dynamics of relatively small proteins and protein complexes with aggregate molecular masses under approximately 50 kDa. The development of new experiments and labeling schemes, coupled with continued improvements in hardware, has significantly reduced this size limitation, enabling atomic-resolution studies of molecular machines in the 1 MDa range. In this Perspective, some of the important advances are highlighted in the context of studies of molecular chaperones involved in protein disaggregation. New insights into the structural biology of disaggregation obtained from NMR studies are described, focusing on the unique capabilities of the methodology for obtaining atomic-resolution descriptions of dynamic systems.

1. INTRODUCTION

Cellular function is critically dependent on a myriad of interprotein interactions involving often highly dynamic components.[1−3] In recent years, considerable progress has been made in identifying many of the pathways in which these proteins participate and in establishing the sequential flow of information from one molecule to the next;[2,4,5] however, much remains to be elucidated at the molecular level. A major difficulty lies in the fact that the complexes formed can be large and transient, precluding their detailed study by many of the structural biology tools that are typically used for stable molecules and interactions.[6,7] Moreover, many of the molecular players undergo significant conformational changes throughout the pathway, further complicating their characterization.[1,2] The tools for studying the dynamics of proteins, especially those that involve sparsely populated and transiently formed states, remain incomplete. Understanding the role of molecular dynamics in function in a great many of cases has, thus, remained an elusive goal of structural biology and molecular biophysics.

Recent developments in solution-based NMR spectroscopy, however, are beginning to change this paradigm. The emergence of new labeling schemes, whereby proteins are selectively labeled with probes of interest, has led to spectral simplification and importantly, in many cases, significant improvements in spectral sensitivity and resolution.[8−12] Synergistically, new pulse schemes have been developed that exploit the labeling patterns in ways that further increase sensitivity and resolution.[13−18] Perhaps the most spectacular example of this is in the use of transverse relaxation optimized spectroscopy (TROSY) to study high-molecular-weight proteins and protein complexes, first described by Pervushin, Wuthrich, and co-workers.[19] Here a protein labeling scheme typically involving uniform ^{15}N incorporation along with a high level of deuteration at non-labile sites is utilized[20] along with pulse sequences that select the slowly relaxing components of magnetization that arise from the effective cancelation of local dipolar and chemical shift anisotropy fields at amide positions. Further applications involving aromatic side chains in proteins and nucleic acids have also been described.[21−23] Methods for labeling proteins with $^{13}CH_3$ methyl groups in an otherwise highly deuterated environment[10,24−32] and experiments that exploit a methyl-TROSY effect have emerged,[14] and a significant number of applications of this technology to studies of molecular machines have been reported in the past several years.[27,30,33−44]

Sensitivity and resolution limitations also continue to be addressed with the development of new hardware, including ultra-high-field magnets and cryogenically cooled probes. Taken together, this has led to sensitivity gains of over a factor of 30, or savings in experimental time of close to 3 orders of magnitude, between the high-field instruments of 25 years ago (500 MHz ^1H frequency) and those of today (1 GHz).

A major strength of solution NMR has always been its sensitivity to a wide spectrum of motion, with spin relaxation experiments probing dynamics ranging from picosecond time scales to seconds.[45−47] Of particular importance in this regard is that atomic resolution dynamics information can be obtained, providing insight into how structure changes with time. In many cases, conformational transitions involving rare states of both proteins and nucleic acids[48−50] can now be studied, yielding detailed structural information about these elusive conformers.

With the improvements in labeling, pulse sequences, and hardware that have led to impressive sensitivity and resolution gains, and with the continued development of approaches for studies of molecular dynamics and conformational transitions, solution NMR has now reached a phase where detailed applications to complex biological systems are feasible.[7] Here we focus on one system, involving the Hsp70/Hsp104 families of molecular chaperones, to illustrate how solution NMR can be used to obtain atomic-resolution details on a highly "dynamic" pathway involving the interaction of multiple protein components. We have deliberately chosen a single system so that sufficient details could be given and have further selected an

Received: October 29, 2015
Published: December 11, 2015

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

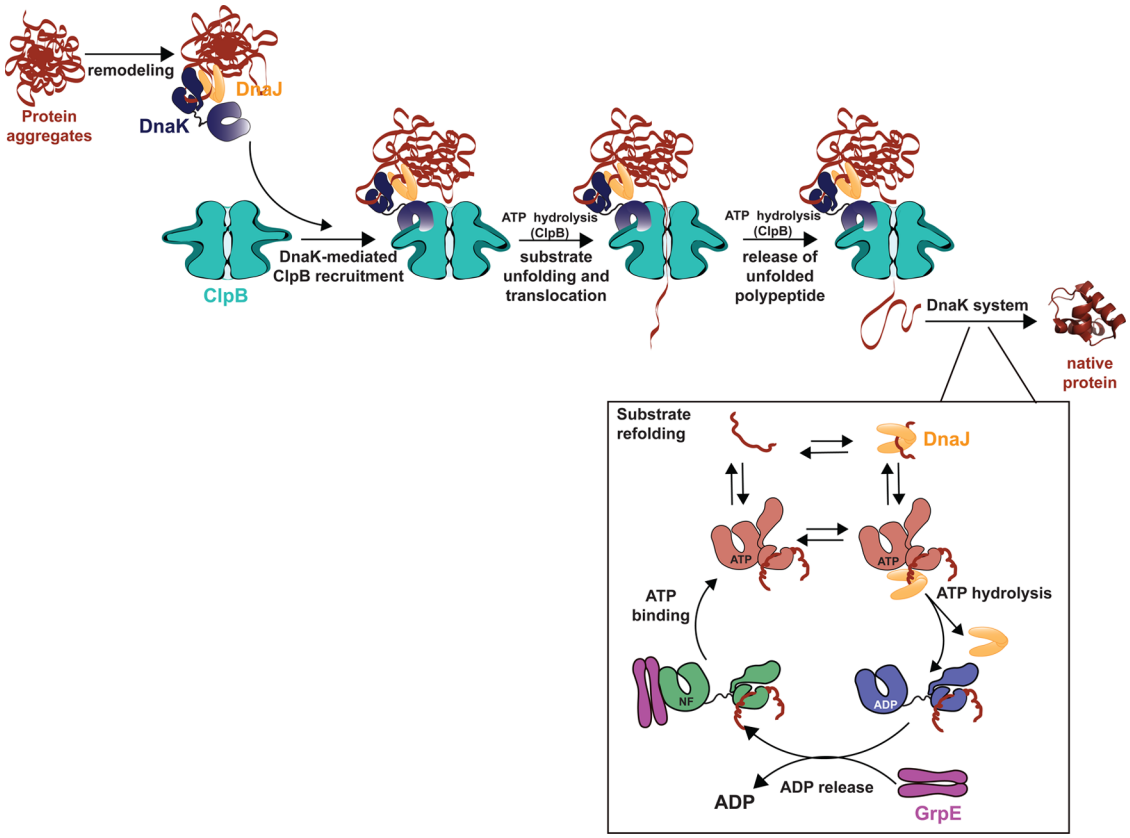

Figure 1. Schematic representation of ClpB/DnaK protein disaggregation and remodeling. Chaperones DnaJ (dimer) and DnaK (comprising both substrate and nucleotide binding domains) are shown in orange and blue, respectively, the protein aggregate in red, and ClpB in turquoise. The inset (black box) depicts the substrate refolding cycle "catalyzed" by the DnaK system that includes the dimeric nucleotide exchange factor GrpE (purple). Details are given in the text.

example from our own research that is well understood to us. The advances in NMR technology which facilitated the described studies will be highlighted, as will the limitations and challenges that remain to be overcome.

2. THE PROBLEM AT HAND

Cellular homeostasis requires the complex interplay between a variety of different factors.[2,4,5] Integral to the process are molecular chaperones that are involved in protein quality control, in maintaining proper protein folding, in protein translocation, in degradation of unfolded proteins, and in preventing and reversing protein aggregation.[51] One such chaperone is the heat shock protein ClpB in *Escherichia coli* (or Hsp104 in eukaryotes) that plays a central role in the remodeling of protein aggregates and, thus, has a fundamental impact on cell physiology, aging, and disease.[52,53]

ClpB/Hsp104 functions in cooperation with a second set of molecular chaperones, the DnaK/Hsp70 system, to unravel toxic, stress-induced protein aggregates.[52,53] A schematic of the ClpB/DnaK remodeling pathway is shown in Figure 1. Initial binding of the small chaperone DnaJ (40 kDa, orange), followed by DnaK (70 kDa, blue) chaperone recruitment, remodels and potentially exposes fragments of the aggregate.[53] DnaK is a two-domain protein comprising substrate (SBD) and nucleotide (NBD) binding domains connected via a flexible linker,[54] with the SBD specifically binding to short, exposed hydrophobic stretches of the aggregate.[55,56] After engagement of the aggregate by DnaJ/DnaK, the ClpB (580 kDa) molecular machine is brought to the site of aggregation via a direct interaction with DnaK.[37,57] Using the energy provided by ATP hydrolysis, ClpB threads each chain of the aggregate through its central pore, thus leading to sequential disentanglement of the aggregated protein.[53,58] The individual unfolded polypeptide chains that are produced in this manner can refold on their own to form functional proteins, or may require one or more passages through the DnaK cycle[56] that is shown in the box in Figure 1. As in the disaggregation pathway, substrates enter the DnaK cycle via interaction with DnaJ; alternatively, they can directly bind to ATP-DnaK. In the DnaK-ATP-bound conformation, the substrate has fast on/off rates and a relatively low affinity for DnaK. Subsequent ATP hydrolysis leads to a large conformational change of DnaK[59,60] that locks the substrate in the bound state,[61] where it is sequestered with high affinity until a nucleotide exchange factor (GrpE in *E. coli*) promotes ADP release,[54] which then allows rebinding of ATP and release of substrate.[56]

The pathways of Figure 1, although informative in identifying key molecular players and in providing a global picture of the disaggregation process, are necessarily qualitative. A long-standing goal has been to move beyond cartoons and quantify each of the steps through the development of atomic-resolution models of the molecular interactions that are critical in each stage of the scheme of Figure 1. Such a complete description does not exist presently, but unique insights are provided through studies that exploit many different structural techniques, including cryo-EM, X-ray crystallography, and solution NMR.[52,53] In subsequent sections the contributions from solution NMR to several of the steps illustrated in Figure 1 are described, emphasizing the unique attributes of the methodology.

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

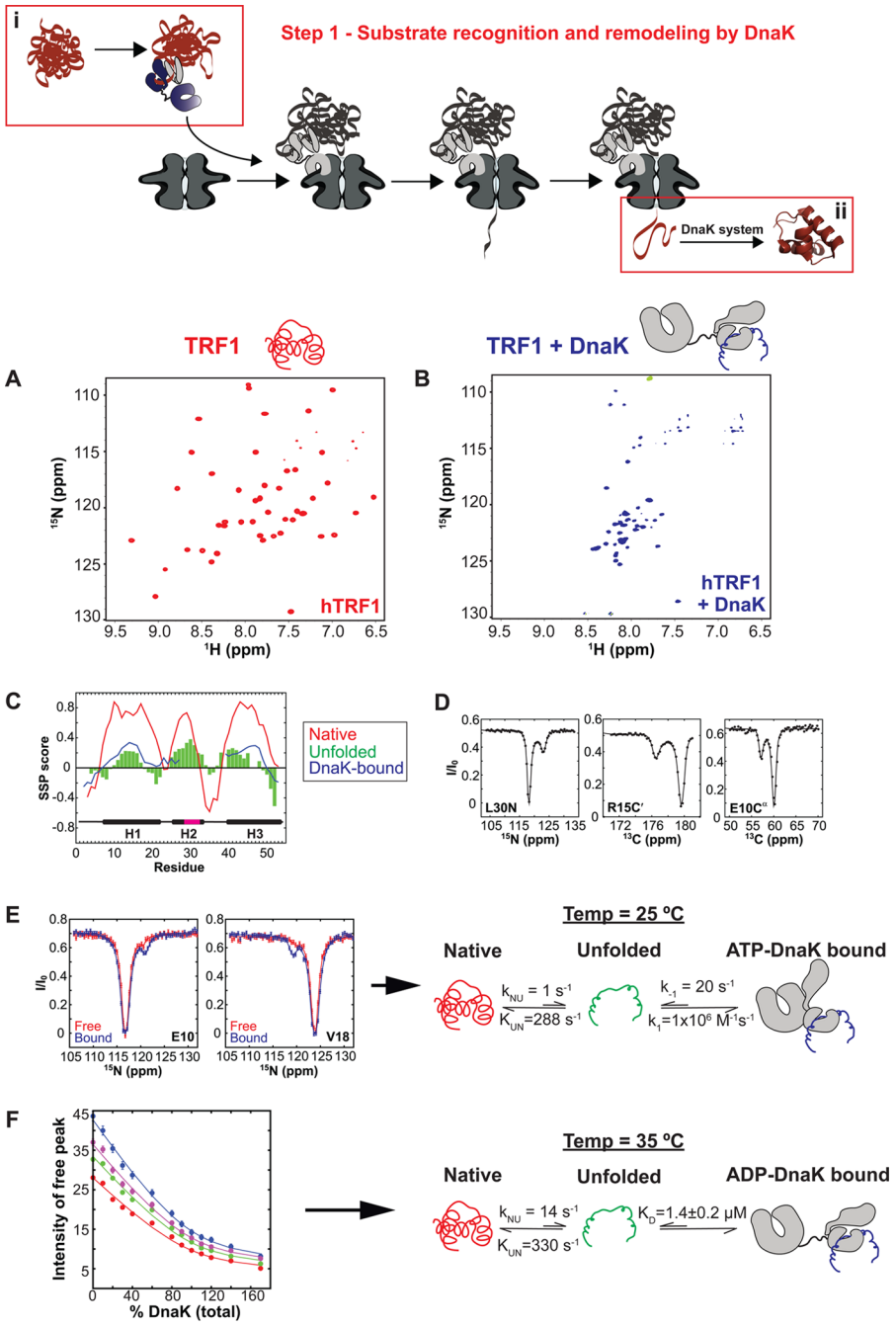

Figure 2. Substrate recognition by the DnaK chaperone. (Top) ClpB/DnaK disaggregation pathway of Figure 1, highlighting in red the steps involved in substrate recognition by DnaK. $^{1}H-^{15}N$ TROSY-HSQC spectra of ^{15}N-labeled hTRF1 (A) unbound and (B) bound to unlabeled ADP-DnaK. Interaction with DnaK results in the global unfolding of the substrate protein. (C) Secondary structure propensity (SSP) scores for the native (red line), unfolded (green bars), and ADP-DnaK-bound (blue line) states of hTRF1. The secondary structure of native hTRF1 is shown on the bottom, with the predicted DnaK binding site in magenta. (D) CEST profiles probing the folding/unfolding transition of hTRF1. Positions of the large and small dips directly correspond to the chemical shifts of nuclei (L30, left; R15, middle; E10, right) in the folded and unfolded states, respectively. (E) Interaction of TRF1 with ATP-bound DnaK. CEST profiles of free (red) and ATP-DnaK-bound hTRF1 (blue) recorded at 25 °C to minimize the contributions from the folding/unfolding equilibrium. Solid lines are fits of the CEST profiles to a three-state model (shown on the right), where substrate first unfolds and then binds to DnaK. The kinetic constants derived from the fits are indicated for the relevant steps in the reaction. (F) Interaction of TRF1 with ADP-DnaK. K_D values were obtained by quantifying the decrease in peak intensities of unbound TRF1 as a function of total DnaK concentration (100% corresponds to 1:1 DnaK:hTRF1). Titration isotherms were fitted to a three-state binding model, shown on the right, with the obtained kinetic and thermodynamic parameters indicated. Figure adapted with permission from Sekhar et al.[62]

3. STEP 1

Substrate Recognition by DnaK. Figure 2 (top) shows a simplified version of the cycles illustrated in Figure 1, highlighting the DnaK–substrate recognition events (red boxes) that are the focus of this section. As described above, DnaK binds to regions of aggregates that are exposed (box i) or to misfolded or unfolded polypeptide chains (box ii).[56] The conformational changes to DnaK that accompany the binding event are well understood and have been characterized in atomic detail by a number of elegant crystallographic studies.[59,60]

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466–1477

Substrate initially binds to an ATP form of DnaK where both substrate and nucleotide binding domains are docked on each other, with the lid that flanks the substrate-binding site in an open state.[59,60] ATP hydrolysis locks the substrate in the binding pocket by repositioning the lid into a closed conformation, with the two DnaK domains becoming disengaged[59,60] (Figure 1, box). While it is clear that there are large structural changes in DnaK that accompany ATP hydrolysis, the conformational properties of the substrate are much less well established, since the only high-resolution structural information is available for peptides bound to DnaK. What conformations are adopted by folding competent cognate substrates, and how do their structures change as a function of the large conformational changes to DnaK that accompany the reaction cycle?

Solution NMR is a powerful technique to address these questions, not only because atomic-resolution information can be obtained, but also because new spin-relaxation NMR methods can be used to probe the dynamic substrate in the bound form that would be difficult to study by other high-resolution methods.[63] Currently it is not feasible to perform high-resolution, quantitative NMR studies of DnaK−aggregate interactions (box i of the pathway), since the aggregates are inhomogeneous in size and typically many MDa in molecular mass. Instead, DnaK−client interactions have been probed using a small (53 residue) three-helix bundle DNA binding domain of the human telomeric repeat-binding factor 1 (hTRF1) protein.[64] The hTRF1 domain studied (heretofore referred to as hTRF1) is well folded in aqueous solution, as can be appreciated by the well-resolved $^{1}H-^{15}N$ HSQC spectrum in Figure 2A. In contrast, once in complex with ADP-DnaK, hTRF1 is globally unfolded, which is made clear from the HSQC spectrum of the complex showing poor chemical shift dispersion in the ^{1}H dimension.[62] Residue-specific secondary structure propensity (SSP) scores[65] were calculated from the assigned ^{1}H, ^{15}N, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and ^{13}CO chemical shifts of hTRF1 in the DnaK-bound state, establishing that regions corresponding to helices 1 and 3 of the folded protein form up to 40% residual helical structure in the bound conformation, Figure 2C (blue).[62] Chemical shift assignments could not be obtained for many of the residues in helix 2 since they make up the DnaK binding site and cross peaks from them become broadened upon formation of the ∼75 kDa DnaK−hTRF1 complex. It is noteworthy that highly deuterated samples of hTRF1 and of DnaK were used in the analysis to minimize relaxation losses in $^{1}H-^{15}N$ experiments that were recorded using TROSY methods,[66] yet significant issues with sensitivity were nevertheless encountered. This points to a limitation using backbone amides as probes that can often be eliminated when methyl groups are used, as will be discussed in subsequent sections.

It is of interest to establish whether the residual structure observed in the bound form is a direct result of binding to DnaK, so that chaperone binding actively promotes secondary structure formation, or whether this level of helicity simply reflects the structural propensities of the hTRF1 primary sequence. To address this question, the chemical shifts of unfolded, unbound hTRF1 must be obtained. This is difficult to accomplish without additives that promote unfolding, and potentially also modify secondary structure tendencies, since the population of the unfolded state in the hTRF1-free form is not sufficiently high to directly observe it in even the most sensitive of NMR experiments.

The development of spin relaxation experiments for "seeing the invisible", i.e., sparsely populated and transiently formed states that do not give rise to observable peaks in NMR spectra, provides an avenue for obtaining the unfolded chemical shifts in this particular case.[63,67] Consider an exchange reaction where the highly populated ground state (G) interconverts with a rare conformationally excited state (E), $G \underset{k_{EG}}{\overset{k_{GE}}{\rightleftharpoons}} E$. One approach to obtaining the invisible E-state spectrum is via chemical exchange saturation transfer (CEST), a method developed over 50 years ago[68] and subsequently used in MRI imaging.[69,70] In the case of protein applications, CEST can be of great utility so long as the fractional population of E, p_E, is on the order of 1% or higher and $50\ s^{-1} < k_{ex} < 400\ s^{-1}$, where $k_{ex} = k_{EG} + k_{GE}$.[63,71] In this approach, a series of spectra are recorded where the position of a very weak magnetic field is applied over a range of chemical shifts, in essence "searching" for excited-state resonances. When the weak field is proximal to the resonance frequency of a peak from E, it perturbs the E-state magnetization, and this perturbation is transferred via the chemical exchange process to the corresponding ground-state peak (from the same residue), leading to an attenuation of the observable ground-state correlation. In a similar manner, placing the weak field at the resonance position of the ground-state correlation significantly attenuates it. The result is a series of CEST profiles (typically one for each residue), consisting of major and minor dips at the resonance positions of exchanging nuclei of the G and E states. As shown in Figure 2D, ^{15}N, ^{13}CO, and $^{13}C^{\alpha}$ chemical shifts of E can be obtained by using weak fields of the appropriate frequencies, with the shifts given by the positions of the minor dips. Additionally, values of p_E and k_{ex} can be extracted by fitting the profiles to the appropriate model of chemical exchange using the Bloch−McConnell equations.[72]

The chemical shifts of the invisible unfolded state of hTRF1, obtained in this manner, were used to calculate residue-specific SSP values (Figure 2C, green), showing similar secondary structural propensities for the unfolded free and DnaK-ADP-bound states of hTRF1 (compare green and blue).[62] Thus, binding of hTRF1 to DnaK-ADP does not alter the structural propensities intrinsic to the primary sequence of the protein outside of the binding region; rather, DnaK "allows" the substrate to explore its intrinsic energy landscape, adopting secondary structure in a protective environment.[62]

During the DnaK cycle (Figure 1 box), the chaperone converts among ATP, ADP, and nucleotide-free (NF) states. Large conformational changes to DnaK are observed between ATP- and ADP-bound conformations,[59,60] with very similar structures for the NF and ADP forms.[73] In order to establish whether the large structural transitions in DnaK that accompany the reaction cycle lead to structural changes in hTRF1, a $^{1}H-^{15}N$ correlation map of hTRF1 was first recorded in the NF bound state, and the spectrum was superimposable with that obtained for the ADP-bound form of the protein.[62] Unfortunately, it is not possible to record the corresponding data set for hTRF1 bound to DnaK-ATP because the affinity for substrate is considerably lower than in the ADP form of the chaperone,[61] and it is difficult to saturate binding. Addition of increasing amounts of DnaK-ATP resulted in samples of high viscosity and very poor quality spectra. This is not an infrequently encountered situation in NMR analyses of molecular interactions, and it often prevents a detailed analysis of the system of interest. CEST can again potentially be extremely useful in these cases because it is possible to detect small populations of molecules (in this case hTRF1 bound to DnaK-ATP) via an abundant ground state that gives high-quality spectra (free hTRF1). Figure 2E compares a pair of CEST profiles recorded on samples of free hTRF1 and hTRF1 with an

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

approximately equimolar amount of DnaK-ATP.[62] Under these conditions, the fraction of DnaK-bound substrate is approximately 10%. The small dips that are observed upon addition of DnaK-ATP derive from bound hTRF1, enabling chemical shifts of this elusive conformer to be obtained.[62] It is noteworthy that experimental conditions have been chosen to minimize the contributions from the folding/unfolding equilibrium of free hTRF1 by lowering the temperature from 35 °C (p_U = 4.2%) to 25 °C (p_U = 0.3%) and therefore stabilizing the folded state. CEST profiles have been fit to a three-state model, where substrate first unfolds and then binds to DnaK, as shown in Figure 2E, enabling the extraction of kinetic and thermodynamic parameters of the binding reaction as well as chemical shifts of the bound state. Notably, very good agreement is observed between substrate chemical shifts in ADP-, ATP-, and NF-DnaK-bound states, indicating that there is no change in the bound hTRF1 structure as DnaK cycles through different conformations.[62]

Thermodynamic parameters can be obtained for the binding of hTRF1 to DnaK-ADP as well. A higher affinity complex is formed in this case relative to DnaK-ATP so that it is possible to achieve much higher amounts of the hTRF1−chaperone complex. Thus, K_D values can be measured simply by quantifying decreases in intensities of unbound hTRF1 peaks (Figure 2F) and, concomitantly, increases in intensities of peaks derived from the DnaK-bound state as a function of the addition of DnaK. Fits of peak intensities to the three-state binding model described above using k_{NU} and k_{UN} values from analysis of CEST profiles give a K_D value of 1 μM, which is approximately a factor of 20 lower than that for binding to DnaK-ATP.[62]

4. STEP 2

DnaK−ClpB Complex Formation. In the previous section we have described how solution NMR methods were used to study the complex between a folding competent cognate substrate and DnaK, focusing on the substrate. Spin relaxation methods that probe rare conformational states were essential for the characterization of the bound hTRF1 substrate. Similar to individual substrate molecules, exposed regions of polypeptides that are part of aggregates bind to the SBD of DnaK prior to engagement with ClpB (Figure 3, top). An important role of DnaK is to present the aggregate to ClpB[55,56] so as to facilitate the pulling reaction that disengages individual protein chains.[53,58] Critical to this process is a physical interaction between DnaK and ClpB,[37,57] the focus in what follows.

ClpB (*E. coli*) or Hsp104 (yeast) is a ∼600 kDa hexameric ATPase, with individual protomers comprised of several different domains,[74] as indicated in Figure 3A. These include an amino-terminal domain (NTD), a pair of nucleotide binding domains, NBD1 and NBD2, and a coiled-coil domain (CCD) extension in NBD1 that is unique to the ClpB/Hsp104 chaperones. Although biochemical studies had established the importance of the CCD for the disaggregation reaction and that the CCD and DnaK interact,[75,76] the molecular details remained to be elucidated.

Solution NMR can be a powerful tool to probe such molecular interactions, although traditionally there have been rather strict size limitations. These result from increased relaxation rates and hence faster signal decay that is associated with large complexes,[13,77,78] such as ClpB−DnaK.[37] In this regard methyl groups have become popular probes of molecular structure and dynamics in these high-molecular-mass systems.[7,79] The utility of methyl groups in this context reflects, in part, the fact that the three-fold methyl rotation about its symmetry axis is an efficient mechanism for averaging dipolar interactions and hence narrowing lines[80,81] and, additionally, that methyl groups are positioned at the end of often long side chains whose dynamics again lead to averaging.[82,83] Equally important is that the inherent spin physics is such that it is possible to create methyl coherences where the dipolar fields constructively interfere, via a dipolar TROSY effect, which significantly attenuates relaxation rates and hence improves both spectral resolution and sensitivity.[14] It can be shown that, in a ^{1}H−^{13}C HMQC pulse scheme, and for a methyl group attached to a protein tumbling in the macromolecular limit, half of the net magnetization relaxes slowly due to cancellation of local fields, with the second half decaying so rapidly that it does not contribute to the observed signal.[14] The use of deuteration is important, as for backbone amide TROSY studies,[66] since the high dilution of proton spins effectively isolates the slowly relaxing coherences from those relaxing more efficiently, ensuring optimal spectral quality. From a practical perspective, the production of highly deuterated proteins with $^{13}CH_3$-labeled methyl groups at any methyl-containing amino acid is now possible using commercially available precursors and protein expression protocols that are well established.[84]

As a first step toward a molecular characterization of the ClpB−DnaK interaction, a fragment of ClpB was constructed, ClpB$^{\Delta NBD2}$, that does not contain the NBD2 domain. Since NBD2 is necessary for ClpB hexamerization,[85] ClpB$^{\Delta NBD2}$ is monomeric and relatively small (65 kDa), enabling very high quality methyl-TROSY spectra to be rapidly obtained from highly deuterated, Ile (δ1), Leu, Val methyl-labeled protein (ILV-ClpB$^{\Delta NBD2}$). As in other studies of protein complexes involving such a labeling scheme[44] only one of the two isopropyl methyl groups of Leu/Val is $^{13}CH_3$, while the other is $^{12}CD_3$.[28] Upon addition of highly deuterated DnaK, a number of peaks in ^{1}H−^{13}C HMQC spectra of ILV-ClpB$^{\Delta NBD2}$, derived from residues localized to the CCD, titrated significantly (Figure 3B), and the titration profiles could be fit to obtain a K_D of approximately 30 μM for the interaction,[37] Figure 3C. In order to establish which residues of DnaK are involved in forming the complex, highly deuterated hexameric ClpB was titrated into a solution of ILV-DnaK, and the changes to positions of cross peaks in the resulting ^{1}H−^{13}C HMQC spectra were measured. Despite the high molecular mass of the complex (650 kDa), the titration data could be accurately quantified, Figure 3D, with fitted K_D values very similar to those from the ILV−ClpB$^{\Delta NBD2}$−DnaK titration, suggesting a 1:1 ClpB hexamer−DnaK stoichiometry.[37] Importantly, all of the chemical shift changes to DnaK localized to the NBD and in particular to domains IB and IIB, Figure 3E.

In order to obtain a model of the ClpB−DnaK complex, a series of paramagnetic relaxation enhancement (PRE) experiments[86,87] was conducted whereby nitroxide spin labels were positioned at different sites in ClpB$^{\Delta NBD2}$ (one at a time) and ^{1}H−^{13}C HMQC spectra of the ClpB$^{\Delta NBD2}$−ILV−DnaK complex recorded and compared to spectra in the absence of the nitroxide. The presence of a spin label leads to a decrease in intensities of cross peaks in NMR spectra in a manner that depends on r^{-6}, where r is the distance between the unpaired electron of the nitroxide and the methyl ^{1}H spin.[86−88] A series of distances between spin labels and ILV methyl probes in DnaK were obtained from which a structural model of the ClpB$^{\Delta NBD2}$−DnaK complex was calculated using the molecular docking program HADDOCK.[89,90] Figure 3F highlights important molecular contacts between DnaK (blue) and ClpB$^{\Delta NBD2}$.[37] Notably Y494,

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

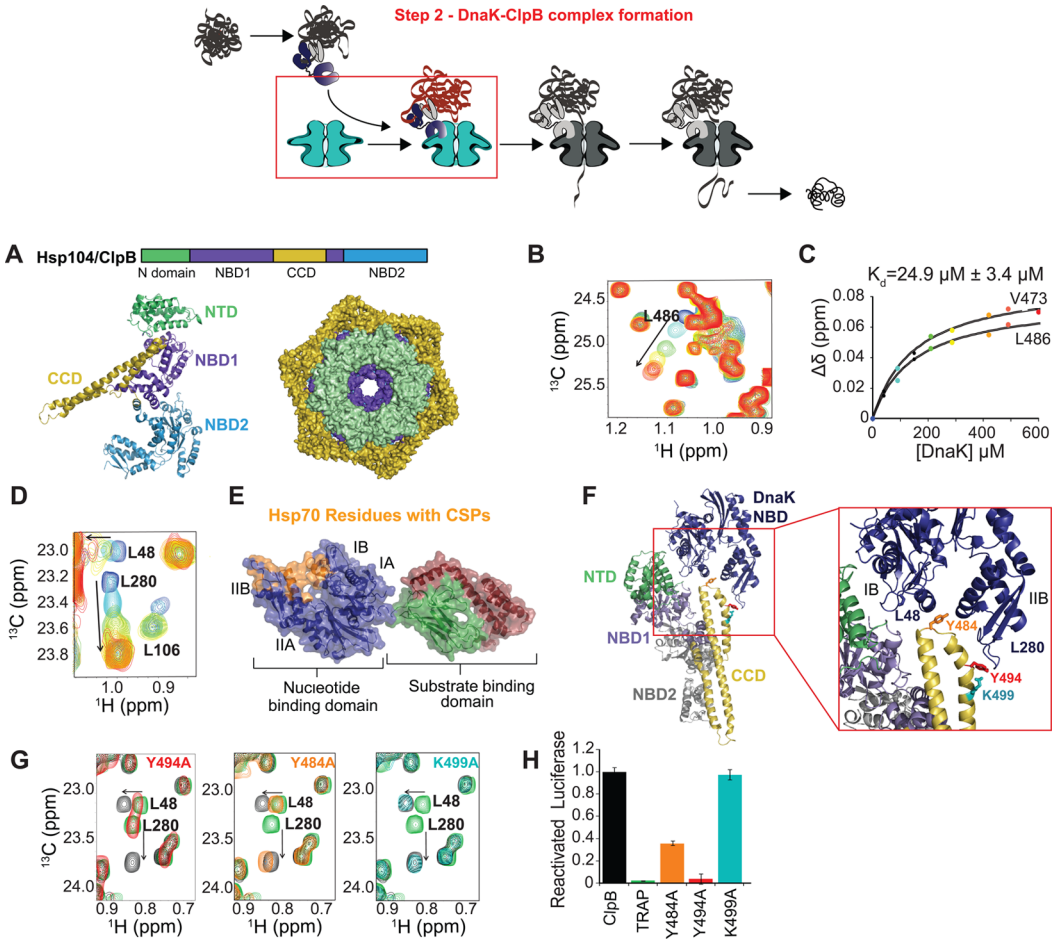

Figure 3. DnaK–ClpB complex formation. (Top) Schematic of Figure 1, highlighting the DnaK–ClpB recognition event in red. (A) Structure and domain organization of the hexameric ClpB chaperone. Domain organization (top) and protomeric structure (bottom left) of the ClpB chaperone (PDB 1QVR[74]). The ClpB protomer consists of an N-terminal domain (NTD, green), two nucleotide binding domains (NBD1 and NBD2, purple and light blue, respectively), and a coil–coil domain insertion (CCD, yellow). The monomers assemble into a hexamer (bottom, right) consisting of three rings formed by NTDs (top ring, green), NBD1-CCD (purple-yellow), and NBD2 enclosing the central pore. (B) NMR titration of ^{2}H, ILV-ClpB$^{\Delta NBD2}$ with deuterated (NMR invisible) DnaK. A series of ^{13}C,^{1}H HMQC data sets was recorded in which a fixed concentration of ClpB$^{\Delta NBD2}$ was titrated with increasing concentrations of DnaK. Selected regions of the resulting ^{13}C–^{1}H correlation spectra are displayed (blue, no DnaK; cyan to red, addition of DnaK in increasing DnaK/ClpB$^{\Delta NBD2}$ ratios from 0.2 to 2.0). DnaK concentration-dependent changes in the position of a methyl resonance of ClpB L486 are indicated by an arrow. (C) Chemical shift titration profiles can be used to calculate dissociation constants (color coding for ClpB V473, L486 as in panel B). (D) NMR titration of ^{2}H, ILV-DnaK (70 kDa) with ^{2}H, hexameric ClpB (580 kDa). ClpB-dependent chemical shift changes for methyl groups of DnaK L48 and L280 are indicated by arrows. (E) DnaK structure (PDB ID: 2KHO[73]), with NBD shown in blue, SBD in green, and the lid section of the SBD in red. Residues with concentration dependent chemical shift changes as a result of ClpB binding are highlighted in orange. (F) HADDOCK-derived model of the DnaK–ClpB complex. DnaK NBD (blue) subdomains IB and IIB interact with the CCD of ClpB (yellow) but not with ClpB NTD (green) or NBD1 (purple) domains. The location of NBD2, truncated in the ClpB$^{\Delta NBD2}$ protein, is shown (gray) for clarity. The inset shows a close-up view of the DnaK–ClpB interface, highlighting Y484 (orange), Y494 (red), and K499 (cyan) residues of ClpB. (G) Key ClpB residues (shown in panel F) were mutated to Ala (Y494A, red; Y484A, orange; and K499A, cyan) and complex formation with DnaK was monitored by recording HMQC experiments. Spectra of ILV-DnaK either free (green) or fully bound (black) to WT ClpB$^{\Delta NBD2}$ are shown in each panel as a reference. DnaK peaks for residues L48 and L280 serve as reporters for the binding of subdomains IB and IIB of DnaK, respectively. (H) Effects of the ClpB mutations in G on protein disaggregation were determined by monitoring the reactivation of luciferase aggregates by hexameric ClpB (WT, black; TRAP-inactive ClpB variant, green; Y494A, red; Y484A, orange; and K499A, cyan) in the presence of the DnaK/DnaJ/GrpE chaperone system. Figure adapted with permission from Rosenzweig et al.[37]

localized to the CCD, and shown previously by mutagenesis to be important for ClpB disaggregation activity, was located at the binding interface. In order to validate the structure a series of ClpB$^{\Delta NBD2}$ point mutants was made where Y484, Y494, and K499 were mutated to Ala. Each mutated protein was added in excess to ILV-DnaK and ^{1}H–^{13}C HMQC spectra recorded. For comparison, spectra of ILV-DnaK either free (green) or fully bound (black) to WT ClpB$^{\Delta NBD2}$ are shown as well, Figure 3G. Focusing on cross peaks for L48 and L280 that, among others, serve as reporters for the binding of subdomains I and II of DnaK, respectively, it is clear that the Y494A mutation abolishes binding to DnaK, while Y484A eliminates binding to subdomain I, although interactions with subdomain II persist.[37] In contrast to these two tyrosine mutations, K499 is removed from the binding interface, and not surprisingly, K499A has no effect on the ClpB–DnaK interaction.[37]

As a further validation of the structure, biochemical experiments were performed where the ability of each of the mutants, in the context of hexameric ClpB, to refold fully aggregated firefly luciferase was assayed in a reaction that requires a tight

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466–1477

collaboration between the ClpB and DnaK chaperone systems.[91,92] Both WT and K499A ClpB, whose binding to DnaK is indistinguishable from WT, were able to fully reactivate luciferase, while Y494A ClpB that showed no binding to DnaK in NMR spectra was completely defective. Y484A ClpB that binds to subdomain II but not to I was able to restore approximately 40% of the luciferase activity.[37]

5. STEP 3

Substrate Recognition by ClpB. In section 4, the physical interaction between DnaK and ClpB was studied by methyl-TROSY NMR. In part, a divide-and-conquer approach[30,44] was used where individual pieces of ClpB were produced that contained the important DnaK binding region localized to the CCD, for example ClpB$^{\Delta NBD2}$,[37] so that the resulting complex (~130 kDa) was sufficiently small that very high quality methyl spectra could be obtained rapidly. Such a strategy is used frequently in studies of many different molecular interactions. In addition, the full ClpB−DnaK complex was investigated, focusing on ILV-DnaK. Although the resulting molecular mass is 650 kDa, only DnaK is labeled and hence NMR visible, so that the inherent spectral complexity is that of a 70 kDa protein. The results unequivocally establish a binding interaction between the NBD of DnaK and the CCD of ClpB[37] that brings the aggregate attached to the DnaK SBD proximal to ClpB for subsequent transfer to the disaggregase, Figure 4 (top, red box). Characterizing the interactions between substrate and ClpB that form the initial stage of the transfer is the focus of what follows.

As shown in Figure 3A, each of the six ClpB protomers is comprised of a pair of nucleotide binding domains that bind and hydrolyze ATP.[74] The energy derived from ATP hydrolysis is then used for threading individual protein chains of the aggregate through a central pore of the hexameric ClpB chaperone.[58] Each of the six NBD1 domains in the complex contributes a conserved tyrosine (Y243 in ClpB), localized to a so-called tyrosine loop, that together form a tyrosine ring that interacts with positively charged and aromatic residues of the threaded substrate.[58,95] Mutating the NBD1 pore tyrosine residues decreases ClpB threading efficiency; however, when combined with deletion of the six NTDs, a complete loss in activity is observed.[96] This suggests that both NTDs and tyrosine loops play an important role in initial substrate engagement.

As a first step toward characterizing NTD−substrate interactions, a series of different potential protein substrates was screened that included aggregates, folded polypeptides, and unfolded proteins. ^{1}H−^{15}N HSQC spectra were recorded of ^{15}N-labeled NTD (16 kDa) in the presence of each target, and binding was assayed quickly by noting changes in the position of cross peaks. Neither aggregates nor folded proteins bound to the NTD, although changes were seen for the intrinsically disordered protein α-casein that is a substrate for ClpB.[97] A series of polypeptides were then selected with different lengths, charge distributions, amino acid compositions, and amounts of residual structure, and a set of HSQC screens were performed. All of the proteins/peptides that showed changes to the spectrum of the NTD bound with affinities in the 100−400 μM range, and the NTD residues that were found to interact with substrates by NMR were localized to a single, highly hydrophobic groove in the domain,[94] Figure 4A. In order to understand how ClpB "chooses" its client proteins, each substrate was ^{15}N labeled, and ^{1}H−^{15}N HSQC spectra were recorded in the presence of unlabeled NTD. Residues whose cross peaks showed significant changes upon addition of NTD were used to generate the distribution of "preferred" client amino acids that is shown in Figure 4B. Notably, hydrophobic, branched residues have the highest preference, and this distribution[94] is quite different from that observed for the tyrosine pore loops, where charged and aromatic amino acids predominate.

The binding of substrate to the NTDs of ClpB is an important step in the disaggregation of some client proteins, perhaps positioning the substrate properly with respect to the pore so that the pulling reaction can occur efficiently. Of interest is how NTD binding affects the stability of the substrate and whether it might "prime" substrate for the subsequent steps in the process. NMR is a powerful tool for studying protein stability,[98] and the effect of NTD binding on substrate was investigated by thermodynamic studies of the Pin1WW domain,[99] which contains a potential NTD binding site. Temperature titrations of free Pin1WW and Pin1WW with excess NTD were performed, Figure 4C, and nearly linear changes in chemical shifts were observed. For the Pin1WW domain that is a rapid folder, the observed peak positions are population-weighted averages of chemical shifts in the folded and unfolded states of the protein and can therefore be used to obtain melting curves as shown in Figure 4D. Addition of NTD significantly decreases the stability of the Pin1WW domain, lowering the melting temperature by 10 °C.[94] These results are consistent with the notion that interaction with NTD primes the substrate for subsequent translocation.

The above studies, exploiting the divide-and-conquer approach and focusing only on the NTD domain, are powerful since they provide important insight into the interaction of substrate with NTD. However, they are limited in this case because the initial substrate engagement involves both the NTD and the tyrosine loops of NBD1 and can thus only be fully studied using the intact ClpB hexamer. Figure 4E shows a selected region of the ^{1}H−^{13}C HMQC spectrum of ILV-ClpB, focusing on cross peaks from Leu and Val.[94] The spectrum illustrates a major difficulty in working with high-molecular-weight proteins where the individual protomer sizes are large (in this case close to 100 kDa). Each ClpB monomer is comprised of 112 Leu and 59 Val residues, and the overlap is prohibitive. Although an improvement is expected with increasing static magnetic field (this spectrum was recorded at 800 MHz and in a 2D ^{13}C,^{1}H-based experiment resolution will scale as $(\nu_o/800)^2$ where ν_o is the ^{1}H frequency at the higher field), it is unlikely to be sufficient. Indeed, on the basis of our studies of a number of large complexes, it may well be that the major issue in NMR studies of molecular machines is one of spectral resolution, as for ClpB, rather than molecular tumbling time, as high-quality methyl spectra were obtained for methyl reporters localized to rigid regions of the 20S core particle proteasome in complex with the 11S regulator that had an aggregate molecular weight of over 1 MDa.[36] The use of segmental labeling strategies offers a potential solution to the resolution problem because individual domains can be labeled as desired, while the remainder of the protein remains "NMR silent".[100−103] Figure 4F shows a much more tractable data set recorded of intact ClpB labeled with ILV only in the NTD and in the linker connecting it with NBD1 that was generated using an expressed protein ligation strategy. Nearly all of the correlations are resolved, and assignment of all cross peaks could be obtained, largely from those of the isolated domain.[94]

Cryo-EM studies of ClpB have suggested that the NTD is at least partially flexible.[104] In order to quantify the extent of flexibility and to establish whether NTD dynamics change with nucleotide state, a set of methyl-based spin relaxation

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

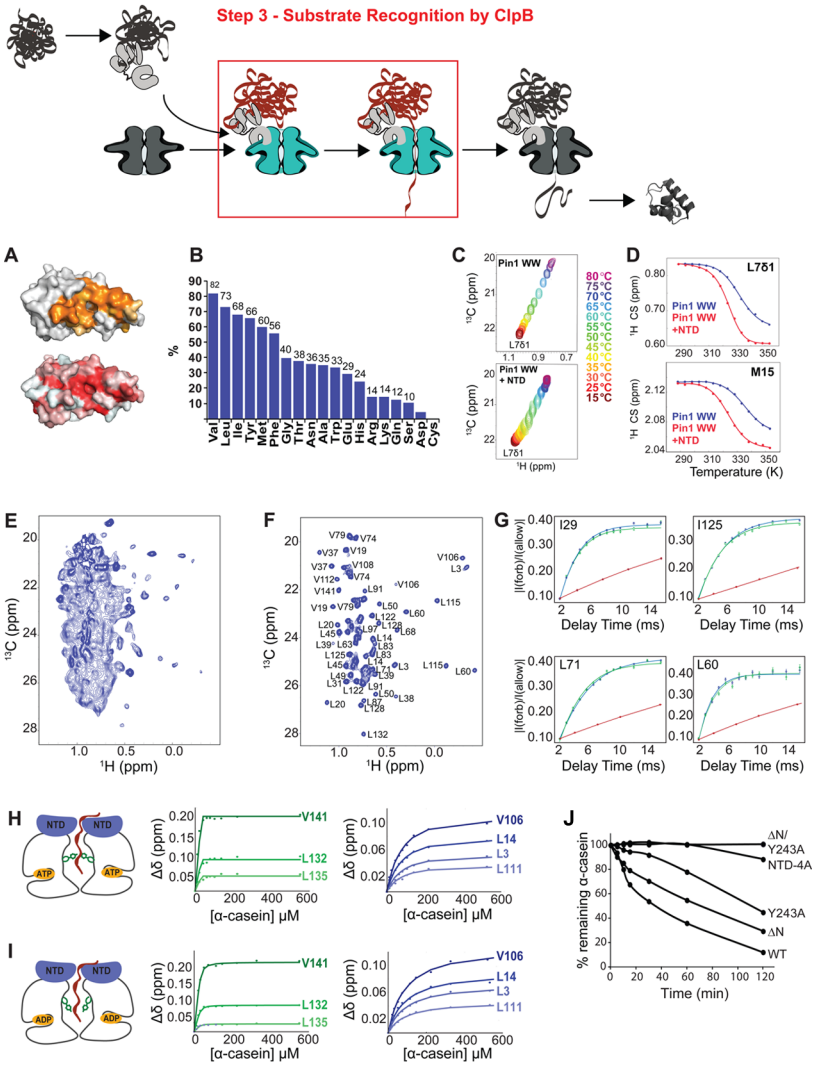

Figure 4. Substrate recognition by the ClpB disaggregating chaperone. (Top) Schematic of Figure 1, highlighting the substrate recognition steps in red. (A) ClpB NTD surface representation (PDB ID: 1QVR[74]) with NTD residues participating in the interaction with unstructured substrates, as established by NMR titration experiments, colored in orange (top). Bottom panel shows ClpB NTD structure colored by residue hydrophobicity (white to red gradient) indicating that the N-terminal domain substrate binding site is enriched in hydrophobic residues. (B) Normalized fraction of amino acids from ClpB-client proteins whose amides show chemical shift changes in the presence of the NTD. Values, normalized against the total number of each amino acid in all substrate sequences examined, establish that the NTD preferentially binds to hydrophobic residues in its client proteins. (C) Selected regions from $^{1}H-^{13}C$ HMQC spectra, highlighting the temperature dependence of chemical shifts for the L7δ1 methyl group of Pin1WW in the absence (top) and presence of ClpB NTD (bottom). (D) Representative melting curves for L7δ1 and M15ε of free (blue) and NTD-bound (red) Pin1WW fitted to standard equations for two-state unfolding.[93] The T_M value of the substrate is reduced by ~10 °C upon NTD binding. (E,F) Selected regions of $^{1}H-^{13}C$ HMQC methyl-TROSY spectra of the 580 kDa hexameric ClpB chaperone uniformly ILV-labeled (E) or with segmental ILV-labeling confined to the N-terminal domain (F). (G) Build-up curves of intensity ratios ($|I_{forb}/I_{allow}|$) from methyl ^{1}H triple-quantum experiments plotted against relaxation delay for selected NTD residues from the hexameric ClpB in ATP- (green) and ADP-bound (blue) states (55 °C), or from the isolated NTD, 25 °C (red). (H,I) Titration of α-casein into samples of (H) ATP- and (I) ADP-bound hexameric ClpB, with segmentally ILV-labeled NTDs. The cartoon representation of ClpB (left) highlights the differences in the positions of the tyrosine pores between the ATP (H) and ADP (I) states. Fits of chemical shift changes of ClpB residues L132, L135, and V141 (green curves) as a function of α-casein concentration establish that the affinity of the tyrosine pores for substrate is significantly higher in the ATP state (H) than in the ADP state (I), while the affinity of NTD−substrate interactions (blue curves) does not change upon ATP hydrolysis. (K) Threading activity of ClpB variants monitored through degradation of α-casein. Notably, significantly slower hydrolysis rates were observed for the NTD mutated ClpB variant (NTD-4A) compared to variants of ClpB lacking the N-terminal domain (ΔN) in its entirety. Figure adapted with permission from Rosenzweig et al.[94]

experiments were performed where the overall tumbling time of the NTD was estimated from the buildup of methyl ^{1}H triple-quantum coherence.[105] This coherence can only be generated efficiently for molecules that tumble slowly, and the buildup rate in this case is directly proportional to the overall tumbling time. Figure 4G shows buildup profiles for a number of NTD methyl groups of ClpB in both ATP (green) and ADP (blue) states. For comparison, similar profiles are also shown for the isolated NTD (red). From a global fit of the relaxation data, correlation times of 82 ± 17 and 87 ± 26 ns were obtained for ATP and ADP states, respectively, while a value of 8.1 ± 0.1 ns was measured for the individual domain.[94] Values of 80−90 ns are approximately a factor of 2 less than the expected value for a 580 kDa rigid protein (~180 ns), suggesting that the NTD is indeed

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

partially flexible and only partly docked to NBD1 in both ATP- and ADP-loaded states.

As described above, the binding of α-casein to the isolated NTD was quantified from $^{1}H-^{15}N$ HSQC spectra recorded as a function of added α-casein to a solution of ^{15}N-NTD, and a K_D value of 100 ± 20 μM was obtained.[94] Similar experiments were also conducted with segmentally labeled hexameric ClpB using methyl group probes. A comparison of ClpB spectra obtained with and without α-casein showed chemical shift changes in the hydrophobic substrate binding groove of NTD as well as an additional set of changes that localized to the linker between the NTD and NBD1.[94] The linker residues showing changes in peak positions are proximal to the tyrosine loops, and these shifts were eliminated using a mutant in which the tyrosines of these loops are replaced by alanine, $ClpB^{Y243A}$. Taken together this provides strong evidence that the second set of shifts reports on substrate binding to the tyrosine loops, with the first set sensitive to NTD binding. It is thus possible to "dissect" the binding pathway by focusing on different sets of reporter residues, an advantage that NMR has over other techniques that measure binding and that are sensitive to only the highest affinity step. Figure 4H shows binding isotherms that report on the interaction between α-casein and the tyrosine loops (green) or NTD (blue) of ATP-ClpB, with all of the data fitted simultaneously to the same binding model.[94] Notably, it was not possible to fit the data under the assumption that a single α-casein substrate binds to the hexamer but rather one where three substrates interact (i.e., one casein binds two NTDs). Very significant differences are apparent in casein binding affinities to NTD (K_D = 46 ± 5 μM) and to the tyrosine pores (K_D < 90 nM), although a precise value for the latter is difficult to obtain because of the very high affinity involved. A similar analysis of α-casein binding to ClpB-ADP was performed, with K_D values of 4.0 ± 0.4 μM (Figure 4I, green) and 85 ± 12 μM (blue) measured for binding to the tyrosine loops and NTDs, respectively.[94] The significantly lower affinity for the loops in the ADP case is consistent with previous binding studies[58] and with cryo-EM structural studies that have established that the critical tyrosines of the pore loops (Y243) are disordered in the ADP-bound state and thus likely not in an orientation that is favorable for high-affinity binding.[106]

The relative importance of the NTD and of the tyrosine pore loops for substrate translocation has been quantified through biochemical studies that make use of a variant of ClpB, termed BAP,[58] that is able to bind to the ClpP protease. BAP transfers proteins, including α-casein, into the proteolytic chamber of the associated ClpP for degradation. Thus, the translocation activities of wild-type ClpB and associated mutants can be established by monitoring the efficiency of substrate degradation over time. Figure 4K shows the results of such a study for the wild-type protein and a number of mutants, focusing on α-casein as the substrate. As expected, a double mutant of ClpB where the NTD is removed and the tyrosine pores rendered defective (ΔN/Y243A) shows no translocation, with only fractional amounts of activity for ΔN or Y243A mutants.[94] Interestingly, a mutant of ClpB involving substitutions of four key substrate-binding residues in the NTD to alanine (NTD-4A) shows almost no translocation activity despite the fact that the tyrosine loops are intact.[94] This suggests that the NTD serves to block the channel entrance prior to substrate binding, preventing unfolded proteins from entering, and, further, that engagement of the NTDs must proceed binding to the tyrosine loops, at least for substrates such as α-casein. This regulatory effect may play a role in ensuring that partially unfolded regions of properly folded, functional proteins are not mistakenly identified by the tyrosine loops and unfolded by ClpB.

6. CONCLUDING REMARKS AND FUTURE PERSPECTIVES

In this publication we have highlighted a number of solution-based NMR approaches for studies of high-molecular-weight protein systems, focusing on the ClpB/DnaK chaperones that play critical roles in protein disaggregation. In addition to providing insight into how the ClpB molecular machine carries out its function, the present work assumes a greater importance when one considers that a molecular-level understanding of how polypeptide aggregates are refolded may aid in the rational design of therapeutic strategies to prevent, slow-down, or reverse the progression of a number of protein misfolding and aggregation-based diseases.

It is our opinion that the goal of solution NMR studies of molecular machines should not be one of generating high-resolution static three-dimensional structures, since this is the domain of cryo-EM and X-ray diffraction. Rather, NMR has a unique role in characterizing systems where highly dynamic interactions are at play, for which atomic-resolution studies are typically not possible with other methods. A number of such examples are found in the ClpB disaggregation pathway described here, and for which NMR has provided insight. These include investigations of highly dynamic protein substrates bound to DnaK, of relatively weak affinity ClpB–DnaK interactions that are nevertheless of critical importance for aggregate transfer from DnaK to ClpB, and of weak yet specific NTD–substrate interactions that are important for the initial engagement of the aggregate particle by ClpB. Outstanding questions relating to the "dynamics" of protein disaggregation include how the handoff of substrate from DnaK to ClpB precisely occurs, how the NTDs of ClpB reorient upon substrate binding so as not to block substrate entry into the central pore of the disaggregase, and the role of the unique CCDs that change position in response to different nucleotide states of ClpB. Molecular machines have many moving parts that can be critical for function, and a range of NMR experiments have been designed for studies of such motions over a wide spectrum of time scales.[45,48,107] One particular case, concerning the motion of the NTDs of ClpB,[94] is discussed here, but other studies showing, for example, that the proteasome is an ensemble of interconverting conformations, each with distinct proteolytic activities,[36] establish the unique role of NMR in studies of large, functionally dynamic systems.

Central to many NMR studies of high-molecular-weight protein complexes has been the development of methods for labeling proteins with $^{13}CH_3$ methyl groups in an otherwise highly deuterated background[9] and methyl-TROSY-based approaches,[7] as described here for ClpB. There are too many applications to discuss in detail presently, but a seminal recent study that is worthy of mention in passing focuses on how the trigger factor chaperone recognizes and maintains nascent polypeptide chains in an unfolded conformation after they emerge from the ribosome.[39] This work provides a wonderful example of the unique role that NMR can play in understanding "dynamic" binding involving an unfolded protein and a molecular chaperone.

Despite the ongoing success of NMR studies of molecular machines over the past several years, there are, of course, aspects that can be improved. As the drive for understanding increasingly complex targets continues, methyl labeling by itself will not be

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466–1477

sufficient, in particular with respect to spectral resolution, as is readily apparent by Figure 4E. One approach has been to make use of precursors that label only pro-*R* or pro-*S* prochiral methyl groups of Val and Leu in proteins.[10] A particularly powerful strategy, and one that will become necessary as the protomer size increases, is segmental labeling, as illustrated in studies of the NTD domain of the intact ClpB hexamer described here. There are a number of powerful methods for generating such molecules, involving either intein-based[102,103,108] or, more recently, sortase-based approaches[101,109,110] that provide the spectroscopist with different tools for simplifying spectra. Many of the large systems that are currently studied by solution NMR techniques are symmetric homo-oligomers.[11,27,35−37,40,41,43,44,87,94,111−116] Relative to hetero-oligomeric complexes, these tend to be particularly advantageous for NMR because spectral complexity is minimized by having multiple copies of the same protomer. In cases where hetero-oligomeric machines are studied, labeling one protomer at a time is likely the method of choice if *in vitro* assembly of the components is possible.[36,117−119] If this is not feasible, a recently developed Lego-NMR approach[120] that enables separate labeling of components *in vivo* shows real promise.

An additional important area of development includes addressing issues regarding the assignment of methyl groups to specific sites in the protein, which remains a current bottleneck in high-resolution NMR studies of molecular machines. Typically assignments are carried out by a "divide-and-conquer" approach whereby individual domains of larger complexes are first assigned in isolation via traditional approaches, with the assignments subsequently transferred to the complex,[30,44,121] or via a combined mutagenesis/NOE-driven approach that takes advantage of high-resolution structural data from other techniques.[40] Pseudo-contact shifts (PCSs) that arise from the interaction of nuclear spins with one or more unpaired electrons introduced to the system of interest[122] can also be exploited for assignment where high-resolution structures are available. Examples include applications to both small[123] and large protein systems.[121] The development of more rigid tags for adding unpaired electrons to the protein under study,[124,125] coupled with approaches whereby tags are positioned at different locations in the protein or where a series of different paramagnetic metals are used to supply the unpaired electrons, may lead to significant advances in methyl group assignment in high-molecular-weight complexes. It is also worth noting that computational approaches[126,127] that take into account all of the available types of NMR data that can be measured are likely to become increasingly important to the assignment process.

In looking back at the progress that has been made over the past several decades in solution NMR studies of proteins, one is reminded of the seemingly unending potential of the methodology. There is constant innovation at the level of methods development from spin physics[128] to labeling,[8,79] to data recording and processing,[129−133] along with new ways of maximally using the restraints obtained[134−138] to study, for example, the "unstructured biology" of intrinsically disordered proteins.[137] If history is any indicator, new advances will continue to decrease molecular weight limitations, enabling even more focused views in the future of how dynamic molecular machines are able to carry out their fascinating functions.

AUTHOR INFORMATION

Corresponding Authors

*rina.rosenzweig@utoronto.ca

*kay@pound.med.utoronto.ca

Notes

The authors declare no competing financial interest.

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.

REFERENCES

(1) Wright, M. A.; Aprile, F. A.; Arosio, P.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. *Chem. Commun. (Cambridge, U. K.)* **2015**, *51*, 14425.

(2) Kim, Y. E.; Hipp, M. S.; Bracher, A.; Hayer-Hartl, M.; Hartl, F. U. *Annu. Rev. Biochem.* **2013**, *82*, 323.

(3) Morimoto, R. I. *Cold Spring Harbor Symp. Quant. Biol.* **2011**, *76*, 91.

(4) Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. *Nature* **2011**, *475*, 324.

(5) Saibil, H. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 630.

(6) Burmann, B. M.; Hiller, S. *Prog. Nucl. Magn. Reson. Spectrosc.* **2015**, *86−87*, 41.

(7) Rosenzweig, R.; Kay, L. E. *Annu. Rev. Biochem.* **2014**, *83*, 291.

(8) Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Mei Ono, A.; Guntert, P. *Nature* **2006**, *440*, 52.

(9) Tugarinov, V.; Kanelis, V.; Kay, L. E. *Nat. Protoc.* **2006**, *1*, 749.

(10) Gans, P.; Hamelin, O.; Sounier, R.; Ayala, I.; Dura, M. A.; Amero, C. D.; Noirclerc-Savoye, M.; Franzetti, B.; Plevin, M. J.; Boisbouvier, J. *Angew. Chem., Int. Ed.* **2010**, *49*, 1958.

(11) Freiburger, L.; Sonntag, M.; Hennig, J.; Li, J.; Zou, P.; Sattler, M. *J. Biomol. NMR* **2015**, *63*, 1.

(12) Lichtenecker, R.; Ludwiczek, M. L.; Schmid, W.; Konrat, R. *J. Am. Chem. Soc.* **2004**, *126*, 5348.

(13) Fernandez, C.; Wider, G. *Curr. Opin. Struct. Biol.* **2003**, *13*, 570.

(14) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420.

(15) Miclet, E.; Williams, D. C., Jr; Clore, G. M.; Bryce, D. L.; Boisbouvier, J.; Bax, A. *J. Am. Chem. Soc.* **2004**, *126*, 10560.

(16) Schanda, P.; Kupce, E.; Brutscher, B. *J. Biomol. NMR* **2005**, *33*, 199.

(17) Lescop, E.; Schanda, P.; Brutscher, B. *J. Magn. Reson.* **2007**, *187*, 163.

(18) Riek, R.; Wider, G.; Pervushin, K.; Wuthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 4918.

(19) Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366.

(20) Fiaux, J.; Bertelsen, E. B.; Horwich, A. L.; Wuthrich, K. *Nature* **2002**, *418*, 207.

(21) Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K. *J. Am. Chem. Soc.* **1998**, *120*, 6394.

(22) Brutscher, B.; Simorre, J. P. *J. Biomol. NMR* **2001**, *21*, 367.

(23) Riek, R.; Pervushin, K.; Fernandez, C.; Kainosho, M.; Wuthrich, K. *J. Am. Chem. Soc.* **2001**, *123*, 658.

(24) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. *J. Biomol. NMR* **1999**, *13*, 369.

(25) Ayala, I.; Hamelin, O.; Amero, C.; Pessey, O.; Plevin, M. J.; Gans, P.; Boisbouvier, J. *Chem. Commun. (Cambridge, U. K.)* **2012**, *48*, 1434.

(26) Sinha, K.; Jen-Jacobson, L.; Rule, G. S. *Biochemistry* **2011**, *50*, 10189.

(27) Velyvis, A.; Kay, L. E. *J. Am. Chem. Soc.* **2013**, *135*, 9259.

(28) Tugarinov, V.; Kay, L. E. *J. Biomol. NMR* **2004**, *28*, 165.

(29) Ayala, I.; Sounier, R.; Use, N.; Gans, P.; Boisbouvier, J. *J. Biomol. NMR* **2009**, *43*, 111.

(30) Gelis, I.; Bonvin, A. M.; Keramisanou, D.; Koukaki, M.; Gouridis, G.; Karamanou, S.; Economou, A.; Kalodimos, C. G. *Cell* **2007**, *131*, 756.

(31) Isaacson, R. L.; Simpson, P. J.; Liu, M.; Cota, E.; Zhang, X.; Freemont, P.; Matthews, S. *J. Am. Chem. Soc.* **2007**, *129*, 15428.

(32) Fischer, M.; Kloiber, K.; Hausler, J.; Ledolter, K.; Konrat, R.; Schmid, W. *ChemBioChem* **2007**, *8*, 610.

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

(33) Zhuravleva, A.; Clerico, E. M.; Gierasch, L. M. *Cell* **2012**, *151*, 1296.
(34) Bista, M.; Freund, S. M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15752.
(35) Imai, S.; Osawa, M.; Mita, K.; Toyonaga, S.; Machiyama, A.; Ueda, T.; Takeuchi, K.; Oiki, S.; Shimada, I. *J. Biol. Chem.* **2012**, *287*, 39634.
(36) Ruschak, A. M.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E3454.
(37) Rosenzweig, R.; Moradi, S.; Zarrine-Afsar, A.; Glover, J. R.; Kay, L. E. *Science* **2013**, *339*, 1080.
(38) Tseng, R.; Chang, Y. G.; Bravo, I.; Latham, R.; Chaudhary, A.; Kuo, N. W.; Liwang, A. *J. Mol. Biol.* **2014**, *426*, 389.
(39) Saio, T.; Guan, X.; Rossi, P.; Economou, A.; Kalodimos, C. G. *Science* **2014**, *344*, 1250494.
(40) Shi, L.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 2140.
(41) Karagoz, G. E.; Duarte, A. M.; Akoury, E.; Ippel, H.; Biernat, J.; Moran Luengo, T.; Radli, M.; Didenko, T.; Nordhues, B. A.; Veprintsev, D. B.; Dickey, C. A.; Mandelkow, E.; Zweckstetter, M.; Boelens, R.; Madl, T.; Rudiger, S. G. *Cell* **2014**, *156*, 963.
(42) Xiao, Y.; Warner, L. R.; Latham, M. P.; Ahn, N. G.; Pardi, A. *Biochemistry* **2015**, *54*, 4307.
(43) Zheng, X.; Perera, L.; Mueller, G. A.; DeRose, E. F.; London, R. E. *eLife* **2015**, *4*, 06359.
(44) Sprangers, R.; Kay, L. E. *Nature* **2007**, *445*, 618.
(45) Mittermaier, A.; Kay, L. E. *Science* **2006**, *312*, 224.
(46) Ishima, R.; Torchia, D. A. *Nat. Struct. Biol.* **2000**, *7*, 740.
(47) Palmer, A. G., III; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *339*, 204.
(48) Sekhar, A.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 12867.
(49) Anthis, N. J.; Clore, G. M. *Q. Rev. Biophys.* **2015**, *48*, 35.
(50) Nikolova, E. N.; Kim, E.; Wise, A. A.; O'Brien, P. J.; Andricioaei, I.; Al-Hashimi, H. M. *Nature* **2011**, *470*, 498.
(51) Duncan, E. J.; Cheetham, M. E.; Chapple, J. P.; van der Spuy, J. *Subcell. Biochem.* **2015**, *78*, 243.
(52) Aguado, A.; Fernandez-Higuero, J. A.; Moro, F.; Muga, A. *Arch. Biochem. Biophys.* **2015**, *580*, 121.
(53) Mogk, A.; Kummer, E.; Bukau, B. *Front. Mol. Biosci.* **2015**, *2*, 22.
(54) Mayer, M. P. *Trends Biochem. Sci.* **2013**, *38*, 507.
(55) Rudiger, S.; Buchberger, A.; Bukau, B. *Nat. Struct. Biol.* **1997**, *4*, 342.
(56) Clerico, E. M.; Tilitsky, J. M.; Meng, W.; Gierasch, L. M. *J. Mol. Biol.* **2015**, *427*, 1575.
(57) Seyffer, F.; Kummer, E.; Oguchi, Y.; Winkler, J.; Kumar, M.; Zahn, R.; Sourjik, V.; Bukau, B.; Mogk, A. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1347.
(58) Weibezahn, J.; Tessarz, P.; Schlieker, C.; Zahn, R.; Maglica, Z.; Lee, S.; Zentgraf, H.; Weber-Ban, E. U.; Dougan, D. A.; Tsai, F. T.; Mogk, A.; Bukau, B. *Cell* **2004**, *119*, 653.
(59) Kityk, R.; Kopp, J.; Sinning, I.; Mayer, M. P. *Mol. Cell* **2012**, *48*, 863.
(60) Qi, R.; Sarbeng, E. B.; Liu, Q.; Le, K. Q.; Xu, X.; Xu, H.; Yang, J.; Wong, J. L.; Vorvis, C.; Hendrickson, W. A.; Zhou, L.; Liu, Q. *Nat. Struct. Mol. Biol.* **2013**, *20*, 900.
(61) Mayer, M. P.; Bukau, B. *Cell. Mol. Life Sci.* **2005**, *62*, 670.
(62) Sekhar, A.; Rosenzweig, R.; Bouvignies, G.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 10395.
(63) Vallurupalli, P.; Bouvignies, G.; Kay, L. E. *J. Am. Chem. Soc.* **2012**, *134*, 8148.
(64) Nishikawa, T.; Nagadoi, A.; Yoshimura, S.; Aimoto, S.; Nishimura, Y. *Structure* **1998**, *6*, 1057.
(65) Marsh, J. A.; Singh, V. K.; Jia, Z.; Forman-Kay, J. D. *Protein Sci.* **2006**, *15*, 2795.
(66) Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wuthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13585.
(67) Fawzi, N. L.; Ying, J.; Ghirlando, R.; Torchia, D. A.; Clore, G. M. *Nature* **2011**, *480*, 268.
(68) Forsen, S.; Hoffman, R. A. *J. Chem. Phys.* **1963**, *39*, 2892.
(69) Ward, K. M.; Aletras, A. H.; Balaban, R. S. *J. Magn. Reson.* **2000**, *143*, 79.
(70) Zhou, J.; van Zijl, P. C. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 109.
(71) Bouvignies, G.; Vallurupalli, P.; Kay, L. E. *J. Mol. Biol.* **2014**, *426*, 763.
(72) McConnell, H. M. *J. Chem. Phys.* **1958**, *28*, 430.
(73) Bertelsen, E. B.; Chang, L.; Gestwicki, J. E.; Zuiderweg, E. R. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8471.
(74) Lee, S.; Sowa, M. E.; Watanabe, Y. H.; Sigler, P. B.; Chiu, W.; Yoshida, M.; Tsai, F. T. *Cell* **2003**, *115*, 229.
(75) Sielaff, B.; Tsai, F. T. *J. Mol. Biol.* **2010**, *402*, 30.
(76) Miot, M.; Reidy, M.; Doyle, S. M.; Hoskins, J. R.; Johnston, D. M.; Genest, O.; Vitery, M. C.; Masison, D. C.; Wickner, S. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 6915.
(77) Sprangers, R.; Velyvis, A.; Kay, L. E. *Nat. Methods* **2007**, *4*, 697.
(78) Wider, G.; Wuthrich, K. *Curr. Opin. Struct. Biol.* **1999**, *9*, 594.
(79) Ruschak, A. M.; Kay, L. E. *J. Biomol. NMR* **2010**, *46*, 75.
(80) Werbelow, L. G.; Marshall, A. G. *J. Magn. Reson.* **1973**, *11*, 299.
(81) Kay, L. E.; Torchia, D. A. *J. Magn. Reson. (1969)* **1991**, *95*, 536.
(82) Nicholson, L. K.; Kay, L. E.; Baldisseri, D. M.; Arango, J.; Young, P. E.; Bax, A.; Torchia, D. A. *Biochemistry* **1992**, *31*, 5253.
(83) Kay, L. E.; Muhandiram, D. R.; Farrow, N. A.; Aubin, Y.; Forman-Kay, J. D. *Biochemistry* **1996**, *35*, 361.
(84) Atreya, H. S., Ed. *Isotope Labeling in Biomolecular NMR*; Springer: Berlin, 2012.
(85) Mogk, A.; Schlieker, C.; Strub, C.; Rist, W.; Weibezahn, J.; Bukau, B. *J. Biol. Chem.* **2003**, *278*, 17615.
(86) Battiste, J. L.; Wagner, G. *Biochemistry* **2000**, *39*, 5355.
(87) Religa, T. L.; Sprangers, R.; Kay, L. E. *Science* **2010**, *328*, 98.
(88) Clore, G. M.; Iwahara, J. *Chem. Rev.* **2009**, *109*, 4108.
(89) Dominguez, C.; Boelens, R.; Bonvin, A. M. *J. Am. Chem. Soc.* **2003**, *125*, 1731.
(90) de Vries, S. J.; van Dijk, A. D.; Krzeminski, M.; van Dijk, M.; Thureau, A.; Hsu, V.; Wassenaar, T.; Bonvin, A. M. *Proteins: Struct., Funct., Genet.* **2007**, *69*, 726.
(91) Glover, J. R.; Lindquist, S. *Cell* **1998**, *94*, 73.
(92) Mogk, A.; Tomoyasu, T.; Goloubinoff, P.; Rudiger, S.; Roder, D.; Langen, H.; Bukau, B. *EMBO J.* **1999**, *18*, 6934.
(93) Fersht, A. *Structure and Mechanism in Protein Science*; W. H. Freeman and Company: New York, 1999.
(94) Rosenzweig, R.; Farber, P.; Velyvis, A.; Rennella, E.; Latham, M. P.; Kay, L. E. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 6872.
(95) Schlieker, C.; Weibezahn, J.; Patzelt, H.; Tessarz, P.; Strub, C.; Zeth, K.; Erbse, A.; Schneider-Mergener, J.; Chin, J. W.; Schultz, P. G.; Bukau, B.; Mogk, A. *Nat. Struct. Mol. Biol.* **2004**, *11*, 607.
(96) Doyle, S. M.; Hoskins, J. R.; Wickner, S. *J. Biol. Chem.* **2012**, *287*, 28470.
(97) Beinker, P.; Schlee, S.; Groemping, Y.; Seidel, R.; Reinstein, J. *J. Biol. Chem.* **2002**, *277*, 47160.
(98) Roberts, G., Lian, L.-Y., Eds. *Protein NMR Spectroscopy: Practical Techniques and Applications*; Wiley: New York, 2011; p 1.
(99) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76.
(100) Shah, N. H.; Muir, T. W. *Chem. Sci.* **2014**, *5*, 446.
(101) Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2670.
(102) Skrisovska, L.; Schubert, M.; Allain, F. H. *J. Biomol. NMR* **2010**, *46*, 51.
(103) Muona, M.; Aranko, A. S.; Raulinaitis, V.; Iwai, H. *Nat. Protoc.* **2010**, *5*, 574.
(104) Carroni, M.; Kummer, E.; Oguchi, Y.; Wendler, P.; Clare, D. K.; Sinning, I.; Kopp, J.; Mogk, A.; Bukau, B.; Saibil, H. R. *eLife* **2014**, *3*, e02481.
(105) Sun, H.; Kay, L. E.; Tugarinov, V. *J. Phys. Chem. B* **2011**, *115*, 14878.
(106) Lee, S.; Choi, J. M.; Tsai, F. T. *Mol. Cell* **2007**, *25*, 261.
(107) Palmer, A. G., 3rd; Massi, F. *Chem. Rev.* **2006**, *106*, 1700.
(108) Cowburn, D.; Shekhtman, A.; Xu, R.; Ottesen, J. J.; Muir, T. W. *Methods Mol. Biol.* **2004**, *278*, 47.

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477

(109) Levary, D. A.; Parthasarathy, R.; Boder, E. T.; Ackerman, M. E. *PLoS One* **2011**, *6*, e18342.
(110) Dorr, B. M.; Ham, H. O.; An, C.; Chaikof, E. L.; Liu, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 13343.
(111) Religa, T. L.; Ruschak, A. M.; Rosenzweig, R.; Kay, L. E. *J. Am. Chem. Soc.* **2011**, *133*, 9063.
(112) Street, T. O.; Zeng, X.; Pellarin, R.; Bonomi, M.; Sali, A.; Kelly, M. J.; Chu, F.; Agard, D. A. *J. Mol. Biol.* **2014**, *426*, 2393.
(113) Neu, A.; Neu, U.; Fuchs, A. L.; Schlager, B.; Sprangers, R. *Nat. Chem. Biol.* **2015**, *11*, 697.
(114) Hagn, F.; Lagleder, S.; Retzlaff, M.; Rohrberg, J.; Demmer, O.; Richter, K.; Buchner, J.; Kessler, H. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1086.
(115) Park, S. J.; Kostic, M.; Dyson, H. J. *J. Mol. Biol.* **2011**, *411*, 158.
(116) Burmann, B. M.; Wang, C.; Hiller, S. *Nat. Struct. Mol. Biol.* **2013**, *20*, 1265.
(117) Kato, H.; van Ingen, H.; Zhou, B. R.; Feng, H.; Bustin, M.; Kay, L. E.; Bai, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 12283.
(118) Mund, M.; Neu, A.; Ullmann, J.; Neu, U.; Sprangers, R. *J. Mol. Biol.* **2011**, *414*, 165.
(119) Audin, M. J.; Dorn, G.; Fromm, S. A.; Reiss, K.; Schutz, S.; Vorlander, M. K.; Sprangers, R. *Angew. Chem., Int. Ed.* **2013**, *52*, 8312.
(120) Mund, M.; Overbeck, J. H.; Ullmann, J.; Sprangers, R. *Angew. Chem., Int. Ed.* **2013**, *52*, 11401.
(121) Velyvis, A.; Schachman, H. K.; Kay, L. E. *J. Am. Chem. Soc.* **2009**, *131*, 16534.
(122) Bertini, I.; Luchinat, C.; Parigi, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 249.
(123) John, M.; Schmitz, C.; Park, A. Y.; Dixon, N. E.; Huber, T.; Otting, G. *J. Am. Chem. Soc.* **2007**, *129*, 13749.
(124) Hass, M. A.; Ubbink, M. *Curr. Opin. Struct. Biol.* **2014**, *24*, 45.
(125) Fawzi, N. L.; Fleissner, M. R.; Anthis, N. J.; Kalai, T.; Hideg, K.; Hubbell, W. L.; Clore, G. M. *J. Biomol. NMR* **2011**, *51*, 105.
(126) Chao, F. A.; Kim, J.; Xia, Y.; Milligan, M.; Rowe, N.; Veglia, G. *J. Magn. Reson.* **2014**, *245*, 17.
(127) Xu, Y.; Liu, M.; Simpson, P. J.; Isaacson, R.; Cota, E.; Marchant, J.; Yang, D.; Zhang, X.; Freemont, P.; Matthews, S. *J. Am. Chem. Soc.* **2009**, *131*, 9480.
(128) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Rance, M.; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed.; Academic Press: San Diego, 1996.
(129) Kazimierczuk, K.; Orekhov, V. *Magn. Reson. Chem.* **2015**, *53*, 921.
(130) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(131) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603.
(132) Hyberts, S. G.; Milbradt, A. G.; Wagner, A. B.; Arthanari, H.; Wagner, G. *J. Biomol. NMR* **2012**, *52*, 315.
(133) Maciejewski, M. W.; Mobli, M.; Schuyler, A. D.; Stern, A. S.; Hoch, J. C. *Top. Curr. Chem.* **2011**, *316*, 49.
(134) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4685.
(135) Shen, Y.; Vernon, R.; Baker, D.; Bax, A. *J. Biomol. NMR* **2009**, *43*, 63.
(136) Wishart, D. S.; Arndt, D.; Berjanskii, M.; Tang, P.; Zhou, J.; Lin, G. *Nucleic Acids Res.* **2008**, *36*, W496.
(137) Krzeminski, M.; Marsh, J. A.; Neale, C.; Choy, W. Y.; Forman-Kay, J. D. *Bioinformatics* **2013**, *29*, 398.
(138) Cavalli, A.; Vendruscolo, M. *J. Biomol. NMR* **2015**, *62*, 503.

DOI: 10.1021/jacs.5b11346
J. Am. Chem. Soc. 2016, 138, 1466−1477