Quantitative dynamics and binding studies of the 20S proteasome by NMR

Remco Sprangers & Lewis E. Kay

The machinery used by the cell to perform essential biological processes is made up of large molecular assemblies. One such complex, the proteasome, is the central molecular machine for removal of damaged and misfolded proteins from the cell. Here we show that for the 670-kilodalton 20S proteasome core particle it is possible to overcome the molecular weight limitations that have traditionally hampered quantitative nuclear magnetic resonance (NMR) spectroscopy studies of such large systems. This is achieved by using an isotope labelling scheme where isoleucine, leucine and valine methyls are protonated in an otherwise highly deuterated background in concert with experiments that preserve the lifetimes of the resulting NMR signals. The methodology has been applied to the 20S core particle to reveal functionally important motions and interactions by recording spectra on complexes with molecular weights of up to a megadalton. Our results establish that NMR spectroscopy can provide detailed insight into supra-molecular structures over an order of magnitude larger than those routinely studied using methodology that is generally applicable.

The proteasome plays a critical role in the degradation of proteins that regulate crucial cellular processes such as cell division, gene expression, signal transduction and apoptosis and is a target for anti-cancer drugs. The 670-kDa 20S core particle (CP) consists of four heptameric rings arranged in an \( \alpha_7 \)-\( \beta_7 \)-\( \beta_7 \)-\( \alpha_7 \) fashion (Fig. 1a). The two outer \( \alpha_7 \) rings form the entrance channel for substrates and the binding sites for the poly-ubiquitin-recognizing \( 19S \) proteasome and the immune-response-related \( 11S \) activator. The interfaces between the \( \alpha_7 \) and \( \beta_7 \) rings enclose the catalytic chamber, where the active sites are sequestered within the lumen of the barrel. Although high-resolution structures of the 20S proteasome from all three domains of life (eukarya, bacteria, archaea) are available, little direct information on the internal motions of the CP can be obtained from these static pictures. However, dynamics are likely to be critical for the function of this enzyme, especially for the \( \alpha \)-domains that gate the entrance of targets for proteolysis.

Methyl transverse relaxation optimized spectroscopy (TROSY)

In general, the structural features of many stable complexes can be determined by electron microscopy and X-ray crystallography, but insight into functionally important motions and transient interactions are more difficult to obtain. NMR spectroscopy is especially suited to the study of dynamics and dynamic interactions over a broad spectrum of timescales; however, applications have been limited to relatively small proteins. The study of megadalton-sized complexes by solution NMR spectroscopy is challenging and requires careful consideration of what probes to use (isotope labelling), along with optimization of experiments that specifically account for the spin properties of these probes. We have recently shown that very-high-quality \( ^1H \), \( ^13C \) methyl spectra can be obtained for highly deuterated proteins that are protonated at isoleucine-\( \delta_1 \), valine and leucine methyl positions (Fig. 1b–e). Here we have used this labelling scheme, along with methyl-TROSY spectroscopy that optimizes both sensitivity and resolution in spectra. Figure 1b shows the heteronuclear multiple quantum coherence (HMHC) spectrum of the isoleucine-\( \delta_1 \) methyl region of the archaeabacterium Thermoplasma acidophilum CP, which was recorded on a 25-\( \mu \)M sample (0.35-mM monomer concentration) at 65 °C in 90 min on an 800-MHz spectrometer with a room-temperature probehead. With the exception of one cross-peak, all of the expected 17 (\( \alpha \)-ring) and 14 (\( \beta \)-ring) correlations from isoleucine residues are observed in the spectrum. There are a number of practical advantages in studying the \( T. \) acidophilum CP. First, it is structurally very similar to eukaryotic versions. Second, all of the \( \alpha \)-subunits are equivalent in this archaeal CP (as are the \( \beta \)-subunits), leading to improvements in both sensitivity and resolution in NMR spectra. Third, its thermal stability allows studies at elevated temperatures, lowering the molecular tumbling time \( t_c \) and thus improving spectral quality (\( t_c \approx 180 \) ns at 65 °C; by comparison for a 15-kDa protein at 30 °C, a \( t_c \) value of \( \approx 8 \) ns would be expected). Fourth, this 205 CP can be reconstituted from separately expressed subunits and it is therefore possible to restrict isotopic labelling to either of the \( \alpha \)- or \( \beta \)-rings, making the other subunits invisible to NMR.

Methyl chemical shift assignments

Here we report the nearly complete chemical shift assignment for isoleucine-\( \delta_1 \), leucine and valine methyl groups in the \( \alpha \)-subunits of the 670-kDa CP, despite the fact that backbone-directed triple-resonance NMR methods fail. The strategy used to solve this assignment problem involves ‘dissecting’ the large system into smaller building blocks (Supplementary Fig. 1). The chemical shifts of the residues in the core of these building blocks will not be affected much by the disruption of the complex provided that the smaller units fold in a similar manner in isolation and in the complex.

Here, the 670-kDa 20S proteasome CP (Fig. 1a–c) was dissected into the 360-kDa double-\( \alpha \)-ring particle (\( \alpha_2\alpha_7 \), Fig. 1a, d) that spontaneously forms from wild-type (WT) \( \alpha \) monomers in the absence of \( \beta \) subunits. Subsequently, a monomeric form of the \( \alpha \) protein (\( \alpha_{\text{monomer}} \), Fig. 1a, e) was generated by introducing mutations that prevent formation of \( \alpha_2\alpha_7 \) (see Methods). This 21-kDa \( \alpha_{\text{monomer}} \) was assigned using standard triple-resonance TROSY-based NMR experiments. Both secondary chemical shifts and nuclear coherence assignments were determined for this protein. The results have been presented graphically in Figs. 2 and 3 for all isoleucine-\( \delta_1 \)-, leucine- and valine-methyl regions.

©2007 Nature Publishing Group

Departments of Biochemistry, Medical Genetics and Chemistry, The University of Toronto, Toronto, Ontario M5S 1A8, Canada.
Overhauser effect (NOE) cross-peaks confirm that the structure of the α-subunit of the proteasome is similar to that of the α-subunit in the crystal structure of the full 20S CP, and indeed, spectra of all three forms of the complex (Fig. 1a) show similarities (Fig. 1c–e, Supplementary Fig. 2). Despite the large molecular weight of the 360-kDa αβγδ complex, correlation maps linking methyl groups with the carbons one and two bonds removed could be recorded using methyl-TROSY versions of previously developed correlated spectroscopy (COSY) experiments (2a). Four chemical shifts ([H]methyl-13Cmethyl and two for 13C-skelet-chain) could therefore be compared between α-subunit and αβγδ where the correspondence is high, assignments of methyl groups could be transferred from α-subunit to αβγδ (Fig. 2a) and confirmed on the basis of NOE patterns (Fig. 2b).

In the next step (Supplementary Fig. 1), additional assignments were obtained for methyl groups proximal to those already assigned, using NOE correlations and the high-resolution crystal structure. Successively, methyl assignments from αβγδ were transferred to the full 670-kDa αβγδ proteasome, based on methyl chemical shifts (Fig. 1c, d), NOE patterns between residues (Fig. 2b) and NOE cross-peaks within residues (Fig. 2c). Finally, resonances for which no NOE or COSY correlations were observed were assigned via mutagenesis.

In our hands this approach is only successful once the majority of assignments are available, because mutations can have substantial effects on the chemical shifts of neighbouring methyl groups (Supplementary Fig. 3). The combined analysis of the data produced assignments for 95% (92 out of 97) and 89% (86 out of 97) of the isoleucine, leucine, and valine methyl groups of αβγδ, and 91% (αβγδ) of these methyl-containing residues are available for the study of site-specific dynamics and interactions.

**Probing molecular dynamics**

Insights into pico- to nanosecond timescale side-chain dynamics in the proteasome were obtained by recording the relaxation properties of both 2H and 13C spins in 13CH2-labelled samples. Figure 3a shows the correlation between 2H- and 13C-derived order parameters squared S2 for the αβγδ complex, which quantify the amplitude of motion (with the amplitude decreasing as S2 increases). We observe an excellent correlation between the two independent measurements, despite the size of the complex (360 kDa, 50 °C, t1/2 = 180 s), so that the measures of the dynamics can be interpreted with confidence (Supplementary Fig. 4).

Values of S2 were also obtained for the 670-kDa CP (65 °C, t1/2 = 180 s) and show a very strong correlation with those in αβγδ (Supplementary Fig. 5). Thus, the β-rings have little effect on either the structure or the dynamics of the α-rings, apart from residues at the αβ interface that are less flexible in the context of the β-rings (for example, V107, V113). We note that although most residues showing increased temperature factors do experience fast timescale mobility, there is little correlation between crystallographic B-factors and S2 values (Supplementary Fig. 6).

Plotting S2 on the structure of the proteasome reveals a cluster of flexible residues on the outside of the barrel, between individual α subunits (Fig. 3b). Interestingly, α-subunits of the CP from *Mycobacterial tuberculosis* and *Rhodococcus* as well as the eukaryotic 26S subunits show sequence deletions in part of this region. This mobile cluster carries a potential nuclear localization-type signal. Although the archaea-bacterium does not possess a nucleus, the flexible signal sequence from the *T. acidophilum* CP has been shown to promote nuclear localization in human cell lines and is conserved in the α2-subunit of the mammalian proteasome. It may be that the mobile sequence was present before nuclei evolved and was the target for different adaptor proteins early in evolution. Irrespective of the identities of the binding partners the dynamics at this site probably play a part in moderating the affinity of interactions by providing an entropic barrier to binding.

Micro- to millisecond timescale dynamics are often correlated with protein function so we studied these motions in the proteasome.
using relaxation dispersion experiments. $^1$H-$^13$C multiple-quantum methyl-TROSY dispersion profiles$^4$ (Supplementary Fig. 7) measured for both $\gamma_2\gamma_2$ and the CP fitted well to the simplest model of chemical exchange ($A \leftrightarrow B$) and similar exchange values $k_{ex}$ ($k_{ex} = k_A + k_B$) were obtained for all residues, which suggests that the source of the exchange involves a single global process. Values of $k_{ex}$ of 1,200 ± 170 and 1,580 ± 90 s$^{-1}$ were obtained for $\gamma_2\gamma_2$ and the CP, respectively, with the small differences probably reflecting the different temperatures at which the experiments were performed (50 °C versus 65 °C). Interestingly, the exchanging residues cluster on the inside of the antechamber and form a surface from the entrance pore of the proteasome towards the catalytic chamber (Fig. 3c); this highly dynamic environment could indirectly facilitate the movement of substrate to the sites of proteolysis by presenting an entropic barrier to binding to the walls of both the antechamber and the entrance pores that would prevent the deposit of protein in these regions. Whether the correlated motions in the antechamber are specific for the proteasome or are more widespread in oligomeric molecular machines has still to be determined; however, studies of many protein complexes are now within the range of these NMR approaches.

In addition to residues inside the antechamber, the amino-terminal regions of the $\gamma$ subunits are mobile on the millisecond timescale (V14 is extremely broadened and no resonance could be assigned to 112, with spectra of WT and 112A CP showing no changes in the isoleucine-$\delta_1$ region). The 12 N-terminal residues are not observed in the crystal structure of the archaeal proteasome$^1$, but the corresponding residues in the yeast and bovine proteasomes$^5, 6$ form a stable conformation that closes the entrance channel and functions as a gate$^{11}$. NMR dispersion data on a mutant protein that lacks residues 1 to 12 (the $\Delta$1–12 CP) establish that these residues have little effect on the observed dynamics inside the antechambers. However, a comparison of HMQC spectra of the WT and $\Delta$1–12 CP show chemical shift changes at the narrowest point of the substrate entrance path (V129; Fig. 3b, c) and inside the antechamber that suggest that the 12 N-terminal residues fold into the antechamber through the entrance channel to form a primitive gate (Fig. 3d). This observation is consistent with electron microscopy data that shows density for the residues around the entrance pore$^{25}$, but goes beyond the available structural data by showing that at least some of these dynamic termini probably fold into the antechamber. It also explains the fact that these ‘invisible’ gating residues influence peptide degradation rates$^{25}$. Interestingly, some of the N-terminal residues of the eukaryotic proteasome also fold back into the lumen of the antechamber$^{9, 10}$ and thereby block substrate entrance$^{11}$.

**Quantitative studies of binding**

Protein machines do not function in isolation and it is important to understand the quaternary organization of these large complexes. NMR is a particularly powerful tool in this regard because many interactions in transient complexes are too weak to be detected and characterized using other structural techniques. Unfortunately, the size limitations associated with traditional NMR experiments have limited progress in the study of large complexes. Recently, however, $^1$H-$^13$N cross-correlated relaxation-induced polarization transfer (CRIPPT) spectroscopy was used to establish the binding site on the 70-kDa GroES of the 800-kDa GroEL$^{26}$.

Here we used methyl assignments to determine the residues of the 670-kDa CP that interact with the 11S activator complex from *Trypanosoma brucei* in the 1.1 MDa 11S-$\alpha_2\beta_2\gamma_2$-11S complex$^5$. Stepwise addition of the 11S complex to the CP leads to the disappearance of a subset of resonances and to the appearance of a new set of resonances (Fig. 4a). Residues affected by binding are

---

**Figure 2 | Details of the assignment strategy.** a, Side-chain carbon chemical shifts obtained from an (H)C(CO)NH TOCSY (total correlation spectroscopy) experiment of the $\gamma$monomer (left two strips) correlate well with the side-chain chemical shifts in $\gamma_2\gamma_2$ obtained from out-and-back $H_\text{methy}C_\text{methy}$ and $H_\text{methy}C_\text{methy}(C_\text{y})C_\text{y}$ methyl-TROSY, COSY-based experiments$^4$ (right four strips). The green stars in the COSY strips indicate the positions of the diagonal peaks that are either not observed or are very weak in the experiment. The labelling pattern and magnetization transfer for the methyl-detected experiments are indicated. b, NOEY strips showing correlations between proximal methyls of V46 and L148, along with the labelling pattern and magnetization transfer (during the 250-ms mixing time). We note that there are no intra-residue NOE cross-peaks, because only one of the two methyl groups is labelled for each residue. c, NOESY strips from an experiment recorded on the $\gamma_3\beta_3\gamma_2$ CP (50-ms mixing time) where both methyls of leucine and valine residues in the $\gamma$-rings are labelled as indicated, giving cross-peaks between methyl groups of the same residue.
In summary, this study establishes that quantitative information of the sort normally reserved for applications to small proteins in solution can be obtained on systems more than an order of magnitude larger than those usually studied by NMR spectroscopy. A combined approach in which X-ray crystallography is used to produce detailed static pictures, followed by NMR studies of motion and ligand interactions promises to be particularly fruitful for the study of supramolecular structures. In particular, by exploiting the methodological advances described here, we have identified and quantified specific sites of motion that may facilitate substrate localization to the catalytic chamber, and potentially modulate interactions of the nuclear localization-type signal with receptors. We have also presented data that is consistent with the N termini of the α-subunits that make up the dynamic gate folding through the entrance pore into the antechamber. In addition, the 11S binding site has been mapped and its affinity measured. Future quantitative studies of the interaction of the proteasome with dynamical substrates, such as α-synuclein, and with substrates inside the lumen of the protease are now feasible, and will lead to additional insights into the function of the proteasome.

Figure 3 | Quantification of dynamics and structure. a, Correlation between $^2\Delta$C values extracted independently from carbon$^{13}$ and deuterium$^2$ relaxation experiments. Errors ($\pm 1$ standard deviation) are indicated with bars. b, Top view of the proteasome with methyl groups coloured according to $^2\Delta$C. A cluster of very flexible side chains is located between the individual monomers (circle, arrow), with the variable loop containing the potential nuclear localization-type signal indicated by an arrow (see text). c, Cross-section of the side view of the proteasome that reveals the inside (lumen). Residues undergoing concerted motion are shown in red and form a continuous cluster in the antechamber. The exchange-broadened V14 in the flexible N terminus is shown in yellow, while the active site threonines are indicated in blue. d, Side view highlighting residues that change chemical shift upon truncation of the first 12 residues of the α subunit (chemical shift differences $\Delta\nu$ between WT and Δ1–12, defined as $(\Delta\nu_{\alpha} + \Delta\nu_{\beta})^{1/2}$ at 800 MHz $^1$H resonance frequency). The location of residues 13–18 in the crystal structure (green) is shown and suggests that the invisible N-terminal 12 residues are located in the lumen of the antechamber. Although the chemical shift differences were measured on $^{13}$C and $^2$H for clarity the complete proteasome is shown.

Figure 4 | Interaction between the 20S proteasome and the 11S activator. a, Methyl-TROSY spectra of $\alpha_2\alpha_7$ (top) and the CP (bottom) during addition of the 11S activator complex. The complex is in slow exchange on the NMR chemical shift timescale; assignments of correlations from the complex are not available. b, Residues whose resonances are affected by the 11S interaction are mapped on the proteasome structure. c, Crystal structure of the 11S-proteasome complex. d, The intensities of resonances during the titration are used to obtain an approximate dissociation constant ($K_D = 12 \pm 10 \mu M$) for the $\alpha_2\alpha_7$/11S interaction that is consistent with the CP titration data (see insets). The decrease in intensity of one of the correlations from L81 is shown in green and the concomitant increase in a ‘bound’ peak indicated in red. Errors are quantified from signal-to-noise in spectra. [Ligand] and [Protein] refer to total ligand and protein concentrations.
METHODS
Sample preparation. Proteins were produced using labelled precursors and growth media as described in the Supplementary Information. All proteins were purified on Ni-NTA resin (Qiagen), followed by cleavage of the purification tag using tobacco etch virus (TEV) protease and gel filtration. The full proteasome was obtained from reconstitution of separately purified NMR-active α-subunits and NMR-inactive β-subunits or from co-expression of both subunits. Proteasome assembly was verified by electron microscopy (data not shown). Final protein concentrations of the NMR samples were between 0.1 and 2.8 mM in monomer (see Supplementary Information) in 100% D2O (or 90% H2O/10% D2O), 25 mM potassium phosphate, pH 6.8, 50 mM NaCl, 1 mM EDTA, 0.03% NaN3, and 2 mM DTT.

NMR spectroscopy. NMR spectra (see also Supplementary Information) were recorded at 30 °C (7αα, 7αβ) or 65 °C (CP = 7αββα) using 600 and 800 MHz Varian Inova spectrometers equipped with cryogenically cooled (600 MHz) or room-temperature (800 MHz) pulsed-field gradient triple-resonance probes; all data sets of the CP were measured at 800 MHz. 1H and 13C chemical shift assignments are referenced against DSS (2,2-dimethyl-2-silapentane-5-sulfonic acid). Backbone and side-chain assignments of ηαα (2Δ–34, R57A, R66A, R130A) were achieved using standard TROSY-based triple-resonance methods11,17. Stereo-specific assignments were made using a 10%13C labelled protein1,12 and, where possible, transferred to 7αβ and the CP on the basis of peak positions and relative intensities in NOE spectra. Methyl-TROSY COSY-based out-and-back data sets were measured on the 7αβ sample using pulse schemes that are very similar to those previously published1,12, with the exception of the methyl 13C evolution period, which was of the HMQC variety. All NOE spectra were recorded using methyl-TROSY-based schemes with 13C chemical shifts obtained in (7αβ, 7αβ) and with mixing times as described in Supplementary Information. Details of relaxation experiments are given in Supplementary Information.

Chemical-shift titration experiments were recorded on a 14 mM 7αβ complex (0.2 mM monomer concentration) and on a 9 mM CP complex (0.13 mM monomer concentration) by stepwise addition of 0.2 mM 11S activator complex (0.2 mM monomer concentration) and Kp values extracted as described in Supplementary Information. All data were processed with the NmrPipe/NmrDraw suite of programs12. Figures displaying molecular structures were made with PyMol (www.pymol.sourceforge.net).

Received 28 July; accepted 6 December 2006.

Published online 21 January 2007.


Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Forman-Kay for discussions and for providing laboratory space, F. Hansen for discussions, J. Rubenstein for electron microscopy images, R. Muhandiram for NMR support and C. Hill for a plasmid of the 11S activator complex. R.S. acknowledges EMBO and the Canadian Institutes of Health Research (CIHR) Training Program in Protein Folding and Disease for fellowships. L.E.K. holds a Canada Research Chair in Biochemistry. Grant support from CIHR and NSERC is acknowledged.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.E.K. (kay@pound.med.utoronto.ca).