Supporting Information

Tracing an allosteric pathway regulating the activity of the HslV protease

Lichi Shi[1] and Lewis E. Kay[1,2]

[1]Departments of Molecular Genetics, Biochemistry and Chemistry, The University of Toronto, Toronto, Ontario, M5S 1A8, Canada

[2]Hospital for Sick Children, Program in Molecular Structure and Function, 555 University Avenue, Toronto, Ontario, M5G 1X8, Canada

Cloning, Expression, and Purification of Proteins The *H. influenzae* HslV gene (1) was cloned into a pET29b$^+$ vector between NdeI and XhoI restriction sites. All mutants were prepared with *PfuTurbo* DNA polymerase using the QuikChange method. Transformed *E. coli* BL21(DE3) cells were grown on 99.9% D_2O, M9 medium using 3g/L d7-glucose as the sole carbon source. Production of [U-^{2}H; Ile$^{\delta1}$-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{12}CD_3$; Met-$^{13}CH_3$; Thr-$^{13}CH_3$]-labeled protein samples was achieved by addition of the following precursors: 100 mg/L methyl-$^{13}CH_3$ Met (Met labeling, (2, 3)), 50mg/L α-ketobutyric acid (Ile labeling, (4)), 60mg/L [α, β-^{2}H; γ-^{13}C]-Thr and 100mg/L d5-glycine (U-[^{2}H,^{12}C], γ2-[$^{13}CH_3$] Thr labeling) (5), 80mg/L α-ketoisovaleric acid ([$^{13}CH_3$,$^{12}CD_3$] Leu, Val labeling, not stereospecific (6)) or 170mg/L 2-hydroxy-2-methyl-D_3-3-oxobutanoate-4-^{13}C ($\delta1,\gamma1$-[$^{13}CH_3$] Leu, Val labeling, stereospecific; (7)) were added 1 hour before induction, as described

previously (8, 9). In cases where proteins with only a subset of label were desired, only the requisite precursors were included in the growth media. Proteins were expressed at 30°C overnight after induction with 0.6mM IPTG at A_{600}=0.8 OD. Cell lysate was sequentially passed through an anion exchange column (Mono Q HR 10/10, Pharmacia), a hydrophobic interaction column (Phenyl Superose HR 10/10, Pharmacia) and a Gel Filtration column (Superdex 200 16/60, Pharmacia) to purify the untagged HslV protein.

NMR experiments All NMR measurements were performed at 40 °C, using 18.8T or 14.0T Varian Inova spectrometers equipped with room-temperature (18.8T) or cryogenically cooled (14.0T) pulsed-field gradient triple-resonance probes. The NMR buffer was 75 mM KCl, 1 mM EGTA, 5mM $MgCl_2$, 20 mM HEPES pH 7.6 and 0.05% azide in 99.9% D_2O for all of the reported experiments except for the pH titration, where a mixture of buffering reagents (monosodium citrate, dipotassium phosphate, Tris and boric acid each at 25 mM) was used in addition to 1 mM EGTA, 5mM $MgCl_2$ and 0.05% azide. ^{13}C-^{1}H correlation spectra were recorded using the HMQC experiment, exploiting a methyl-TROSY effect that is particularly beneficial for applications to high molecular weight proteins (10).

Carr-Purcell-Meiboom-Gill (CPMG) ^{13}C relaxation dispersion experiments were recorded on samples labeled as [U-^{2}H; Ile$^{\delta1}$-$^{13}CH_3$; Leu,Val-$^{13}CH_3$(*pro-R*)/$^{12}CD_3$(*pro-S*) or Leu,Val-$^{13}CH_3$/$^{12}CD_3$; Met-$^{13}CH_3$; Thr-$^{13}CH_3$], at 18.8 and 14.0T, using a previously published pulse scheme (11) and analyzed as described previously (8, 12). Briefly, a constant-time relaxation delay of 20 ms was used, that included a variable number of ^{13}C refocusing pulses. Data sets were recorded with

CPMG frequencies ranging from 50 Hz to 1000 Hz. ^{1}H CPMG relaxation dispersion profiles were measured on a WT HslV sample labeled as [U-^{2}H; Ile$^{\delta1}$-$^{13}CHD_2$; Leu,Val-$^{13}CHD_2$,$^{13}CHD_2$; Met-$^{13}CHD_2$], using a previously published pulse scheme (13). A constant-time relaxation delay of 10 ms was used, along with CPMG frequencies ranging from 100 Hz to 1000 Hz. Dispersion data were fitted using an in-house written program (ChemEx, available upon request) assuming a two-site global exchange model.

Translational diffusion coefficients were measured by recording a series of 1D ^{13}C-edited spectra, 25^{o}C, using a pulse scheme that is similar to an ^{15}N-edited experiment published previously with ^{15}N and ^{13}C pulses interchanged (14). A constant-delay diffusion period (after initial encoding of magnetization) was set to 200 ms. The resulting ^{1}H signal was integrated over the methyl ^{1}H frequency range to obtain intensities as a function of encoding/decoding gradient strength.

Rapid (ps-ns) time-scale dynamics were quantified on $^{13}CH_3$-methyl labeled HslV samples using a ^{1}H triple-quantum based relaxation violated coherence transfer scheme, described previously (15). Relaxation delays (during which time differential ^{1}H relaxation occurs) of 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 10.0, 12.0, 14.0 ms were used. A correlation time for the assumed isotropically tumbling HslV complex was set to 84 ns, 40 °C, based on the measured value of 120 ns for the *T. acidophilum* $\alpha_7\alpha_7$ complex, 50 °C (8) taking into account the viscosity dependence of D_2O as a function of temperature, as well as the molecular weight differences between the two particles.

Methyl Group Assignment Over 95% of the Ile (11/11), Leu (18/19), Met (3/3), Thr (11/11) and Val (15/16) correlations in spectra of HslV, 40^{o}C, were assigned by a combined mutagenesis and NOE analysis that made use of the crystal structure of the HslV dodecamer (1) (PDB ID: 1G3K). This approach was used, rather than a divide and conquer methodology that was successful in the case of the 20S CP (8, 9), because of difficulties in producing single subunits of HslV. After several months of effort, including 26 single mutations targeting 11 positions involved in inter-subunit salt bridges and hydrophobic interactions, a mutagenesis strategy was adopted, as described below.

Each of 44 selected methyl containing residues (all I,L,M,V except M66, M94, I120, I123, I148) were mutated one at a time, but only 26 of them (I96V, I106V, I109V, I155Q, I159V, I169V, L22I, L37I, L43I, L59I, L71V, L72V, L88I, L91I, L95I, L104V, L121I, L153I, M27A, V12I, V34I, V42A, V76A, V138I, V149A, V161I) led to unambiguous assignments. Interestingly, L78 could not be detected as a new peak for the L78I mutant in the Ile region was not observed and no peak disappeared in the Leu region. Other mutants either did not express soluble protein or samples led to spectra of very poor quality that could not be used for assignment. Methyl-TROSY based 3D NOE experiments (pulse programs available upon request) that record methyl ^{13}C,^{1}H chemical shifts as, C[i]-H[i]-NOE-H[j] or C[i]-NOE-C[j]-H[j] were obtained on a 1 mM [U-^{2}H; Ile$^{\delta 1}$-$^{13}CH_3$; Leu,Val-$^{13}CH_3$(pro-R),$^{12}CD_3$(pro-S), Met-$^{13}CH_3$]-labeled sample (subunit concentration) with a mixing time of 250 ms. A total of 75 NOE links could be established, where all the 4 possible pathways were observed between the two proximal methyl groups (C[i]-H[i]-NOE-H[j],C[j]-H[j]-

NOE-H[i] and C[j]-NOE-C[i]-H[i],C[i]-NOE-C[j]-H[j]). Another 21 residues (2 Met, 5 Ile, 6 Leu, 8 Val) were assigned, with all of the 75 NOEs interpreted as originating from a pair of methyl groups within 11Å (C-C distance) based on the crystal structure. In order to link intra-residue methyl groups (i.e. *pro-R*/*pro-S* for Leu,Val) a second sample with [$^{13}CH_3$,$^{13}CH_3$] labeling of L,V methyl groups was prepared and an NOE data set recorded with a short mixing time (50 ms). Threonine assignments were obtained by NOE analysis of data sets recorded on a [U-^{2}H; Ile$^{\delta 1}$-$^{13}CH_3$; Leu,Val-$^{13}CH_3$(*pro-R*)/$^{12}CD_3$(*pro-S*); Met-$^{13}CH_3$; Thr-$^{13}CH_3$]-labeled sample (250 ms mixing time), as described in the legend to Supporting Information Figure 3.

As a final note we wish to emphasize that the production of multiple mutants for assignment, at least in this case, was not prohibitively expensive. Central to the success of this approach was the use of a codon-optimized plasmid (OptimumGene™ GenScript USA Inc.), so that protein yields on the order of 100 mg per L of D_2O-based M9 bacterial culture could be obtained. As a result all the mutants were expressed in 100 mL culture, with final NMR sample (monomer) concentrations ranging from 0.3 mM to 1.0 mM. The cost of producing each sample (D_2O, deuterated glucose, precursors for methyl $^{13}CH_3$ labeling) was roughly $60 and high quality NMR spectra could be obtained for the majority of the mutants within approximately 1-2 hours.

Peptidase assay HslV protease activity was measured by using a fluorogenic peptide substrate, N-carbobenzoxy-Gly-Gly-Leu-amido-4-methylcoumarin (Z-GGL-AMC). A SpectraMax M5 microplate reader was used (96 well plates), with each reaction followed for 10 minutes in 20-second increments. Each well contains 90 μL

substrate at different concentrations incubated at 40 °C in D_2O based NMR buffer or pH titration buffer + 20% DMSO. 10µL of HslV protease (usually at a concentration of 1 mg/mL, but higher/lower for less/more active mutants) was added to start the reaction. The measured fluorescence (λ_{ex}=380 nm, λ_{em}=460 nm) was converted to product concentration by first reacting a known amount of substrate with excess protease, with the saturating fluorescence proportional to the product concentration (equal to the starting substrate concentration). All the reactions were in the steady-state regime and a constant reaction rate could be extracted from each kinetics data set. The v-[S] curve was fitted to a simple Michaelis-Menten equation, but only the catalytic efficiency k_{cat}/K_m could be reliably extracted as K_M >> [S]. The added DMSO (20%), necessary to keep the Z-GGL-AMC substrate soluble at high concentration, did not affect protease activity (the activity was identical when either 5% or 20% DMSO was present for substrate concentrations up to 50µM). Each reaction was repeated three times to estimate the standard errors in the rates.

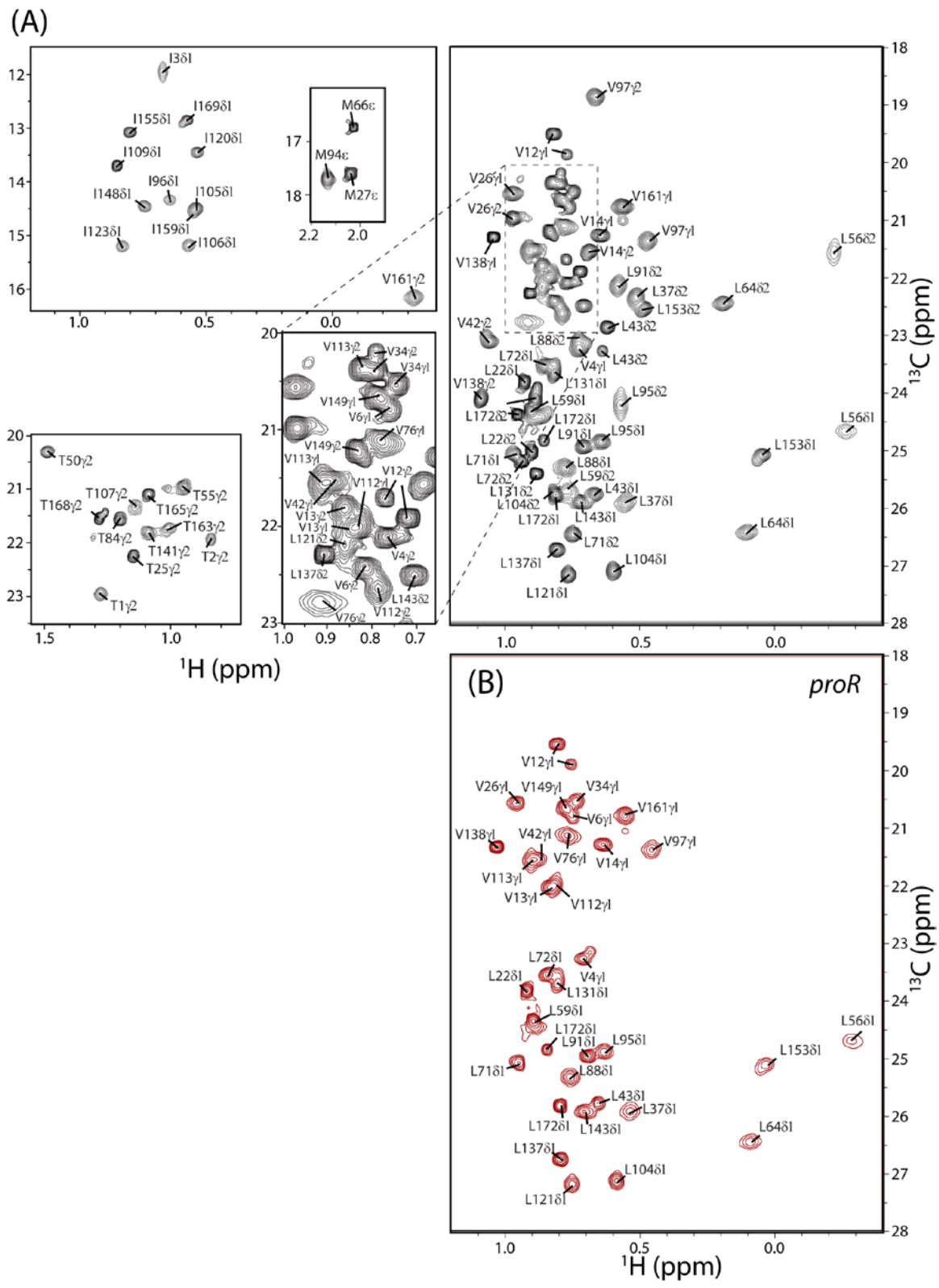

(A)
(B)
proR
^{1}H (ppm)
^{13}C (ppm)

Figure S1. (A) ^{13}C,^{1}H HMQC spectra of [U-^{2}H; Ile$^{\delta1}$-$^{13}CH_3$; Leu,Val-$^{13}CH_3$/$^{12}CD_3$; Met-$^{13}CH_3$]-labeled HslV or [U-^{2}H; Ile$^{\delta1}$-$^{13}CH_3$; Met-$^{13}CH_3$; Thr-$^{13}CH_3$]-labeled HslV, 18.8T, 40 °C, with assignments as indicated. (B) Portion of ^{13}C-^{1}H correlation map of WT [U-^{2}H; Ile$^{\delta1}$-$^{13}CH_3$; Leu,Val-$^{13}CH_3$(*proR*)/$^{12}CD_3$(*proS*)]-labeled HslV, 18.8 T, 40^{o}C. Correlations are labeled as shown.

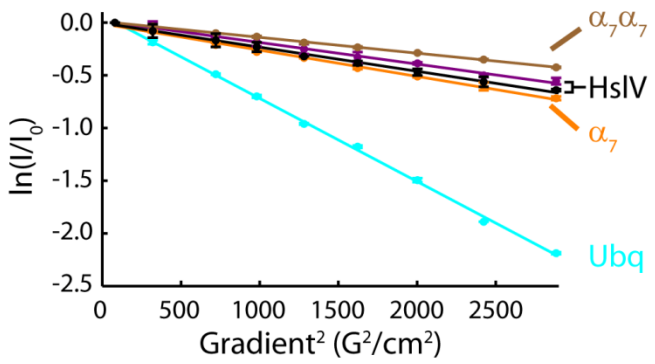
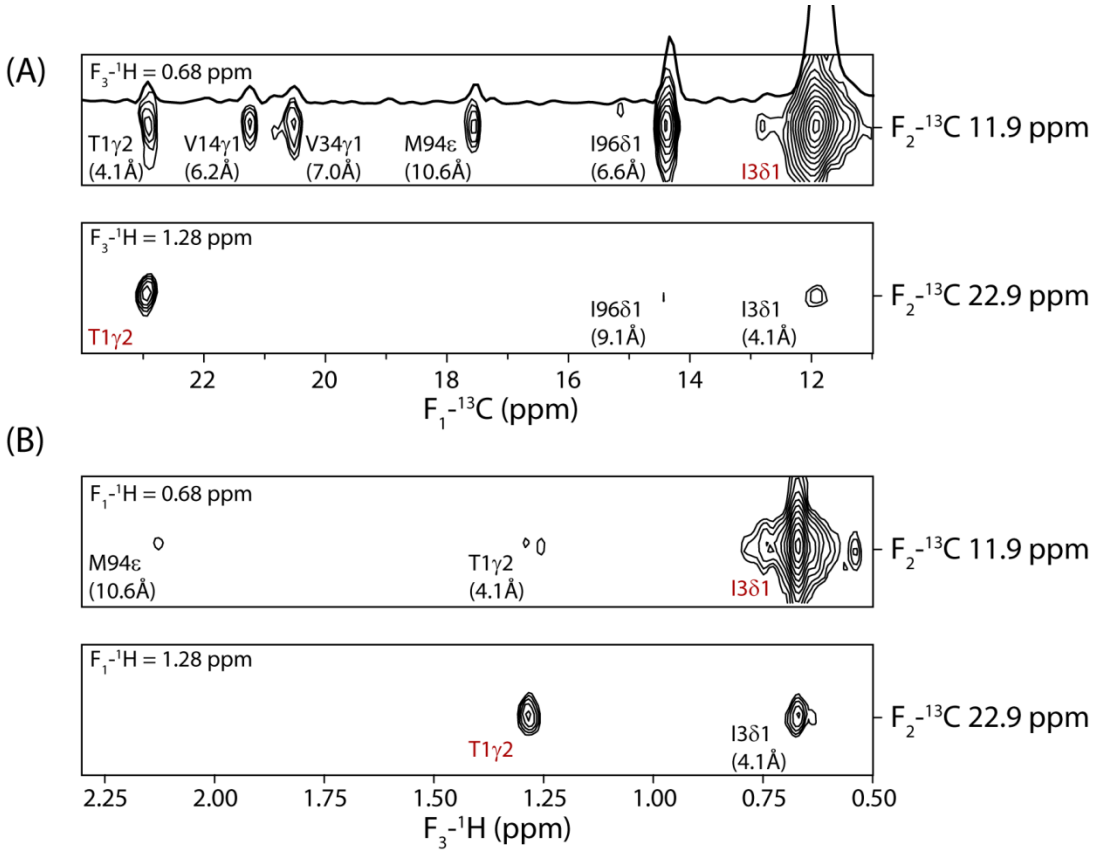


Figure S2. The hydrodynamic radius of HslV is within the expected range for a particle with molecular weight 230 kDa. Translational diffusion coefficients of WT and mutant HslVs were measured by pulsed field gradient NMR spectroscopy, using a 1D ^{13}C-edited pulse scheme that selects for the methyl groups in a $^{13}CH_3$-methyl labeled sample (see NMR Experiments). The slope of the ln(intensity) *vs.* gradient profile in each of the curves is proportional to the diffusion constant or inversely proportional to the cube root of the molecular weight, assuming that each of the diffusing particles is spherical. Translational diffusion constants were measured on samples of (1) the ½ proteasome from *T. acidophilum*, $\alpha_7\alpha_7$, corresponding to a pair of heptameric α-rings (MW=360 kDa), (2) α_7, a single α ring (MW=180 kDa), (3) Ubq, Ubiquitin (MW=8.9 kDa) as standards, and subsequently on samples of HslV. The two extreme cases for the HslV samples (L88I purple, WT black) are shown. Diffusion coefficients for HslV are (10^{-7} cm^2/s): L71V (3.9±0.2), L88I (3.6±0.2), L91I (3.9±0.2), I159V (4.0±0.4), WT (4.1±0.2).


(A)
F3-1H = 0.68 ppm
T1γ2 (4.1Å)
V14γ1 (6.2Å)
V34γ1 (7.0Å)
M94ε (10.6Å)
I96δ1 (6.6Å)
I3δ1
F2-13C 11.9 ppm
F3-1H = 1.28 ppm
T1γ2
I96δ1 (9.1Å)
I3δ1 (4.1Å)
F2-13C 22.9 ppm
22 20 18 16 14 12
F1-13C (ppm)
(B)
F1-1H = 0.68 ppm
M94ε (10.6Å)
T1γ2 (4.1Å)
I3δ1
F2-13C 11.9 ppm
F1-1H = 1.28 ppm
T1γ2
I3δ1 (4.1Å)
F2-13C 22.9 ppm
2.25 2.00 1.75 1.50 1.25 1.00 0.75 0.50
F3-1H (ppm)

(C)

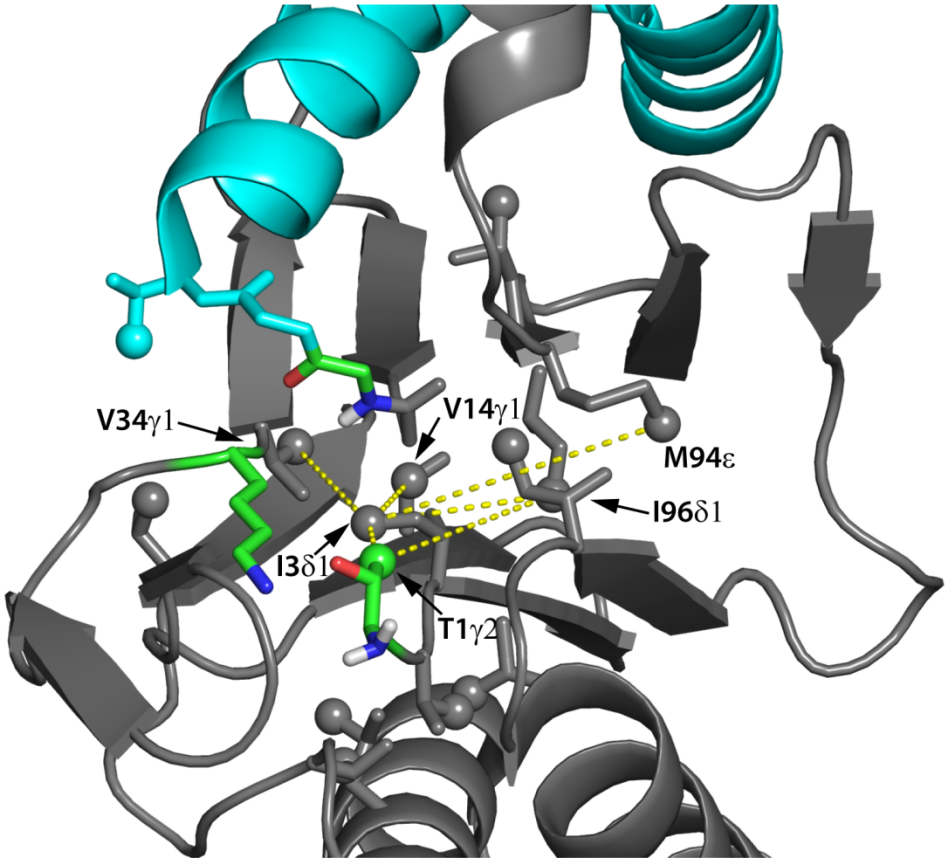

V34γ1
V14γ1
M94ε
I96δ1
I3δ1
T1γ2

Figure S3. Assignments based on NOE patterns: T1 as an example. (A) Selected planes from a 3D ^{13}C-^{13}C-^{1}H NOE experiment, recording the ^{13}C chemical shift of the originating methyl in F_1, followed by an NOE period (250 ms) and the ^{13}C-^{1}H chemical shifts of the destination methyl group in F_2-F_3, respectively. The top panel highlights NOEs from methyl groups (F_1) proximal to I3δ1, whose shifts are recorded in F_2/F_3. An NOE peak is observed between the methyl groups of T1 and I3 (δ1), as expected (4.1Å separation). The symmetry related correlation connecting I3δ1 to T1γ2 is shown in the second plane. (B) Planes recorded from a 3D ^{13}C-edited NOESY data set (^{1}H-^{13}C-NOE-^{1}H). The NOEs connecting T1 and I3 are highlighted. The distances between the methyl groups in the figure are indicated in parentheses. Methyl-TROSY based data sets were recorded on a [U-^{2}H; Ile$^{\delta1}$-$^{13}CH_3$; Leu,Val-$^{13}CH_3$(*proR*)/$^{12}CD_3$(*proS*); Met-$^{13}CH_3$; Thr-$^{13}CH_3$]-labeled HslV sample 18.8 T, 40^{o}C (1 mM in monomer). In all slices the methyl group from which the diagonal peak derives is identified in red. (C) Portion of the structure of an HslV monomer (1) focusing on the contacts between T1γ2 and adjacent methyl groups.

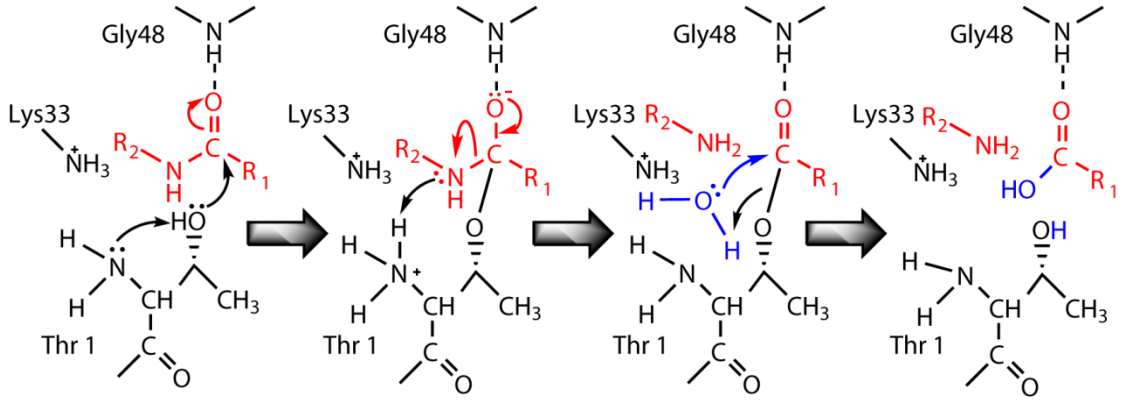


Figure S4. Proposed mechanism of HslV catalyzed proteolysis (1, 16). Key HslV residues, T1, K33 and G48 are highlighted, along with the substrate (red) and a water molecule that is believed to participate in the reaction. Adapted from ref (17).

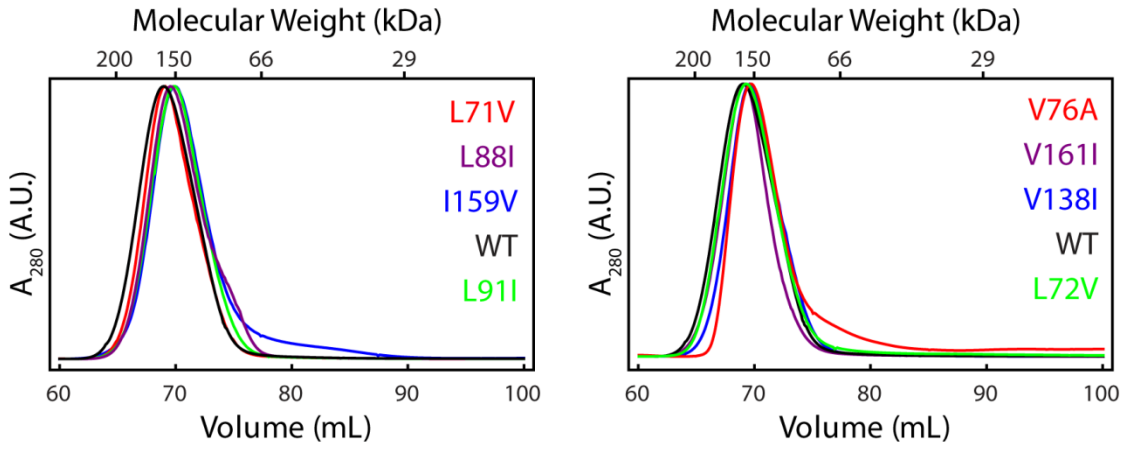


Figure S5. Similar gel-filtration elution profiles are obtained for WT and mutant HslVs, suggesting identical oligomeric structures for the proteins used in the present study. The elution peak positions for a mixture of protein standards and the molecular weights of each standard (amylase 200 kDa, alcohol dehydrogenase 150 kDa, albumin 66 kDa, and carbonic anhydrase 29 kDa) are as indicated. All HslV constructs examined have very similar elution profiles. Notably HslV elutes later than expected, likely reflecting the shape of the double ring structure.

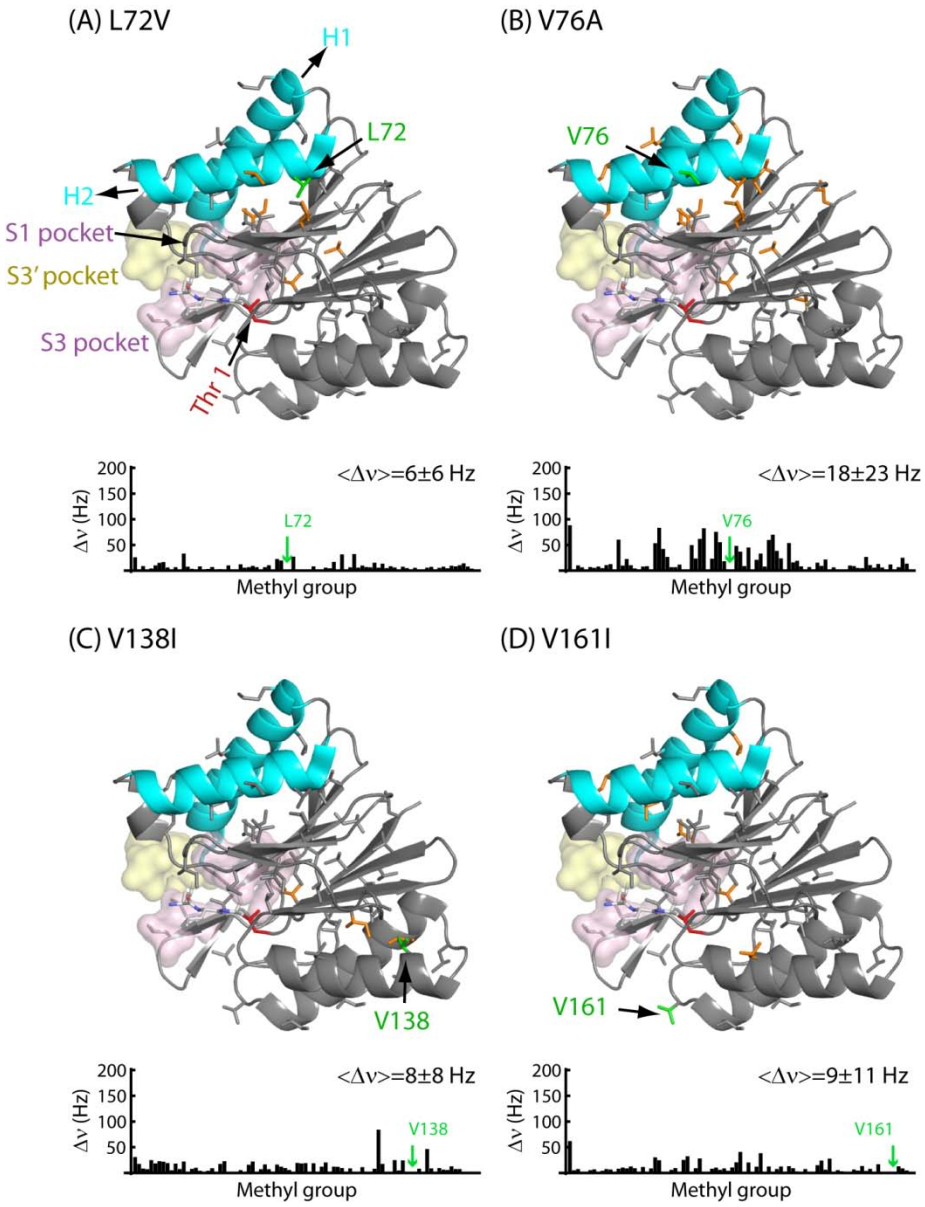


Figure S6. Allosteric chemical shift perturbations via mutagenesis at positions (A) 72 (L72V), (B) 76 (V76A), (C) 138 (V138I) and (D) 161 (V161I). Chemical shift changes relative to the wild-type protein are indicated in bar chart form and color-coded on the structure. Details are as for Figure 3 of the main text.

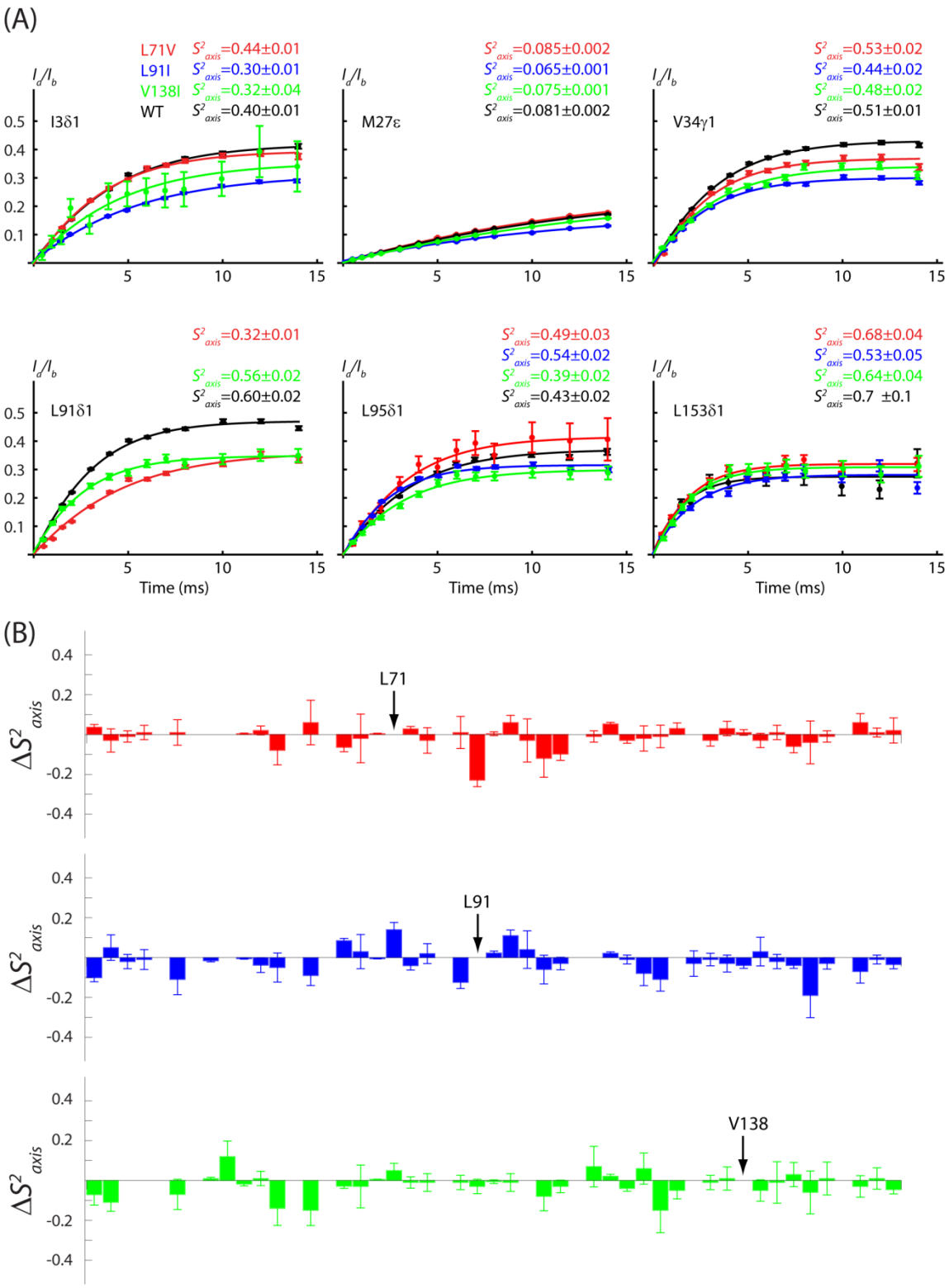


Figure S7. Measurement of differential amplitudes of motion of the methyl group symmetry axes (S^2_{axis}) in L71V (20x active), L91I (12 x less active) and V138I (~ wt

activity) HslV. Values of S^2_{axis} were obtained from an approach that measures the time dependencies of sums (I_b) and differences (I_a) of magnetization derived from methyl ^{1}H single-quantum transitions, as described by Sun *et al.* (15) (A) Ratio of I_a and I_b as a function of build-up time for selected methyl groups in each of the mutants examined. Note that build-up rates increase with increasing values of S^2_{axis}. (B) $\Delta S^2_{axis} = S^2_{axis}(\text{mutant}) - S^2_{axis}(\text{WT})$ values are plotted as a function of methyl group (x-axis) for mutant = L71V (red), L91I(blue), V138I (green).

REFERENCES

1. Sousa MC*, et al.* (2000) Crystal and solution structures of an HslUV protease-chaperone complex. *Cell* 103(4):633-643.
2. Gelis I*, et al.* (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131(4):756-769.
3. Religa TL, Sprangers R, & Kay LE (2010) Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science* 328(5974):98-102.
4. Goto NK, Gardner KH, Mueller GA, Willis RC, & Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. *Journal of biomolecular NMR* 13(4):369-374.
5. Velyvis A, Ruschak AM, & Kay LE (2012) An economical method for production of (2)H, (13)CH3-threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. *PloS one* 7(9):e43725.
6. Tugarinov V & Kay LE (2004) An isotope labeling strategy for methyl TROSY spectroscopy. *Journal of biomolecular NMR* 28(2):165-172.
7. Gans P*, et al.* (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chem Int Ed Engl* 49(11):1958-1962.
8. Sprangers R & Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* 445(7128):618-622.
9. Ruschak AM & Kay LE (2012) Proteasome allostery as a population shift between interchanging conformers. *Proceedings of the National Academy of Sciences of the United States of America* 109(50):E3454-3462.
10. Tugarinov V, Hwang PM, Ollerenshaw JE, & Kay LE (2003) Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *Journal of the American Chemical Society* 125(34):10420-10428.
11. Korzhnev DM, Kloiber K, & Kay LE (2004) Multiple-quantum relaxation dispersion NMR spectroscopy probing millisecond time-scale dynamics in proteins: theory and application. *Journal of the American Chemical Society* 126(23):7320-7329.
12. Korzhnev DM, Kloiber K, Kanelis V, Tugarinov V, & Kay LE (2004) Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *Journal of the American Chemical Society* 126(12):3964-3973.
13. Baldwin AJ, Religa TL, Hansen DF, Bouvignies G, & Kay LE (2010) 13CHD2 methyl group probes of millisecond time scale exchange in proteins by 1H relaxation dispersion: an application to proteasome gating residue dynamics. *Journal of the American Chemical Society* 132(32):10992-10995.

14. Choy WY*, et al.* (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *Journal of molecular biology* 316(1):101-112.
15. Sun H, Kay LE, & Tugarinov V (2011) An optimized relaxation-based coherence transfer NMR experiment for the measurement of side-chain order in methyl-protonated, highly deuterated proteins. *The journal of physical chemistry. B* 115(49):14878-14884.
16. Lowe J*, et al.* (1995) Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. *Science* 268(5210):533-539.
17. Velyvis A & Kay LE (2013) Measurement of active site ionization equilibria in the 670 kDa proteasome core particle using methyl-TROSY NMR. *Journal of the American Chemical Society* 135(25):9259-9262.