Supporting Information

Measurement of active site ionization equilibria in the 670 kDa proteasome core particle using methyl-TROSY NMR

Algirdas Velyvis and Lewis E. Kay

Sample Preparation: Plasmids expressing the wild type β-subunit (βWT), the β subunit T1C mutant (βT1C) and the S95C version of the α-subunit (αS95C) of *T. acidophilum* proteasome have been described previously[1,2]. All proteasome samples used in this study were constructed by mixing αS95C subunits with various β-subunits produced as described below. Mutant versions of β-subunits (βT3S,T16S,T44S and βT-5S,T-2S,D166N) were prepared by using a QuikChange kit (Agilent, Inc.) with the appropriate primers. The D166N mutation results in an inactive proteasome, so that the prosequence that is normally cleaved during the final stage of proteasome maturation remains intact (prosequence residues are denoted by negative numbers, starting from G-1). We have added the Thr to Ser pair of mutations in βT-5S,T-2S,D166N to remove the possibility of interference in spectra from what would presumably be very intense resonances from T-5 and T-2.

U-[^{2}H] α–subunit was prepared as described before[1,3] and very briefly involved protein expression in *E. coli* BL21(DE3) CodonPlus(RIL) cells in 99%-D_2O M9 medium with 3 g/L d_7-glucose and purification via Ni-chelating resin. The His-tag was subsequently cleaved with TEV protease, followed by further purification of the α-subunit on a Superdex 200 gel-filtration column.

U-[^{2}H], Ileδ1-$^{13}CH_3$, Thrγ2-$^{13}CH_3$ labeled β–subunits were prepared by a previously published protocol[1,3]. Briefly, protein was expressed in *E. coli* BL21(DE3) CodonPlus(RIL) cells in 99%-D_2O M9 medium with 3 g/L d_7-glucose, and by addition of 50 mg/L of [α-^{2}H;β-^{2}H;γ-^{13}C]-Thr, 50 mg/L of alpha-ketobutyric acid

([methyl-$^{13}CH_3$] labeled, deuterated at the 3 position), and 100 mg/L of d_5-glycine 1 hour prior to the induction of protein expression. The resultant protein was purified on a Ni-chelating resin, and subsequently the NusA- and His-tags were cleaved with TEV protease, with further protein purification on a MonoQ column, 50 mM Tris pH=8. Purified α- and β-subunits were combined to reconstitute the 20S CP, incubated for 12-24 hours at 37 °C and purified on a Superdex 200 gel-filtration column. Samples of βT1C and βT3S,T16S,T44S proteasomes were exchanged into 99.9% D_2O buffer with 25 mM K-phosphate, 50 mM NaCl, 1 mM EDTA, and 0.3 g/L NaN_3 pH*=6.8.

For the pH titrations reported in the text proteasome samples were dissolved in a 99.9% D_2O, 25 mM Na-malonate, 25 mM citric acid, 25 mM NaH_2PO_4, 25 mM Tris and 25 mM boric acid, that has buffering capacity over the range of pHs used to generate the titration profiles. All NMR samples contained 0.5-1 mM DSS for chemical shift referencing. Notably, the activity of wild-type proteasome in the pH titration buffer, as measured by the hydrolysis of the peptide Suc-LLVY-AMC, is not changed relative to phosphate- or tris-only buffers. Samples were adjusted to pH* 10.0 before the start of the titration with ~1M NaOD in D_2O and subsequently adjusted during the course of the titration by the addition of small (1-10 μL) volumes of ~1M DCl in D_2O. pH* values were measured at room temperature using a H_2O-filled glass electrode calibrated against standard H_2O buffers at pH=4.0, 7.0 and 10.0.

All ^{1}H-^{13}C HMQC spectra were measured at 70 °C, 18.8 T. Each spectrum of the pH titration profile was recorded in approximately 1 hour (0.5 mM protein

subunit concentration). Data sets were processed using the NMRPipe suite of programs[4] and analyzed using NMRView[5]. Uncertainties in measured chemical shifts were determined by calculating the standard deviation of chemical shifts (separately for ^{1}H and ^{13}C) measured for Ile202 (that does not titrate) from the 8 points of the pH titration profile.

Analysis of NMR titration data for a system with two titratable sites

Figure 3A illustrates a general dissociation scheme involving a pair of titratable groups with microscopic dissociation constants, K_i ($i \in \{a,b,c,d\}$). As described in the text the protonation state of each group is denoted as '0' (deprotonated) or '1' (protonated), so that '00' indicates a completely deprotonated molecule. It is straightforward to show that $K_aK_c = K_bK_d$, so that only three of the K_i values are independent. It follows from the scheme in Figure 3A that

$$K_a = \frac{[00][H^+]}{[10]},\ K_b = \frac{[00][H^+]}{[01]},\ K_c = \frac{[10][H^+]}{[11]},\ K_d = \frac{[01][H^+]}{[11]} \qquad \text{[S1]}$$

and that the fraction of the '00' state, f_{00}, for example is given by

$$f_{00} = \frac{[00]}{[00]+[01]+[10]+[11]} = \frac{K_aK_c}{K_aK_c + (K_c + K_d)[H^+] + [H^+]^2} \qquad \text{[S2]}$$

with similar expressions for f_{01}, f_{10}, and f_{11}.

In the limit where deprotonation/protonation is fast on the NMR chemical shift timescale, as is most often the case[6], the chemical shift of the titrating NMR probe can be written as,

$$\delta_{obs} = \sum_i f_i\delta_i \quad i \in \{00,01,10,11\}$$

$$= \frac{\delta_{11}[H^+]^2 + (\delta_{10}K_c + \delta_{01}K_d)[H^+] + \delta_{00}K_aK_c}{K_aK_c + (K_d + K_c)[H^+] + [H^+]^2}, \qquad \text{[S3]}$$

that reduces to

$$\delta_{obs} = \frac{\delta_{11}[H^+]^2 + (\delta_{10}K_{A,2} + \delta_{01}K_{KA,1})[H^+] + \delta_{00}K_{A,1}K_{A,2}}{K_{A,1}K_{A,2} + (K_{A,1} + K_{A,2})[H^+] + [H^+]^2} \qquad [S4]$$

when $K_a = K_d = K_{A,1}$ and $K_b = K_c = K_{A,2}$ (*i.e.*, binding is not cooperative). It is straightforward to show that Eq [S4] can always fit data derived from a 2 site titrating system, even when $K_a \neq K_d$ (*i.e.*, $K_a = \alpha K_d$) and $K_b \neq K_c$ (*i.e.*, $K_b = \alpha K_c$), so long as $\alpha \leq 1$. This can be seen by comparing Eqs [S3] and [S4] and requiring that

$K_aK_c = K_{A,1}K_{A,2}$

$K_d+K_c = K_{A,1}+K_{A,2}$ [S5]

from which it follows that

$$K_{A,1} = \frac{K_d + K_c + \sqrt{(K_d + K_c)^2 - 4\alpha K_dK_c}}{2},\ K_{A,2} = \frac{K_d + K_c - \sqrt{(K_d + K_c)^2 - 4\alpha K_dK_c}}{2}\ . \qquad [S6]$$

Note that for $\alpha < 1$ proton dissociation is less efficient when the coupled second site is already deprotonated, as is almost certainly the situation in any real system. In this case it follows that $(K_d + K_c)^2 - 4\alpha K_dK_c \geq 0$ and a solution (real values) for K_{A1}, K_{A2} can be found. In the general case, titration profiles from a system comprised of a pair of titrating groups can always be fit to extract a pair of (macroscopic) dissociation constants, as well as three chemical shifts[6].

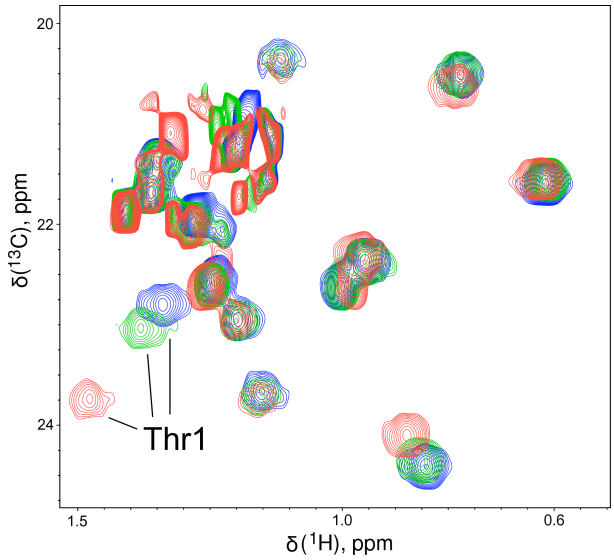
For fitted values of K_{A1}, K_{A2} it can be shown that

$$K_{d,c} = \frac{(K_{A,1} + K_{A,2}) \pm \sqrt{(K_{A,1} + K_{A,2})^2 - \frac{4K_{A,1}K_{A,2}}{\alpha}}}{2} \qquad [S7]$$

so that real number solutions are found for $\alpha \geq \alpha_{min}$, where $\alpha_{\min} = \frac{4K_{A,1}K_{A,2}}{(K_{A,1} + K_{A,2})^2}$. Figure S3 plots values of $K_{d,c}$ and $K_{a,b}=\alpha K_{d,c}$ as a function of $-\log(\alpha)$ showing how the

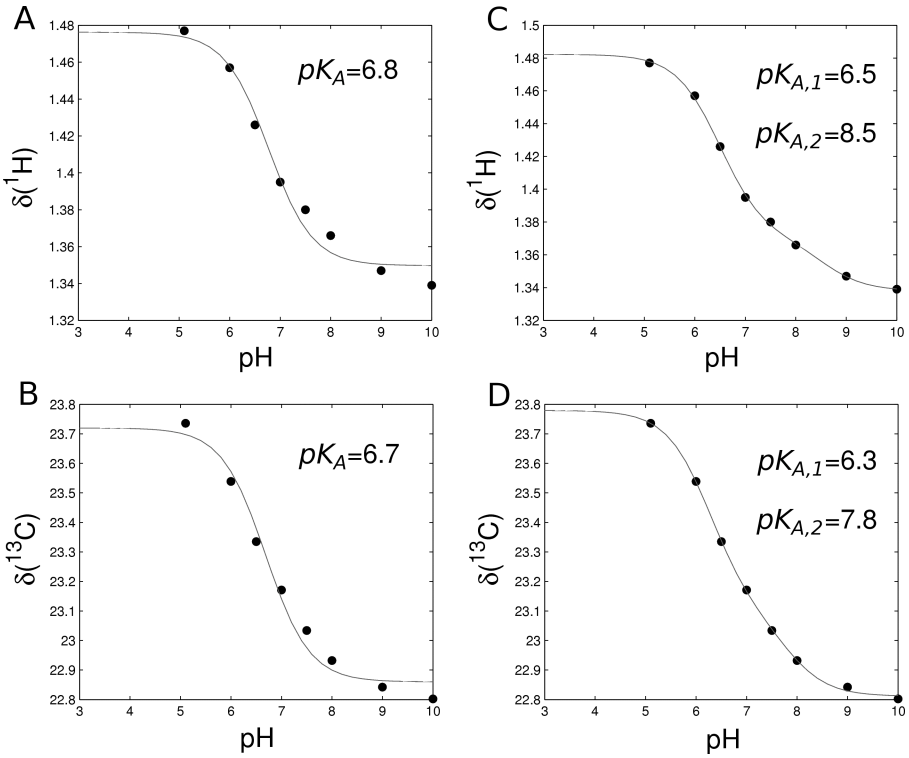
microscopic dissociation constants in a coupled titrating system can vary as a function of α. Values of $pK_{A,1}$ = 6.3 and $pK_{A,2}$ = 7.7, obtained from fits of titration data assuming a non-cooperative system, have been used. Of particular interest is that abstraction of the first proton from a fully protonated system ('11'), corresponding to deprotonation of the terminal amine of Thr1 (pK_d), is only very modestly affected by the extent of cooperativity (α). The overall picture that the Thr1 amine is largely deprotonated at neutral pH is thus invariant to the fitting model used.

Supporting Figure 1. Superposition of ^{13}C-^{1}H HMQC correlation maps of the 20S CP focusing on the Thr region recorded at pH 5 (red), 7.5 (green), and 10 (blue), 18.8T, 70 ^{o}C.

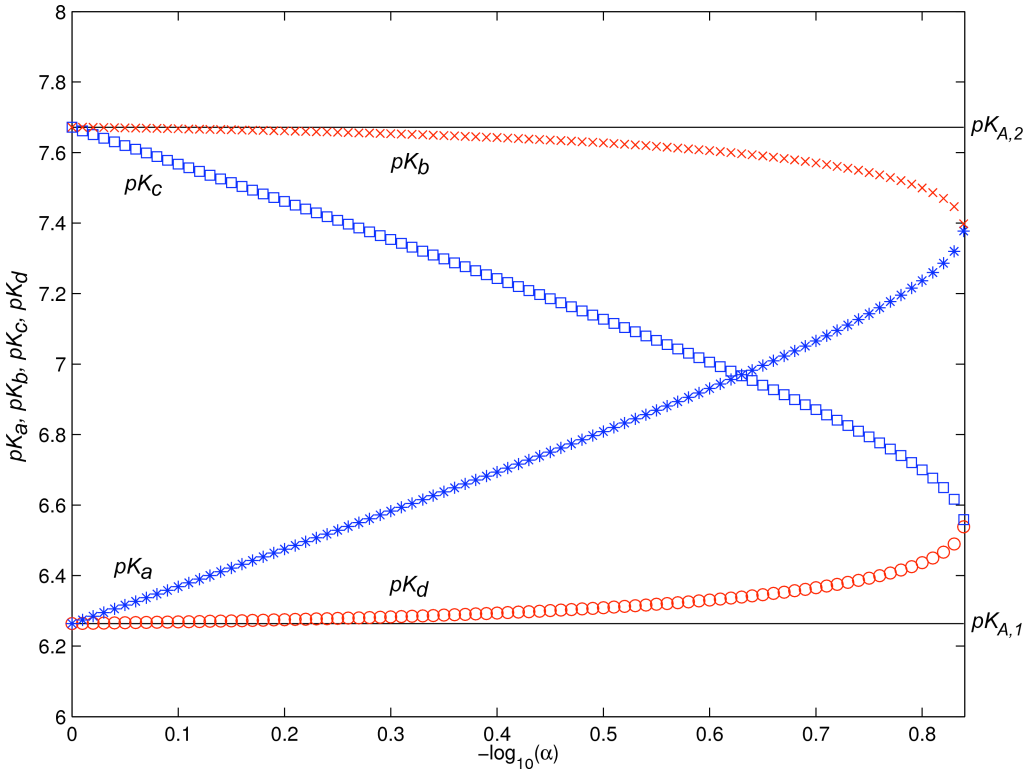


Supporting Figure 2. Fits of titration data from Thr1 assuming a single titrating site. Best fits (solid lines) of experimental data (circles) are indicated for both ^{1}H (A) and

^{13}C (B) chemical shifts. The quality of data fitting is improved significantly if a more complex model involving two titratable moieties is assumed, such as in (C,D). In (C,D), ^{1}H and ^{13}C titration data are fitted separately, with similar pairs of $pK_{A,1}$ and $pK_{A,2}$ values obtained from fits of both nuclei.



Supporting Figure 3. Values of pK_a (blue, *), pK_b (red, x), pK_c (blue, box) and pK_d (red, o) as a function of –log(α). Values of $pK_{A,1}$ and $pK_{A,2}$ used in the calculation are as indicated. Details are as described above.



Supporting references:

(1) Ruschak, A. M.; Religa, T. L.; Breuer, S.; Witt, S.; Kay, L. E. Nature 2010, 467, 868–71.

(2) Ruschak, A. M.; Kay, L. E. Proc Natl Acad Sci U S A 2012, 109, E3454–62.

(3) Velyvis, A.; Ruschak, A. M.; Kay, L. E. PLoS One 2012, 7, e43725.

(4) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. J Biomol NMR 1995, 6, 277–93.

(5) Johnson, B. A.; Blevins, R. A. J Biomol NMR 1994, 4, 603–14.

(6) McIntosh, L. P.; Naito, D.; Baturin, S. J.; Okon, M.; Joshi, M. D.; Nielsen, J. E. J Biomol NMR 2011, 51, 5–19.