

UCSF

UC San Francisco Previously Published Works

Title

The $Y\Phi$ motif defines the structure-activity relationships of human 20S proteasome activators

Permalink

<https://escholarship.org/uc/item/08t4k2m8>

Journal

Nature Communications, 13(1)

ISSN

2041-1723

Authors

Opoku-Nsiah, Kwadwo A
de la Pena, Andres H
Williams, Sarah K
et al.

Publication Date






2022

DOI

10.1038/s41467-022-28864-x

Peer reviewed

The $Y\Phi$ motif defines the structure-activity relationships of human 20S proteasome activators

Kwadwo A. Opoku-Nsiah ^{1,2}, Andres H. de la Pena³, Sarah K. Williams ^{1,2}, Nikita Chopra^{1,4,5}, Andrej Sali ^{1,4,5}, Gabriel C. Lander ³ & Jason E. Gestwicki ^{1,2}✉

The 20S proteasome (20S) facilitates turnover of most eukaryotic proteins. Substrate entry into the 20S first requires opening of gating loops through binding of HbYX motifs that are present at the C-termini of certain proteasome activators (PAs). The HbYX motif has been predominantly characterized in the archaeal 20S, whereas little is known about the sequence preferences of the human 20S (*h20S*). Here, we synthesize and screen ~120 HbYX-like peptides, revealing unexpected differences from the archaeal system and defining the *h20S* recognition sequence as the Y-F/Y ($Y\Phi$) motif. To gain further insight, we create a functional chimera of the optimized sequence, NLSYYT, fused to the model activator, PA26^{E102A}. A cryo-EM structure of PA26^{E102A}-*h20S* is used to identify key interactions, including non-canonical contacts and gate-opening mechanisms. Finally, we demonstrate that the $Y\Phi$ sequence preferences are tuned by valency, allowing multivalent PAs to sample greater sequence space. These results expand the model for termini-mediated gating and provide a template for the design of *h20S* activators.

¹Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA, USA. ²Institute for Neurodegenerative Disease, University of California San Francisco, San Francisco, CA, USA. ³Department of Integrative Structural and Computational Biology, Scripps Research Institute, La Jolla, CA, USA. ⁴Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA. ⁵Quantitative Biosciences Institute, University of California San Francisco, San Francisco, CA, USA. ✉email: jason.gestwicki@ucsf.edu

The proteasome is a critical regulator of protein homeostasis that degrades ~90% of all eukaryotic proteins¹. The enzymatic activity of this system is carried out by the 20S proteasome (20S), a cylindrical, complex composed of four stacked rings enclosing an axial channel. To be degraded, potential substrates must first diffuse through a pore at the center of the distal α -rings before encountering the peptidase sites within the inner β -rings. N-terminal extensions of the α -subunits gate entry into the pore, limiting the degradation of bystander proteins². In turn, this barrier creates a key regulatory role for the proteasome activators (PAs)^{3–5}, large particles that bind the 20S, open the gates, and facilitate substrate selection and entry. One evolutionarily conserved^{6–8} way that PAs achieve this goal is by using a tripeptide motif at their extreme C-termini, which is characterized by a hydrophobic amino acid, followed by a tyrosine and then any amino acid (HbYX). HbYX motifs open the gates by docking into pockets located between adjacent α -subunits of the 20S (termed α -pockets) (Fig. 1a)^{9–11}. Unlocking the details of HbYX recognition at this protein-protein interaction (PPI) will deepen our understanding of the gate opening mechanism and enable novel strategies to regulate protein degradation in cells.

The HbYX model for gate opening was pioneered from studies by Goldberg and colleagues using the C-terminal sequence of the archaeal PA, proteasome-activating nucleotidase (PAN)⁹. This work, and subsequent structures, revealed key roles for the penultimate Tyr residue and the terminal carboxylate of the HbYX motif^{12,13}. However, unlike the archeal system, the human 20S proteasome (*h20S*) is a hetero-oligomer, containing seven distinct α -pockets for binding to the HbYX motif. Thus, it is not clear whether the same structure-activity relationships (SAR) that govern the recognition of the HbYX motif in the archaeal system are conserved in the *h20S*. Addressing this question in the context of natural PAs has been challenging. For example, the structural topology of PA700/19 S, the eukaryotic homolog to PAN, means that its six distinct C-termini are not allowed to sample different α -pockets¹⁴. Moreover, there are confounding effects of cooperativity, ATP hydrolysis^{15–17}, and contributions from the additional subunits of PA700^{18–20}. We envisioned that one way to circumvent these issues and reveal the underlying SAR might be to use synthetic peptides instead of native PAs. Indeed, HbYX-containing peptides derived from the C-termini of the human Rpt subunits of PA700, have been shown to bind the α -pockets and stimulate turnover of substrates by the 20S^{9,21,22}. Inspired by that approach, we hypothesized that a library of peptides could be used to understand the SAR of the HbYX motif in the *h20S*. Importantly, beyond the important insights into the molecular mechanisms of *h20S* gate opening, such investigations might be expected to inform the design of small molecules that bind the α -pockets²³.

In this study, we design and synthesize ~120 peptides derived from Rpt5's C-terminus and evaluate their ability to stimulate the peptidase activity of the *h20S* in vitro. This analysis reveals sequence preferences that differ from the canonical HbYX motif as derived in archaea. We refer to this re-defined preference as the Y Φ motif. To better understand the structural underpinnings of the Y Φ motif, we graft an optimal sequence, NLSYYT, to the C-termini of an inert PA platform and solve a 2.9 Å resolution structure of the PA bound to the *h20S* by cryo-electron microscopy (cryo-EM). Remarkably, the same orientation of NLSYYT is observed in five of the seven α -pockets, suggesting a conserved set of molecular interactions. Analysis of the bound Y Φ motif reveals specific inter- and intramolecular contacts, which are involved in molecular recognition and gate opening. Finally, using a series of chimeras, we find that the valence of the PA displaying the Y Φ motif (e.g. monomer vs. heptamer) tunes these

sequence preferences, with multivalent PAs able to overcome otherwise non-ideal sequences. Together, these studies reveal mechanisms of termini-dependent gate opening in the *h20S* and establish a consensus sequence for monovalent activators.

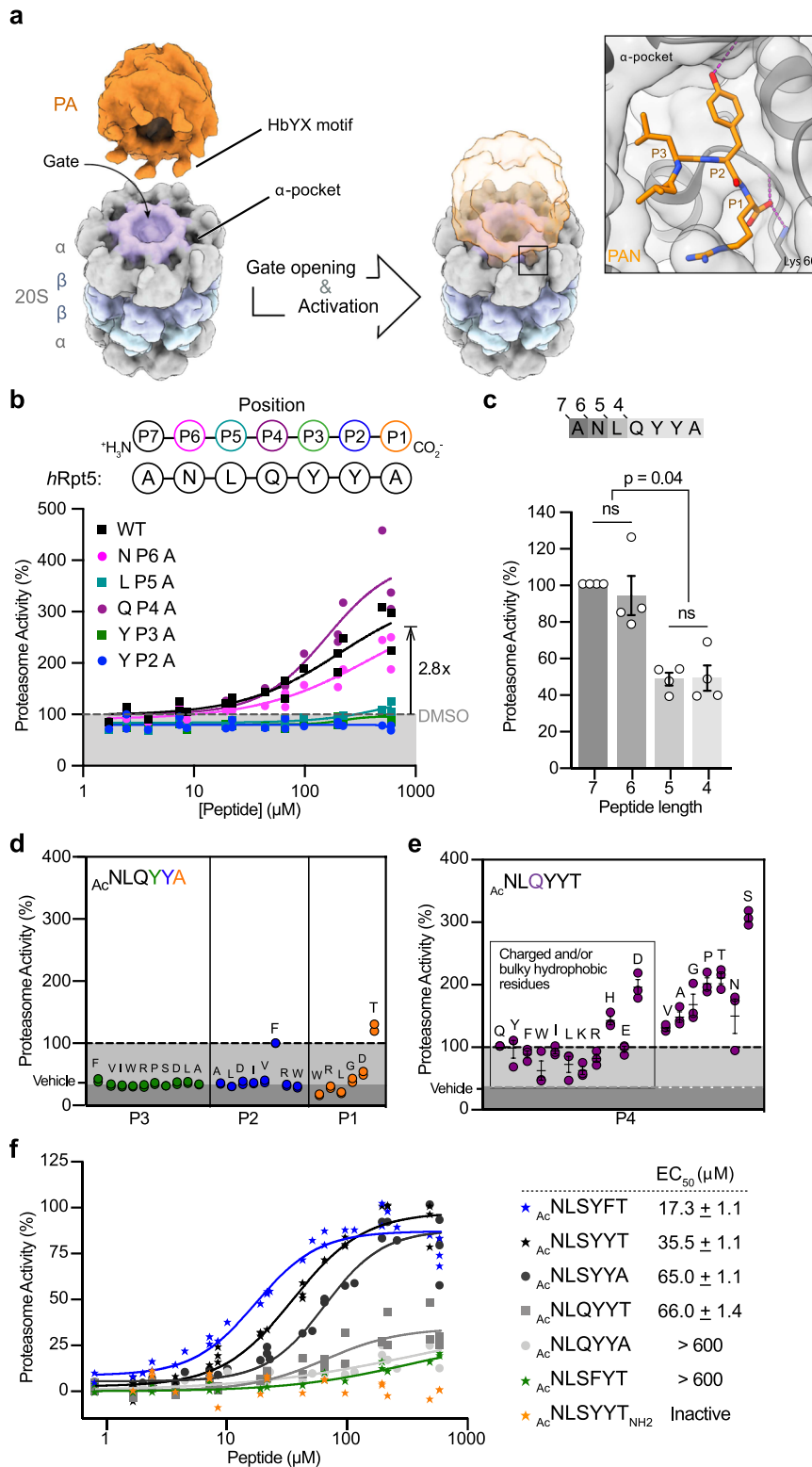
Results

Nomenclature of the HbYX motif and α -pockets. Throughout this work, we will refer to the carboxy-terminal residue of the HbYX motif as termed P1, the next residue P2, etc. In this parlance, the canonical HbYX model is defined as having a hydrophobic residue at P3, a preference for Tyr at P2, and any residue at P1. Another key part of the HbYX motif is that it contains the terminal carboxylate at the P1 position, which forms a critical salt bridge with a conserved cationic side chain at the base of the α -pocket; for example α Lys66 of the *Thermoplasma acidophilum* 20S (Fig. 1a)⁹. Unless otherwise noted, we will use the residue numbering of the *T. acidophilum* 20S.

Establishing a pipeline for measuring peptide-mediated activation of the *h20S*. To determine the SAR of the HbYX motif in the *h20S* system, we synthesized, characterized, and assayed a library of peptides derived from the C-terminus of human Rpt5 with the native sequence: ANLQYYA. For each peptide, its capacity to accelerate substrate turnover by purified *h20S* was monitored by the hydrolysis of a fluorogenic substrate (suc-LLVY-amc). Historically, peptide-based activators have been shown to exhibit atypical dose responses, including partial inhibition of the 20S at higher concentrations²⁴. Similarly, sodium dodecyl sulfate (SDS) and other detergent-like small molecules have been reported to activate the 20S nonspecifically; these otherwise useful solubilizing agent are typically excluded from buffers in proteasome activity assays²⁵.

With these considerations in mind, we first optimized the buffer conditions and establish a triage pipeline to remove non-specific activators. These efforts were inspired by previous reports to identify buffers in which the difference between the 20S activities in the basal and stimulated conditions are maximized^{5,26,27}. Using this starting condition (see Methods), we observed that high concentrations of either peptide or known small-molecule activators would inhibit, rather than stimulate, substrate turnover by the *h20S*. We reasoned that this effect was likely due to aggregation at the higher concentrations²⁸. Supplementing the assay buffer with a non-ionic surfactant (0.01% Pluronic F-68[®]) attenuated the loss in dose-responsiveness caused by high concentrations of activators (Supplementary Fig. 1a–c). Yet, we also identified peptides displaying atypical dose responses that could not be resolved with the solubilizing agent (e.g. Cys- and Met-containing peptides) (Supplementary Fig. 1d) and these peptides were therefore excluded from further study. Peptides that yielded visibly cloudy solutions at 10 mM in DMSO were also omitted (Supplementary Fig. 1e). Finally, the remaining peptides were evaluated for solubility by dynamic light scattering (DLS) (see Methods) (Supplementary Table 1). Together, these triage steps restricted the sequence space that could be studied, but also minimized contributions from non-specific mechanisms.

Structure-activity relationships (SAR) of *hRpt5* peptides deviate from the HbYX model. In the first series of peptides, we performed an alanine mutational scan of *hRpt5*: ANLQYYA. Consistent with previous reports²¹, the parent peptide (termed wildtype or WT) stimulated turnover of suc-LLVY-amc by ~3 \times (Fig. 1b). An EC₅₀ value could not be determined, due to relatively weak potency, so, at this stage, we compared peptides based on their relative ability to stimulate hydrolysis rate. Replacing



either the P3 ('Hb') or P2 ('Y') positions with Ala completely ablated activity, in agreement with the HbYX model⁹. Surprisingly, replacing the P4, P5 or P6 residues with Ala also had appreciable effects, including an unexpected enhancement in activity by Ala substitution at P4 (1.3× over WT), complete ablation of activity at P5, and modest impairment at P6 (~80% of WT) (Fig. 1b). A series of N-terminal truncations revealed that the 6- and 7-mer peptides were significantly more stimulatory

than shorter ones (Fig. 1c), supporting a role for residues upstream of the tripeptide motif⁹.

Guided by these findings, we designed a series of hexapeptides to probe the SAR in more detail. Peptides were synthesized and N-terminally acetylated (Ac), which generally enhanced activity over the free amine (Supplementary Fig. 2a). Acetylated peptides were then assessed at a single dose (250 μ M) against the h20S to obtain an initial overview of their relative activities. In the first set

Fig. 1 Activation of *h20S* by *hRpt5*-derived peptides reveals unexpected SAR. **a** Schematic of proteasome activation by the C-terminal HbYX motif of a model PA. HbYX motifs are known to dock into α -pockets located between subunits of the 20 S α -rings (gray). This interaction opens the gates of the 20 S, allowing substrate entry. The inset summarizes the interactions between PAN's HbYX motif (orange) and the α -pocket (gray surface/cartoon) of the archaeal 20S (PDB ID: 3IPM). **b** Alanine scanning revealed residues of Rpt5 C-terminus that make key contributions to the stimulation of the *h20S*. Activation of the *h20S* (4 nM) was measured by an increase in the hydrolysis rate (RFU/s) of the fluorescent substrate, suc-LLVY-amc (10 μ M) upon addition of *hRpt5*-derived peptides. Data are normalized to DMSO-treated *h20S* and plotted individually ($n = 3$). **c** Proteasome stimulation by peptides (250 μ M) from an N-terminal truncation series. Results are normalized to ANLQYYA and the average of four independent experiments (open circles; $n = 4$) is plotted with error reported as s.e.m. *P*-values were calculated using two-tailed unpaired *t*-test: * $p < 0.04$. **d** Scatterplot of the relative activities of N-terminally acetylated *hRpt5* hexapeptides (250 μ M) sampling modifications along the P3 (green), P2 (blue), and P1 (orange) residues. Data are normalized to Ac-NLQYYA and plotted individually ($n = 2$). **e** Scatterplot of maximal activity of hexapeptides sampling all proteinogenic amino acids at the P4 position (purple). Data are normalized to Ac-NLQYYT and the average of three independent experiments (circles; $n = 3$) is plotted with error reported as s.e.m. **f** Combination of optimized residues (underlined) from the SAR campaign. Data are normalized to Opt5 (Ac-NLSYYT, star) and plotted individually ($n = 3$ or 4). EC₅₀ values were calculated from four independent experiments and reported as mean \pm s.e.m.

of comparisons, we varied the P1, P2, and P3 positions, replacing the wild-type residue with chemically and/or structurally diverse amino acids. We found that the *h20S* had sequence preferences at the P1 position, with Thr favored over other residues that were tested. More importantly, Tyr seemed to be exclusively required at the P3 position; such that even other hydrophobic residues, including Leu and Ile, were not tolerated. Lastly, either Phe or Tyr was preferred at the P2 position (Fig. 1d). When we compared these sequence preferences to those reported for the archaeal 20S, we found the *h20S* had both similarities and significant differences (Supplementary Fig. 3)⁹.

We were particularly intrigued by the gain in activity caused by Ala at P4 (see Fig. 1b), so we generated a focused series of P4-substituted hexapeptides to explore this position further. In this collection, the P1 was uniformly replaced with Thr, which we found to improve aqueous solubility and activate better than Ala, allowing us to obtain saturable stimulation curves and calculate EC₅₀ values (Supplementary Fig. 4a). We found that none of the P4 substitutions completely ablated activity, suggesting that, unlike P2 or P3, the requirements at the P4 residue are more permissive. Nonetheless, modifications at P4 impart the greatest modulatory effect on the rate of hydrolysis (up to 3 \times over WT), with positively charged or bulky side chains generally exhibiting lesser stimulation (Fig. 1e). Additionally, residues known to disrupt α -helical character (e.g. Gly and Pro) were well-tolerated at the P4 position, consistent with previous observations²⁹. Lastly, we explored the contributions of the P5 and P6 position. Even though we had previously found that truncations of these residues diminished activity, residues substitutions at P5 or P6 had modest effects (see Supplementary Fig. 2b–d), suggesting that the side chain identity was less important in these positions. Next, we incorporated the optimal substitutions at both P1 and P4 to probe whether they are synergistic. In these experiments, we were particularly interested in understanding which substitutions might enhance the apparent EC₅₀ (as a pseudo approximation of affinity) and which ones might impact the rate of hydrolysis. The results showed that Thr at P1 improved only the EC₅₀, while Ser at P4 enhanced both EC₅₀ and the rate of hydrolysis. Combining both substitutions had an additive effect (Fig. 1f), yielding an optimized sequence Ac-NLSYYT (EC₅₀ = 35.5 \pm 1.1 μ M). For brevity, we refer to this sequence as Opt5.

Based on our peptide study, the penultimate tyrosine residue is not essential for activating the *h20S*. Previous work had demonstrated that certain HbYX-containing PAs can function with a Phe or Tyr (F/Y—denoted as Φ) at P2³⁰, however, the predicted importance of the Tyr at the P3 position seemed unique. To test this idea, we substituted either the P2 or P3 Tyr residues of Opt5 (referred to here as Opt5^{YY} for clarity) with Phe (Opt5^{YF} or Opt5^{FY}, respectively) and measured the ability of these peptides to stimulate *h20S*. Removal of the P3 hydroxyl (Opt5^{FY}) dramatically reduced activity by \sim 70%, while loss of the

P2 hydroxyl (Opt5^{YF}) had no appreciable effect on turnover and actually improved potency by \sim 2-fold relative to Opt5^{YY} (EC₅₀ = 17.3 \pm 1.1 μ M) (Fig. 1f). Previous work had also suggested that a Phe at P2 might be favored³⁰, so we assessed the generality of this finding by testing two additional peptide sequences (Ac-NLSY Φ A and Ac-NLGY Φ T), noting up to a 3-fold improvement in potency by Phe over Tyr in both cases (Supplementary Fig. 4b).

To this point, measurements of proteasome activity were restricted to the suc-LLVY-amc probe, which measures chymotryptic-like activity. If the optimized peptide Opt5 induces gate opening, we would expect that it would also stimulate the tryptic-like activity. Indeed, we found that treatment with Opt5 and other peptides promoted the tryptic-like activity of the *h20S*, as measured using boc-LRR-amc (Supplementary Fig. 5a). Moreover, these peptides also stimulated hydrolysis of a longer, nonapeptide substrate (FAM-LFP), which is known to require gate opening for its entry into the proteasome³¹ (Supplementary Fig. 5b). Taken together, these findings suggest that *h20S* activation by *hRpt5*-like peptides occurs through interactions with at least the last four residues, with additional contributions from P5 and P6. The sequence requirements are expanded from the HbYX motif and are summarized as the Y Φ model (see below).

PA26^{E102A}-Opt5 induces terminus-dependent gate opening of the *h20S*. To understand the molecular basis for the sequence preferences in the Y Φ motif, we attempted to determine the structure of *h20S* bound to Opt5 peptides by cryo-EM but were unable to obtain high-resolution structures. Rather, inspired by the previous studies^{8,13,30,32}, we sought to use the homoheptameric activator PA26 as a scaffold for the display of multiple copies of Opt5 (NLSYYT). Instead of C-terminal HbYX motifs, PA26 has an activation loop that displaces the N-terminal gating residues through direct contacts with the reverse-turn loop at α Pro17 in the α -ring (Fig. 2a)^{3,4}. An Ala substitution in the activation loop (E102A) renders PA26 inactive⁸. Previous studies showed that the capacity of the disabled-loop mutant (PA26^{E102A}) to induce gate opening was rescued by grafting PAN's HbYX motif in place of the native C-termini^{13,30}. Thus, we envisioned likewise using PA26^{E102A} to enable high-resolution structural studies of termini-dependent gate opening.

Archaeal PA26 is capable of activating the *h20S*³³, however, it remains unclear whether the E102A mutation impairs this function, as has been reported for the archaeal 20S. Using the *T. brucei* PA26, we verified that PA26^{E102A} was unable to stimulate the *h20S*. Moreover, PA26^{E102A}'s ability to bind *h20S* was significantly diminished, as measured by biolayer interferometry (BLI) (Fig. 2b). Next, we genetically installed NLSYYT in place of the last six native residues of PA26 (PA26^{E102A}-Opt5)

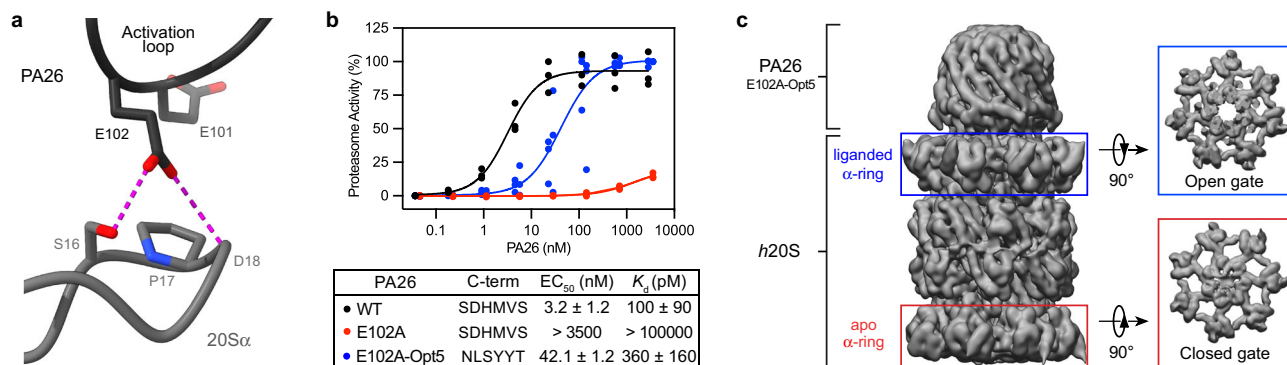


Fig. 2 PA26^{E102A-Opt5} induces terminus-dependent activation and gate opening of the *h20S*. **a** Schematic of how PA26's activation loop (black cartoon) directly repositions the Pro17 reverse turn of the 20S α -subunit (gray cartoon) to induce HbYX-independent gate opening (PDB ID: 1YA7). **b** Stimulation of *h20S* by PA26 constructs: wild-type PA26 (wt, black), disabled-loop mutant (PA26^{E102A}, red), and a disabled-loop mutant with C-terminal Opt5 (NLSYYT) sequence (PA26^{E102A-Opt5}, blue). Data are normalized to PA26^{E102A-Opt5} and plotted individually ($n = 3$). Reported EC₅₀ is a mean of EC₅₀ values calculated from three independent experiments ($n = 3$) with error reported as s.e.m. K_D values were determined by BLI from three independent experiments ($n = 3$) and reported as the mean \pm s.e.m. **c** PA26^{E102A-Opt5} induces gate opening. The 3D reconstruction of the singly capped PA26^{E102A-Opt5}-*h20S* complex at 2.9 Å resolution from 234,960 particles shows gate opening (blue box) relative to the apo, closed gate α -ring on the opposite side (red box).

to generate a PA that now bound the *h20S* ($K_D = 360 \pm 160$ pM) and potently stimulated its activity ($EC_{50} = 42.1 \pm 1.2$ nM) via its C-terminus (Fig. 2b). To probe how the binding of Opt5 activates *h20S*, we then determined the structure of the *h20S*-PA26^{E102A-Opt5} complex by cryo-EM to an overall resolution of 2.9 Å (Supplementary Figs. 6 & 7 and Supplementary Table 2). The cryo-EM structure resolves a single particle of the PA26^{E102A-Opt5} heptamer bound to an α -ring of the *h20S*. At this site, the N-terminal extensions of the α -subunits were displaced from the central pore (Fig. 2c) and the diameter of the pore was widened by 3.8 Å relative to the apo α -ring (Supplementary Table 3), reminiscent of other open gate structures⁸. Consistent with findings from an analogous study with an archaeal 20S, binding of PA26^{E102A-Opt5} at one α -ring of the *h20S* did not allosterically open the distal α -ring gate³².

Cryo-EM resolves open gate conformation of the *h20S*. Using this structure, we explored the potential mechanisms contributing to the opening of the *h20S* gates. In other systems, the open gate conformation is regulated by a cluster of conserved N-terminal residues (α Tyr8, α Asp9, α Pro17, and α Tyr26)^{7,8}. Specifically, these residues are repositioned away from the central pore during gate opening and are anchored in that position by a characteristic set of intra- and intermolecular contacts. However, these residues are less conserved in humans, particularly in $\alpha 1$ and $\alpha 2$ (α -subunits are labeled to match the numbering used in yeast 20S), where the canonical α Tyr8 and α Asp9 are replaced with $\alpha 1$ Phe9 and $\alpha 2$ Ser7, respectively. We wondered whether these differences might impact either the gating mechanism or the extent of opening. In the liganded α -ring of our structure, the N-termini of $\alpha 5$, $\alpha 6$, and $\alpha 7$ formed the expected, ordered clusters at the α -subunit interface, including a repositioned α Pro17 within CH- π distance from α Tyr26 and α Tyr8, consistent with the fully open state^{19,20}. However, the equivalent residues in $\alpha 1$ and $\alpha 2$ formed clusters with fewer contacts likely resulting from displaced polar contacts in the noncanonical $\alpha 1$ Phe9 and $\alpha 2$ Ser7 residues, respectively (Supplementary Fig. 8a). In addition, the N-terminal residues were poorly resolved in $\alpha 1$ and $\alpha 2$ relative to N-termini from the ordered clusters. Interestingly, we noted that although subunits neighboring $\alpha 1$ and $\alpha 2$, namely $\alpha 3$ and $\alpha 4$, have canonical N-terminal sequences, these conserved residues formed limited contacts within clusters and/or had side chains that could

not be structurally resolved. Furthermore, $\alpha 2$, $\alpha 3$, and $\alpha 4$, the subunits that primarily seal the pore's entrance of the closed gate apo *h20S*, displayed disordered N-terminal extensions that seem to partially occlude the pore, unlike the ordered and fully displaced N-termini of reported open states (Supplementary Fig. 8b). Such conformations have uniquely been observed in human 26 S proteasome structures^{16,17}, however, it was not entirely clear if the 19S's C-termini were incapable of fully opening the gate. These observations support previous claims³⁰ and further demonstrate how noncanonical residues in the human α -subunits limit gate opening by displacing stabilizing interactions, which propagates disorder along the α -ring to, in turn, destabilize a fully opened gate state in the *h20S*.

To discern the conformational changes associated with gate opening, we aligned the β -rings of our structure to that of apo *h20S* (PDB ID: 4R3O)³⁴ because the β -subunits remain relatively unchanged during activation^{4,13}. We observed that the Ca atoms of all seven α Pro17 residues were radially displaced between 1.0 and 3.7 Å during gate opening. Consistent with the analyses of the N-terminal gate positions, the partially opened $\alpha 1$ and $\alpha 2$ N-termini had the smallest shifts in α Pro17 (1.0 and 1.3 Å, respectively) (Supplementary Fig. 8c). The modest displacement of $\alpha 2$ Pro15 contrasts to the relatively large shifts typically observed by the N-terminus of the yeast $\alpha 2$ subunit³⁰. This difference between yeast and human is interesting to note because both species share a noncanonical Ser in $\alpha 2$; thus, it is possible that replacement of α Tyr8 with $\alpha 1$ Phe9, which is unique to *h20S*, is more restrictive of the fully open conformation.

We next examined how the termini might be inducing gate opening in our structure. First, we confirmed that the disabled activation loop of PA26^{E102A-Opt5} does not interact with α Pro17, suggesting that this structure allows the determination of terminus-specific mechanisms, as intended. At the occupied α -pockets, most of the P2 and P3 Tyr side chains of Opt5 made H-bond contacts with residues located in the reverse turn between α Pro17 and helix 0 (Supplementary Fig. 9a). This finding suggests that residues of Opt5 act directly on the reverse-turn loop to open the gate. We also detected that α -subunits were rotated an average of $1.5 \pm 0.4^\circ$ about the axial channel (Supplementary Fig. 9b) and observed significant side chain rearrangements throughout the α -ring (Fig. 3a), which were induced by PA26^{E102A-Opt5}. Indeed, both rigid body rotations and induced-fit conformational changes have been attributed to terminus-dependent gate opening^{12,13,17,32}.

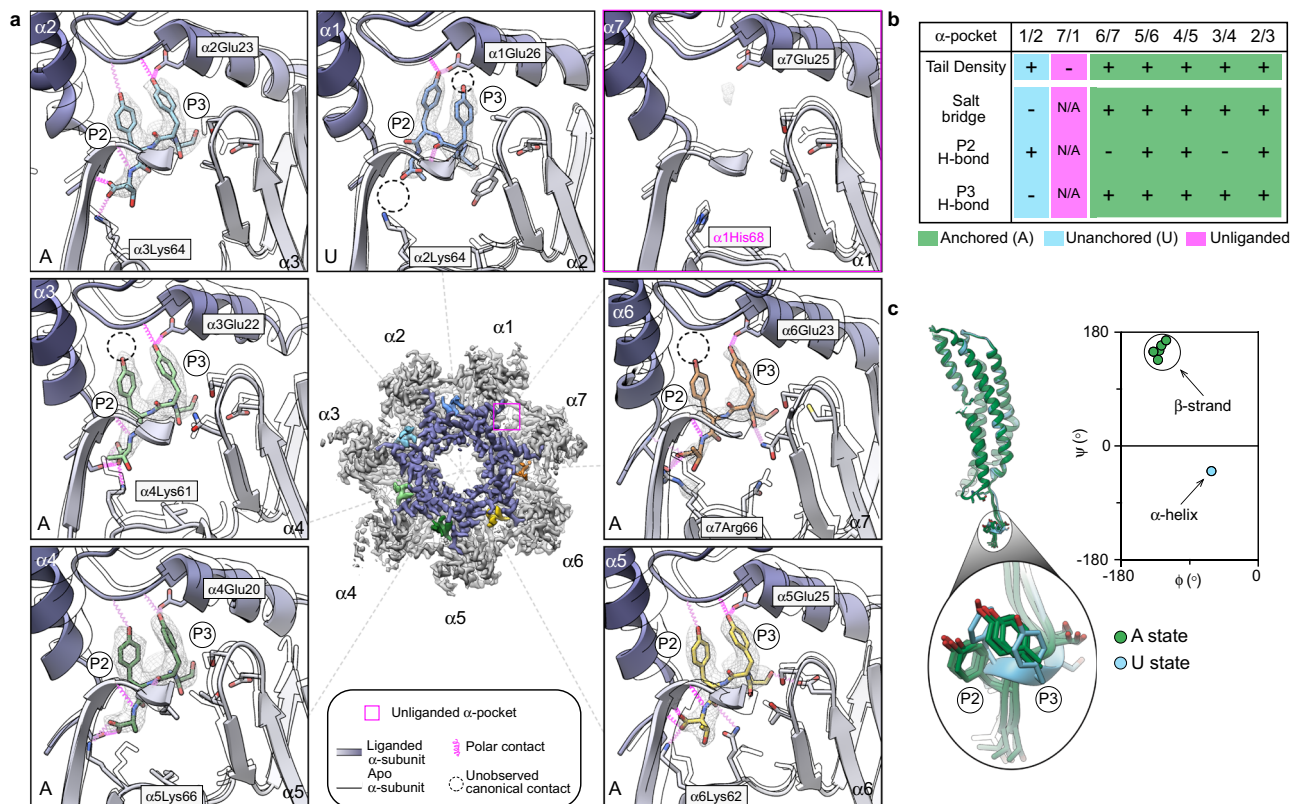


Fig. 3 Ensemble of C-terminal interactions promote gate opening in the *h20S*. **a** Opt5 sequence in the *h20S* α -pockets. The α -subunits are numbered according to the yeast proteasome. *Center*, a top-view 3D density map of the liganded α -ring displaying color-coded densities for each C-tail. The empty α -pocket ($\alpha 7/\alpha 1$) is marked by a red dotted circle. Close-up views of each C-terminal Opt5 sequence (stick) is superimposed over their corresponding cryo-EM densities (mesh). PA26^{E102A}-Opt5 induces terminus-dependent conformational changes throughout the inner (purple) and outer (gray) regions of the α -pockets relative to the apo *h20S* (outlined) (PDB ID: 4R30). Predicted interactions (pink coils) and unobserved canonical interactions (black dotted circles) are denoted. Anchored (A) or unanchored (U) Opt5 binding states are labeled at the bottom-left corner of each panel (see text). **b** Summary table of the interactions for each α -pocket. **c** Opt5 binds with distinct conformations in the A and U states. Opt5 forms an α -helix in the U state ($\alpha 1/\alpha 2$ -pocket), while it adopts a β -strand conformation in the A state. Overlaying the C-tails highlights the distinct backbone conformations between the A and U states, while also showing the striking similarities between the A states, especially in the position of the P2 and P3 Tyr side chains. See Supplementary Table 2 for relative densities of the C-tails (**b**) and torsion angles (**c**).

Overall, we conclude that termini-mediated activation of the *h20S* largely proceeds through the conserved gating mechanism observed in other proteasomes; but with distinctions in the asymmetric gate opening of the α -rings, potentially linked to the noncanonical $\alpha 1$ Phe9.

Opt5-bound α -pockets reveal key contacts for terminus-dependent *h20S* activation.

We next examined the direct interactions of Opt5 in the *h20S* α -pockets. This analysis was aided by the matched stoichiometry of C-termini to α -pockets, together with the fact that the α -pockets were locally resolved to better than 2.9 Å resolution (see Supplementary Fig. 7d). Six of the seven α -pockets contained well-defined density corresponding to the termini of PA26^{E102A}-Opt5. As previously predicted^{8,20,35}, the $\alpha 7/\alpha 1$ -pocket (interfacing the $\alpha 7$ and $\alpha 1$ subunits) did not contain density for Opt5, likely because this pocket does not have a canonical Lys residue ($\alpha 1$ His68). Of the remaining six sites, the termini in the $\alpha 1/\alpha 2$ -pocket had a noticeably distinct structure. Specifically, its terminal carboxylate was displaced (5.1 Å) from the ϵ -amino group of $\alpha 1$ Lys63 and therefore unable to form the expected salt bridge, which has previously been shown to be critical for binding of native PAs and HbYX peptides^{9,12}. To verify the importance of the negatively charged terminal carboxylate to overall *h20S* stimulation, we amidated the C-terminus of Opt5 peptide and confirmed that it lost all activity in turnover

assays (see Fig. 1f). Hence, we categorized the C-termini in the $\alpha 2/\alpha 3$, $\alpha 3/\alpha 4$, $\alpha 4/\alpha 5$, $\alpha 5/\alpha 6$ and $\alpha 6/\alpha 7$ pockets as binding in an anchored (A) state and the C-termini in the $\alpha 1/\alpha 2$ -pocket as binding in the unanchored (U) state (Fig. 3a, b).

The peptide backbone of Opt5, of both A and U states, were well-resolved for the P1, P2, P3, and P4 positions, enabling further assessment. First, we noticed that the backbone conformations of the bound C-termini are different between the A and U states. The dihedral angles of the P2-P3 residues of the bound Opt5 form a β -strand structure for the A state but a right-handed α -helix for the U state (Fig. 3c). In a recent study of the archaeal 20S-PAN complex, the terminal residues of PAN in PA26^{E102A}-PAN adopt α -helices when docked into α -pockets of the *T. acidophilum* 20S³², suggesting that the β -strand conformation might be specific to eukaryotic 20S proteasomes and/or Y Φ motifs. Next, we noticed that the five C-termini that bound in the A state are nearly identical; each predicted to make between five and nine hydrogen bonds in the α -pocket, including the critical salt bridge (Fig. 3a). In contrast, there are fewer contacts (3 H-bonds and no salt bridge) made between the U state peptide and the $\alpha 1/\alpha 2$ -pocket. Another distinction between the U and A states is the configuration of predicted H-bonds between Opt5 and residues in the α -pocket, most notably, α Glu25. In the U state, H-bonding occurs between the carboxy side chain of $\alpha 1$ Glu26 and the hydroxyl of the P2 Tyr. In the A state, hydrogen

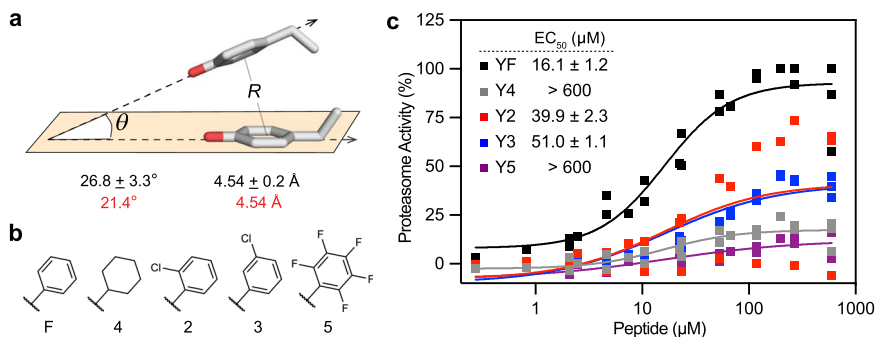


Fig. 4 P3 and P2 residues make intramolecular π -stacking interactions within the *h20S* α -pockets. **a** Experimental and theoretical distances (R) and angles of incidence (θ) for adjacent P2 and P3 tyrosine residues³⁷. Values are of bound Opt5 from the cryo-EM structure ($n = 6$) and reported as mean \pm s.d. See Supplementary Table 3 for individual measurements. **b** Structure of the side chains from unnatural amino acids used in this study. **c** P2 modifications to the π electron density decreases the stimulatory effect of Opt5. Data are normalized to Opt5^{YF} and plotted individually ($n = 2$ to 4). Reported EC₅₀ is a mean of EC₅₀ values calculated from two to four independent experiments with error reported as s.e.m.

bonding is mediated between α Glu25 and the P3 Tyr side chain, which allows for additional contacts to the α -pocket with the hydroxyl group of the P2 Tyr (*i.e.* $\alpha 2/\alpha 3$, $\alpha 4/\alpha 5$, & $\alpha 5/\alpha 6$) (Fig. 3b). The α Glu25 is highly conserved from *T. acidophilum* to humans, including across all seven distinct α -subunits (see Supplementary Fig. 8a), further highlighting the potential importance of this H-bond contact. Together, these studies suggest that polar contacts within the α -pocket, notably H-bonding between the P3 Tyr and α Glu25, are involved in the terminus-mediated gate opening of the *h20S*.

Intramolecular stacking of P3 and P2 residues regulates α -pocket engagement. Considering the heterogeneity of the α -pockets in the *h20S*, we were struck by how the C-termini adopted strikingly similar orientations. For example, when we overlay the six bound C-termini from the PA26^{E102A}-Opt5-*h20S* structure, the P2 and P3 Tyr residues adopt a nearly identical, stacked orientation (see Fig. 3c). Moreover, we noted that of the subset of reported structures of eukaryotic PA-20S complexes^{10,36}, those with C-terminal Y Φ motifs also depicted stacking of the aromatic P2 and P3 residues. In addition, Rpt5 and PA200/Blm10 displayed comparable dihedral angles, which resembled the backbone configuration of Opt5 to suggest that the distinct positioning of the Tyr residues may be a common feature of Y Φ motifs (Supplementary Fig. 10). We next noted how the average distance (R) between the centroids of the P2 and P3 Tyr rings (4.5 ± 0.2 Å) and the average angle (θ) between the rings' planes ($26.8 \pm 3.3^\circ$) closely matched energy-minimized calculations for a Tyr-Tyr dipeptide adopting an off-centered parallel, π stack (Fig. 4a)³⁷. Thus, we postulated that intramolecular π stacking facilitates terminus-dependent gate opening by orienting Opt5 to a competent conformation within the α -pocket. To test this hypothesis, we explored how perturbations to the putative π stack affected the ability for peptide analogs of Opt5^{YF} to stimulate *h20S* activity (Fig. 4b). In dose-response assays, replacement of the Phe side chain with a cyclohexyl group (4) at the P2 position, which should no longer π stack with the P3 Tyr side chain, reduced activity by 80% relative to Opt5^{YF}, (Opt5^{Y4}, EC₅₀ > 600 μ M) (Fig. 4c). Likewise, an inversion of the aromatic quadrupole through a pentafluoro phenylalanine (5) substitution at the P2 position (Opt5^{Y5}) significantly reduced potency (EC₅₀ > 600 μ M) and activity (by 80%) relative to Opt5^{YF}. Less drastic perturbation of the π electron density, through the installation of ortho- or meta-chloro groups (2 and 3; Fig. 4b), had intermediate consequences (Fig. 4c). Although we cannot rule out steric effects, these experiments suggest that intramolecular π -stacking interactions between the Y

and Φ residues contribute to Opt5 activity, perhaps by orienting the motif for engagement with α Glu25.

Valency tunes the sequence preferences for proteasomal activation. Native PAs and other 20S-binding partners are often multivalent, displaying multiple C-termini from a central scaffold. One effect of this valency is that binding avidity is increased, likely to promote cellular assembly of proteasome complexes^{18,26,27,38,39}. However, a less appreciated effect of valency in biological recognition is that it also tunes specificity⁴⁰, amplifying some preferences and allowing more variability in others. To determine if valency influences the sequence preferences of C-terminal recognition by the *h20S*, we generated and tested additional chimeras of the heptavalent PA26^{E102A}, emphasizing on modifications at the P2 and P3 positions because of their aforementioned contributions. Unlike the monovalent peptides, Phe substitution of the P3 Tyr did not affect PA26^{E102A}-Opt5's capacity to stimulate the *h20S* and minimally altered its potency (PA26^{YY} & PA26^{FY}, EC₅₀ = 42.1 ± 1.2 & 33.7 ± 1.2 nM, respectively). Similar to the trends observed with Opt5 peptides, Phe substitution of the P2 Tyr slightly improved the EC₅₀ (PA26^{YF}, 23.5 ± 1.2 nM) (Fig. 5a). Interestingly, the binding affinity for *h20S* remain relatively unchanged when either the P2 or P3 Tyr residues were replaced with Phe. We also noted that PA26^{YY} has a k_{on} that is ~ 3 x slower than that of PA26^{YF} and PA26^{FY}; and yet, the calculated K_d for PA26^{YY} falls within error of the others, owing to its significantly slower k_{off} (Supplementary Fig. 11b). We then generated an alternate, monovalent PA by grafting peptide sequences of interest onto the C-terminus of maltose-binding protein (MBP). Tethering the Opt5 sequence (NLSYYT) to the monovalent MBP, we found that the Opt5 chimera (MBP^{YY}) equipotently stimulated *h20S* (EC₅₀ = 41.8 ± 1.2 μ M) relative to Opt5^{YY} peptide. Likewise, the P2 Phe mutant was more potent (MBP^{YF}, EC₅₀ = 10.0 ± 1.2 μ M), mirroring the SAR from the monovalent peptides. The P3 Phe mutant exhibited decreased activity (MBP^{FY}, EC₅₀ > 400 μ M), however, the activity of MBP^{FY} was relatively diminished compared to its peptide equivalent (Fig. 5b). Together, we conclude that valency appears to allow the *h20S* to be more permissive of missing contacts in the PA's C-termini, whereas monovalent ones largely adhere to the Y Φ motif.

Next, we interrogated the impact valency might have on the relative importance of intramolecular π stacking by generating P3 or P2 Ala mutations in all three PA-types: monovalent peptide, monovalent MBP, and multivalent PA26^{E102A}. In stimulation assays, Ala mutations at either P3 or P2 completely inactivated the monomeric tools: peptides (Opt5^{AY} and Opt5^{YA}) and MBP chimeras (MBP^{AY} and MBP^{YA}) (Supplementary Fig. 4c, d). Thus, without intramolecular π stacking, monovalent PAs seem unable to stimulate *h20S*. However, in the context of multivalent

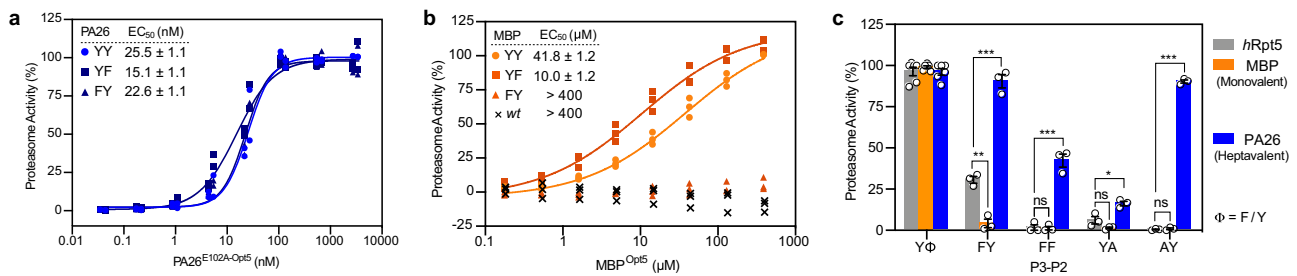


Fig. 5 Valency tunes the sequence preferences of *h2OS* PAs. **a** Proteasome activity of heptameric PA26^{E102A}-Opt5 (YY, circle) and C-terminally modified variants, YP2F mutant (YF, square) and YP3F mutant (FY, triangle). **b** Similar to **a** except using C-terminally modified mutants of MBP, tested alongside wildtype MBP (wt, 'x'). Data are normalized to YY and plotted individually ($n = 3$ or 4). Reported EC₅₀ is a mean of EC₅₀ values calculated from three independent experiments ($n = 3$ or 4) with error reported as s.e.m. **c** Proteasome stimulation by monovalent Opt5 peptides (gray), monovalent MBP^{Opt5} chimeras (orange), and multivalent PA26^{E102A}-Opt5 (blue). YΦ is a composite of YY and YF activators. Data are calculated from three independent experiments (see Supplementary Fig. 4g, h), normalized to YY, and plotted individually (open circles, $n = 6$ or 3) with error reported as s.e.m. *P*-values were calculated using unpaired *t*-test: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

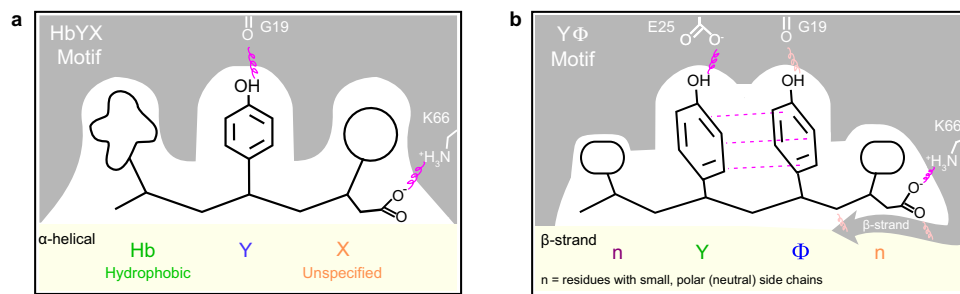


Fig. 6 The YΦ motif model for terminus-dependent *h2OS* activation. **a** Model summarizing the major features of the HbYX motif, for comparison. **b** Model of the $nY\Phi n$ (YΦ) motif summarizes the proposed mechanism for termini-mediated activation of the *h2OS*. The critical interactions (red) include polar contacts (squiggle) between Tyr residues and α Glu25 and α Gly19 and π stacking between the Tyr residues (dotted line). The carboxylate interaction with α Lys66 is also important. Secondary interactions (pink) are denoted for the P1 (orange), P2 (blue), P3 (green), and P4 (purple) residues. In addition, the P5 and P6 are likely to play a role. Note that multivalency seems to allow PAs to, in some cases, overcome a subset of these requirements.

PA26^{E102A}, disrupting π stacking with a P3 Ala mutation (PA26^{AY}) had no effect on hydrolysis (EC₅₀ ~8.3 nM). Rather, the introduction of the P2 Ala mutation (PA26^{YA}) was sufficient to diminish stimulation by 80% relative to PA26^{YY} (EC₅₀ > 3000 nM) (Fig. 5c; Supplementary Fig. 4e). Thus, multivalent PAs do not seem to require either P3 Tyr or intramolecular π stacking to induce gate opening. Collectively, these data show that valency reduces the stringency for termini-mediated activation of the *h2OS* (Fig. 5c).

Discussion

Pioneering studies of the HbYX motif in the archaeal system (Fig. 6a)^{9,12,13}, have contributed significantly to our understanding of 20S activation. Herein, we aimed to deepen that knowledge by characterizing how Rpt5-derived peptides stimulate the *h2OS*. This effort uncovered narrower preferences for the P1 'X' and P3 'hydrophobic' residues and revealed an unanticipated role for the P4 position, suggesting that the *h2OS* exhibits distinct sequence preferences. Although only a subset of amino acids could be explored because of solubility criteria, only Ala and Thr were strongly preferred at P1, a result that contrasts with the relatively broad specificity exhibited in the archaeal system (see Supplementary Fig. 3). A potential reason for this preference was revealed by structural studies, which showed that Thr in the P1 position engages in H-bonding within *h2OS* α -pockets. In contrast, the structural rationale for the preference at P4 is less clear. Tentatively, small, aliphatic residues might be preferred over bulkier residues due to potential steric clashes with a neighboring β -strand-turn in the α -pockets. In addition, a subset of polar amino acids, such as Ser, Thr, and Asp, were also relatively

preferred at P4 (see Fig. 1e), indicating that H-bonding in that region may also be involved. However, further work is needed to elucidate the importance of these interactions, including whether they contribute to binding, gate opening or both. Finally, another key observation was that a Tyr was prioritized at P3, whereas the P2 position accepted Φ . Structural and activity analyses suggested that these preferences might be based on a requirement for intermolecular H-bonds with the α -pocket (esp. a conserved glutamate) and intramolecular π -stacking interactions. Together, these structure-function observations led us to propose the YΦ motif as an addendum to the HbYX model, which is focused on the *h2OS* (Fig. 6b).

In our *h2OS*-PA26^{E102A}-Opt5 structure, we noted that the C-termini bind in two different states (A and U). Although we concluded that PA valency influences which interactions mediate proteasome activation, we speculate that the A and U states might represent putative docking poses for Opt5^{YΦ} peptides occupying α -pockets of the *h2OS*, providing potential mechanistic insight into peptide-mediated gate opening. For instance, to explain the unprecedented contribution of the P3 Tyr in peptide/monomeric PAs we conjectured that Opt5^{FY} likely binds the *h2OS* in the U state. Hence, we postulate that the diminished activity of Opt5^{FY} is ultimately mediated by displacement of the salt bridge, which is attributed to fewer polar interactions within the α -pocket, namely the displaced H-bond caused by replacing the P3 Tyr with a Phe residue. Altogether, this would suggest that H-bond formed between the P3 Tyr and α Glu25 acts as an anchor that promotes the YΦ motif in the open gate competent A state. In support of this model, reported structures of both yeast²⁰ and human³⁶ 26 S proteasomes, the P3 & P2 residues of the Rpt3 subunit (F & Y,

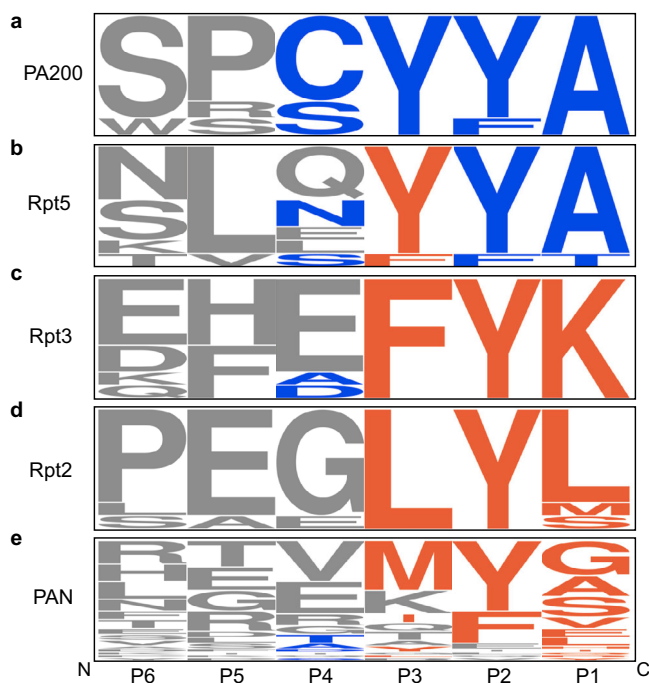


Fig. 7 The YΦ motif is strictly conserved in the eukaryotic, monovalent activator PA200. **a–e** WebLogo (<http://weblogo.berkeley.edu>) (48) representation of the last six residues of **(a)** PA200 ($n = 7$), **(b)** Rpt5 ($n = 9$), **(c)** Rpt3 ($n = 9$), **(d)** Rpt2 ($n = 9$), and **(e)** PAN ($n = 41$). Amino acid residues at the P1, P2, P3, and P4 (Y only) positions are colored according to sequence preferences of the HbYX (orange) or nYΦn (blue) motifs. Unspecified, along with P5 and P6 residues are in gray. The height of each residue represents the relative frequency. The complete entry list in Supplementary Table 4.

respectively), depict displaced salt bridges and C-termini that adopt α -helical turns, as observed in the U state. Moreover, Rpt3 is the least stimulatory of the Rpts with HbYX motifs^{21,22}, further suggesting that interactions observed in our structure of the heptameric PA26^{E102A}-Opt5 could apply to binding of and activation by Opt5 peptides. However, it is plausible that the U state may simply be a binding pose unique to the $\alpha 1/\alpha 2$ -pocket. The $\alpha 1/\alpha 2$ -pocket is also the designated docking site for Rpt3's C-terminus, potentially suggesting that the U state is an artifact of this specific PPI and neither specific to interactions of the YΦ motif nor implicated in C-termini-dependent gate opening. Additional structures, particularly of peptide-bound complexes, will be required to deepen our understanding of the YΦ gating mechanism.

Using a series of monovalent (free peptides and MBP chimeras) and heptameric (PA26^{E102A}) scaffolds, we explored how valency tunes sequence preferences. We found that the multivalent PA26^{E102A} was able to overcome a subset of the requirements suggested by the YΦ motif, most strikingly, a Tyr residue at the P3 position. This finding implies that monovalent, eukaryotic activators, such as PA200/Blm10¹⁰, might have distinct requirements at their C-termini. As a first step in asking that question, we aligned the C-terminal sequences of naturally occurring PAs from multiple organisms (Supplementary Table 4). Indeed, PA200 ended in residues that closely match the YΦ motif (Fig. 7a). Moreover, a recent structure revealed that the C-terminus of PA200 adopts both the β -strand and the intramolecular off-centered π stacking⁴¹. Thus, we postulate that monovalent PAs will strictly adhere to the YΦ model. A different picture emerged from the alignment of C-terminal sequences

from multivalent PAs, including archaeal PAN and the three Rpt subunits of PA700 containing HbYX motifs (Rpt2, Rpt3, & Rpt5) from multiple organisms. This exercise showed that the C-terminus of Rpt5 closely resembled the YΦ motif, except that the P3 position accommodates a Phe residue, and the P4 position samples a wider range of residues (Fig. 7b). Notably, Rpt5 is the most stimulatory of the Rpt subunits^{21,22}, perhaps supporting the influence of the YΦ motif on *h20S* activation, even in the context of some multivalent PAs. The less stimulatory Rpt2 and Rpt3 subunits, in contrast, displayed more divergent sequences from the YΦ motif, most notably because Tyr was not conserved at the P3 position (Fig. 7c, d). Even less consensus was observed when we aligned the C-termini of available PAN sequences. Specifically, the P2 tended to be Φ , while the residues at the remaining positions seemed arbitrary (Fig. 7e). Unlike PA700, PAN is homo-oligomeric and so no individual monomer (e.g. Rpt5) is predicted to make an outsized contribution to overall binding or stimulation. Thus, in the case of PAN, valency seems to largely override the requirements of both the YΦ and the more permissive HbYX motifs. Nonadherence to the canonical HbYX model was similarly noted in an examination of another hexavalent regulator of the archaeal 20S, Cdc48³⁹. Together, these observations suggest that C-terminal sequence identity, plus a contribution from valency, combine to dictate the PPIs between PAs and the proteasome. Further analysis of the SAR at these PPIs is important because additional binding partners of the 20S, with HbYX-like motifs of varying sequences and valences, have been identified in eukaryotes, but the full extent to which they regulate the proteasome remains limited^{39,42–44}. Hence, it may be useful to create sub-categories of HbYX-like models, in which the valency of the PA is a key, defining characteristic.

Loss of proteasome function is implicated in many devastating proteinopathies, including neurodegenerative disorders⁴⁵. Consistent with this idea, boosting 20S activity, by introducing either native⁴⁶ or engineered activators⁴⁷, has been found to be protective in cell-based disease models. These observations have motivated campaigns to discover drug-like molecules that mimic the activity of PAs²³. This current work could be important in that effort, by defining the SAR associated with *h20S* α -pockets and providing a template for the rational design of pharmacological proteasome activators^{48,49}. For example, we suspect that the arrangement of the YΦ motif, through π stacking, minimizes the entropic costs of binding for Opt5. For molecules that conform to this pharmacophore model, potency should be improved by rigidification of the bioisosteric equivalent to the Tyr-Tyr π stack. We also found that Opt5 docked into 6 of the 7 available α -pockets with remarkable similar structural states. This consensus structure might be a good starting point for understanding the requisite occupancy and binding interactions for inducing gate opening with a small molecule. Thus, the current work provides a potential blueprint for designing novel proteasome activators, while also extending our knowledge of the mechanisms of *h20S* activation.

Methods

Reagents. Human 20S proteasome was purchased from the Proteasome Center. pET28 6His PA26 (49 V) cloning vector was a gift from Philip Coffino. pET 6His MBP TEV LIC cloning vector (1 M) was a gift from Scott Gradia (Addgene plasmid # 29656).

Strains and Plasmids. The *E. coli* strain Top10 was used for propagating plasmids. BL21 (DE3) cells were used for the expression and purification of recombinant proteins. Primers are listed in Supplementary Table 5.

Peptide synthesis. Peptides were synthesized by Fmoc solid-phase peptide synthesis on a Syro II peptide synthesizer (Biotage) at ambient temperature and atmosphere on a 12.5 μ M using either pre-loaded Wang resin or Rink amide resin (Sigma-Aldrich). Coupling reactions were run with 4.9 eq. of HCTU (*O*-(1H-6-

chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate), 5 eq. of Fmoc-AA-OH and 20 eq. of *N*-methylmorpholine (NMM) in 500 μ l of *N,N*-dimethyl formamide (DMF). Fmoc-AA-OH was double coupled for 8 min while shaking for each position. Fmoc deprotection was conducted with 500 μ l 40% 4-methylpiperidine in DMF for 3 min, followed by 500 μ l 20% 4-methylpiperidine in DMF for 10 min and six washes with 500 μ l of DMF for 3 min. Acetylation of the N-terminus was achieved by reacting 20 eq. acetic anhydride and 20 eq. *N,N*-diisopropylethylamine (DIPEA) in 1 mL DMF for 2 h while shaking. Peptides were cleaved with 500 μ l of cleavage solution (95% trifluoroacetic acid (TFA), 2.5% water and 2.5% triisopropylsilane) while shaking for 2 h. Crude peptides were precipitated in 15 ml cold 1:1 diethyl ether: hexanes and air-dried overnight. Crude peptides were solubilized in a 1:1:1 mixture of DMSO: water: acetonitrile, filtered, and purified by high-performance liquid chromatography (HPLC) on an Agilent Pursuit 5 C18 column (5 mm bead size, 150 \times 21.2 mm) using Agilent PrepStar 218 series preparative HPLC. The mobile phase consisted of A, 0.1% TFA in water and B, 0.1% TFA in acetonitrile. Peptides were purified to >95% homogeneity confirmed by liquid chromatography-mass spectrometry before solvent was removed by lyophilization. Peptides were resuspended in 1:1 water: acetonitrile, lyophilized again in tared tubes and stocks were stored at -20°C .

Dynamic light scattering (DLS). From 10 mM stocks in DMSO, three 4-fold serial dilutions of each peptide were prepared in assay buffer that was filtered with Millex-GS 0.22 μ m sterile filter unit to final concentrations of 600, 150, and 37.5 μ M. A final volume of 20 μ l was added to a 384-well plate (Corning, 3540) and the unsealed plate was centrifuged at 25 $^{\circ}\text{C}$ for 45 s at 500 \times g. Colloidal aggregation was measured using a Wyatt Technologies DynaPro Plate Reader II (acquisition time of 2 s, 10 acquisitions, with auto attenuation, and temperature controlled at 25 $^{\circ}\text{C}$). Peptides that were suspected of forming colloidal aggregates based on DLS, visual inspection of DMSO stocks, and/or atypical activity profiles (Supplementary Fig. 1) were excluded from further analysis.

Protein expression and purification in bacteria. All proteins were produced in *E. coli* BL21 (DE3) and stored at -80°C .

Wildtype (WT) PA26 and the E102A mutant (PA26^{E102A}), which was generated by site-directed mutagenesis, constructs (*T. brucei*, 49 V, N-terminal 6His tag) were expressed from a pET28 vector. *E. coli* were grown in terrific broth (TB) with the requisite antibiotic at 37 $^{\circ}\text{C}$, induced with 1 mM isopropyl β -d-1-thiogalactopyranoside (IPTG) in log phase, and grown for an additional 3 h. Cells were harvested by centrifugation, resuspended in His binding buffer (20 mM Tris pH 7.9, 20 mM NaCl, 10 mM imidazole) supplemented with protease inhibitors, sonicated, and clarified by centrifugation. Clarified lysate was applied to Ni-NTA His-Bind Resin (Novagen). The resin was washed with binding buffer and then eluted in batches with buffer containing increasing amounts of imidazole (up to 300 mM). Purified protein was treated with 1 mM dithiothreitol (DTT) for 1 h at ambient temperature and imidazole was removed by overnight dialysis into storage buffer (20 mM Tris pH 8.0, 200 mM NaCl) for later use.

Chimeric PA26^{E102A} constructs were cloned from PA26 WT and integrated into pMCSG7 vector with a N-terminal 6His tag through LIC for protein expression. *E. coli* were grown in TB with the requisite antibiotic at 37 $^{\circ}\text{C}$, induced with 1 mM IPTG in log phase, cooled to 18 $^{\circ}\text{C}$ and grown overnight. Cells were harvested by centrifugation, resuspended in His resin binding buffer (20 mM Tris pH 7.9, 20 mM NaCl, 10 mM imidazole) supplemented with protease inhibitors, sonicated, and clarified by centrifugation. Clarified lysate was applied to Ni-NTA His-Bind Resin (Novagen). The resin was washed with binding buffer and then eluted in batches with buffer containing increasing amounts of imidazole (up to 800 mM). Purified protein was treated with 1 mM DTT for 1 hr at ambient temperature and imidazole was removed by overnight dialysis into storage buffer for later use.

MBP WT and chimeric MBP (N-terminal 6His tag) were expressed from pET and pMCSG7 constructs, respectively. *E. coli* were grown in TB supplemented with 0.2% glucose (without glycerol) with the requisite antibiotic at 37 $^{\circ}\text{C}$, induced with 1 mM IPTG in log phase, cooled to 25 $^{\circ}\text{C}$ and grown overnight. Cells were harvested by centrifugation, resuspended in amylose resin binding buffer (20 mM Tris pH 7.8, 200 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA)) supplemented with protease inhibitors, sonicated and clarified by centrifugation. Clarified lysate was applied to amylose resin (New England Biolabs, E8021S). The resin was washed with binding buffer and then eluted with 200 mM maltose in binding buffer. Purified protein was dialyzed into a storage buffer for later use.

Protein concentrations were determined using Pierce bicinchoninic acid (BCA) Protein Assay Kit (Thermo Scientific, 23225) and bovine serum albumin (BSA) as a standard. Purified proteins and BSA were serially diluted into storage buffer and then 10 μ l of the sample was incubated with 100 μ l of the BCA working reagent at 37 $^{\circ}\text{C}$ for 30 min. Absorbance was measured at 562 nm using a Spectramax M5 microplate reader (Molecular Devices; SoftMax Pro 6.5.1). A standard curve was generated from BSA measurements and the concentrations were determined for each purified protein in Microsoft Excel (16.16.23).

Proteasome activity assays. Stimulation of *h20S* proteasome (Proteasome Center) was assayed in 384-well plates (Greiner Bio-One, 781209) using the fluorogenic

peptide substrates suc-LLVY-amc (AnaSpec, AS-63892), boc-LRR-amc (AdipoGen, AG-CP3-0014) or FAM-LFP (5-FAM-AKVYPYPMEEK(QXL520)-NH₂; AnaSpec) in assay buffer containing 50 mM Tris pH 7.5, 10 mM MgCl₂, 200 μ M adenosine triphosphate (ATP), 1 mM DTT, and 0.01% Pluronic F-68[®] (Gibco Life Technologies, 24040032) in a total volume of 30 μ l. Human 20S (*h20S*) (final concentration of 4 nM) was incubated in the presence or absence of activators (peptides, 0.2–600 μ M; MBP 0.2–400 μ M; PA26, 0.04–3500 nM) at ambient temperature for 5 min. Substrate (suc-LLVY-amc, 10 μ M; boc-LRR-amc, 20 μ M; or FAM-LFP, 100 nM) was added immediately before reading. Fluorescence intensity of suc-LLVY-amc and boc-LRR-amc (excitation, 355 nm; emission, 440 nm; cutoff 435 nm) or FAM-LFP (excitation, 490 nm; emission, 520 nm; cutoff, 515 nm) were monitored at 30 to 60 s intervals for 30 min at 25 $^{\circ}\text{C}$ using a Spectramax M5 microplate reader (Molecular Devices; SoftMax Pro 6.5.1). The hydrolysis rate was calculated from the slope of the curve between 100 and 500 s (500 and 1000 s for boc-LRR-amc) in arbitrary units (RFU per s).

Data were processed and fit in GraphPad Prism 8.1.2. The baseline hydrolysis was normalized to the total mean activity for the lowest concentration of every activator assayed in a given plate and the maximal hydrolysis was normalized as reported. The normalized activity was plotted relative to log₁₀(activator). Data was fit to the model for log(agonist) versus response (variable slope). In Eq. (1), $X = \log[\text{activator}, \mu\text{M}]$ except for PA26-based activators, which are in nanomolar.

$$Y = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{(1 + 10^{(\log EC_{50} - X) \times \text{HillSlope}})} \quad (1)$$

Biotinylation of *h20S*. Commercially available *h20S* was dialyzed into PBS (Gibco Life Technologies) pH 7.5 and biotinylated with 20 eq of NHS-Biotin (Thermo Scientific, 20217) at 4 $^{\circ}\text{C}$ for 2 h. Biotin-conjugated *h20S* was isolated from excess, unreacted NHS-Biotin with a PD-10 desalting column (Amersham Biosciences), and concentrated with a centrifugal filter concentrator (MilliporeSigma, UFC5100). Aliquots of biotinylated *h20S* were snap-frozen and stored at -80°C for later use.

Binding kinetics analysis. Biolayer interferometry (BLI) data of PA26 activators were measured using an Octet RED384 (ForteBio). The reactions were carried out in black 384-well plates (Greiner Bio-One, 781209) at 25 $^{\circ}\text{C}$ with a volume of 85 μ l per well in BLI buffer (assay buffer containing 0.2% (w/v) BSA (Sigma)). Biotinylated *h20S* (10 nM) were immobilized on streptavidin (SA) biosensor. Serial dilutions of PA26 in BLI buffer were used as analyte. Association was observed by immersing loaded biosensors into solutions of analyte for 450 s. No binding was detected of the analyte to unloaded sensors. Dissociation was observed by transferring the sensor to a well-containing binding buffer and no analyte for 1200 s. Binding affinities (K_d) and kinetic parameters (k_{on} and k_{off}) were calculated from a global fit (2:1 heterogeneous ligand model) of the data using the Octet data analysis software (Octet Data Analysis 7.1).

Cryo-electron microscopy grid preparation. *h20S* proteasome and PA26^{E102A}-Opi5 activators were mixed by dilution into a buffer system (20 mM Tris pH 7.5, 20 mM NaCl, 10 mM KCl, 1 mM DTT, and 0.025% Nonidet P-40) to a final concentration of 2 μ M and 4 μ M, respectively. The mixture was centrifuged at 25 $^{\circ}\text{C}$ for 30 min at 12,000 \times g in a tabletop centrifuge and placed on ice.

Three microliters of the *h20S*-PA26^{E102A}-Opi5 solution were applied to R2/2 400-mesh grids (Quantifoil) that had been plasma treated for 25 s using a glow discharger (Electron Microscopy Sciences) operated under atmospheric gases doped with amylamine. The grids were manually blotted to near dryness with Whatman no. 1 filter paper inside a cold room (4 $^{\circ}\text{C}$, 95% humidity). The sample application and blotting process was repeated twice more to increase on-grid protein concentration. After the third blot, the grid was gravity plunged into liquid ethane using a homemade system and stored under liquid nitrogen⁵⁰.

Cryo-electron microscopy data collection and image processing. Cryo-EM data were acquired using the Legion 3.3 software for automated data acquisition⁵¹ using a Titan Krios (Thermo Fisher) equipped with a K2 Summit (Gatan) direct electron detector in counting mode (Supplementary Table 2). Movies were collected by navigating to the center of a hole and image shifting a beam of \sim 600 nm diameter to 10 targets situated at the periphery of the 2- μ m hole. A total of 13,329 movies were recorded at a nominal magnification of 29,000 \times (1.03 \AA magnified pixel size) and composed of 29 frames (250 ms per frame, \sim 50 e^{-} \AA^{-2} per movie). Movie collection was guided by real-time assessment of image quality using the Appion 3.3 image processing environment⁵². Frame alignment and dose weighting were performed in real-time using UCSF Motioncor2⁵³. Estimation of the contrast transfer function (CTF) was performed on aligned, unweighted, micrographs using Gctf 1.06⁵⁴. All data post-processing steps were conducted in RELION 3.0^{55,56}.

Single particle analysis was performed in RELION 3.0. Particle picking was conducted using the AutoPick function and resulted in 2,650,143 particle picks that were extracted using a box size of 420 \times 420 pixels, down-sampled to 70 \times 70 pixels, for reference-free 2D classification. 497,630 particles belonging to the 2D classes demonstrating features characteristic of secondary structural elements were subjected to 3D refinement and subsequent 3D classification ($k = 8$). A 3D template of the PA26-bound archaeal proteasome (PDB ID: 1YAU)⁸ was lowpass-

filtered to 60 Å and used to guide the initial 3D refinement and 3D classification. 288,515 particles corresponding to 3D classes without artefactual features were chosen for further data processing. To minimize the detrimental effects of pseudo-symmetry (C2) on resolution, the raw particles were C2 symmetry expanded, 3D refined, and a python script was used to determine the x- and y-shifts required to reposition the proteasome gate at the center of the particle box. The particles at the new center were re-extracted to serve as individual asymmetric units without down-sampling in a box of 256 × 256 pixels, resulting in a total of 577,030 particles. A round of reference-free 2D classification enabled us to remove the ends of h20S particles that lacked an activator molecule. This combined expansion and classification approach yielded 521,860 particles. We subjected these particles to 3D classification ($k = 8$) and selected 384,939 particles whose parent 3D classes resembled a fully assembled h20S-PA26^{E102A-Opt5} complex. Following removal of particle duplicates with a python script, 326,676 particles were CTF and beam tilt refined. Further 3D classification without alignment ($k = 10$) revealed two classes with an unresolved alpha subunit helix. We attributed this finding to rotational misalignment around the C7 pseudo-symmetry, the longitudinal axis of the complex and discarded these particles. The remaining 247,362 particles were then sorted and pruned by z-score, yielding 234,960 particles for final 3D refinement (2.9 Å) (Supplementary Figs. 6 and 7).

Atomic model building. The atomic model was built by using the h20S and PA26 (PDB IDs: 5A0Q and 1YAU, respectively) as templates and rigid body fitting each subunit with Chimera into the electron density map. Subunits for which there was no density as a result of the symmetry expansion and re-extraction processing approach, were removed from the template. The modified templates were subject to one cycle of morphing and simulated annealing in PHENIX 1.10.1, followed by a total of 5 real-space refinement macrocycles with atomic displacement parameters, secondary structure restraints, local grid searches, and global minimization. After PHENIX refinement, manual real-space refinement was performed in Coot 0.8.9.1. Multiple rounds of real-space refinement in PHENIX (five macrocycles, Ramachandran and rotamer restraints, no morphing, no simulated annealing) and Coot were performed to address geometric and steric discrepancies identified by the RCSB PDB validation server. To ensure atomic models were not overfit as a result of real-space refinement, map-to-model FSCs were calculated with PHENIX 1.10.1⁵⁷ (Supplementary Fig. 7b). All images were generated using UCSF Chimera 1.14 and PyMol 2.3.2.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The structural data generated in this study have been deposited in the Electron Microscopy Databank under EMD-22259 and in the Protein Databank (PDB) under 6XMJ. The raw and processed data generated in this study are provided in the Source Data file. The reporting summary for this Article is available as a Supplementary Information file. All data supporting the findings of this study is available from the corresponding author upon reasonable request. Source data are provided with this paper.

Received: 10 July 2020; Accepted: 11 February 2022;

Published online: 09 March 2022

References

- Kwon, Y. T. & Ciechanover, A. The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem. Sci.* **42**, 873–886 (2017).
- Groll, M. et al. A gated channel into the proteasome core particle. *Nat. Struct. Biol.* **7**, 1062–1067 (2000).
- Zhang, Z. et al. Identification of an activation region in the proteasome activator REGalpha. *Proc. Natl Acad. Sci. U.S.A.* **95**, 2807–2811 (1998).
- Whitby, F. G. et al. Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature* **408**, 115–120 (2000).
- Köhler, A. et al. The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release. *Mol. Cell* **7**, 1143–1152 (2001).
- Löwe, J. et al. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Sci. (80-)*. **268**, 533–539 (1995).
- Förster, A., Whitby, F. G. & Hill, C. P. The pore of activated 20S proteasomes has an ordered 7-fold symmetric conformation. *EMBO J.* **22**, 4356–4364 (2003).
- Förster, A., Masters, E. I., Whitby, F. G., Robinson, H. & Hill, C. P. The 1.9 Å structure of a proteasome-11S activator complex and implications for proteasome-PAN/PA700 interactions. *Mol. Cell* **18**, 589–599 (2005).
- Smith, D. M. et al. Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's a Ring Opens the Gate for Substrate Entry. *Mol. Cell* **27**, 731–744 (2007).
- Sadre-Bazzaz, K., Whitby, F. G., Robinson, H., Formosa, T. & Hill, C. P. Structure of a Blm10 Complex Reveals Common Mechanisms for Proteasome Binding and Gate Opening. *Mol. Cell* **37**, 728–735 (2010).
- Majumder, P. et al. Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. *Proc. Natl Acad. Sci. USA.* **116**, 534–539 (2019).
- Rabl, J. et al. Mechanism of Gate Opening in the 20S Proteasome by the Proteasomal ATPases. *Mol. Cell* **30**, 360–368 (2008).
- Yu, Y. et al. Interactions of PAN's C-termini with archaeal 20S proteasome and implications for the eukaryotic proteasome-ATPase interactions. *EMBO J.* **29**, 692–702 (2010).
- Tian, G. et al. An asymmetric interface between the regulatory and core particles of the proteasome. *Nat. Struct. Mol. Biol.* **18**, 1259–1267 (2011).
- Wehmer, M. et al. Structural insights into the functional cycle of the ATPase module of the 26S proteasome. *Proc. Natl Acad. Sci.* **114**, 1305–1310 (2017).
- Chen, S. et al. Structural basis for dynamic regulation of the human 26S proteasome. *Proc. Natl Acad. Sci. USA.* **113**, 12991–12996 (2016).
- Zhu, Y. et al. Structural mechanism for nucleotide-driven remodeling of the AAA-ATPase unfoldase in the activated human 26S proteasome. *Nat. Commun.* **9**, 1–12 (2018).
- Adams, G. M., Crotchett, B., Slaughter, C. A., DeMartino, G. N. & Gogol, E. P. Formation of proteasome-PA700 complexes directly correlates with activation of peptidase activity. *Biochemistry* **37**, 12927–12932 (1998).
- Eisele, M. R. et al. Expanded Coverage of the 26S Proteasome Conformational Landscape Reveals Mechanisms of Peptidase Gating. *Cell Rep.* **24**, 1301–1315.e5 (2018).
- De la Peña, A. H., Goodall, E. A., Gates, S. N., Lander, G. C. & Martin, A. Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation. *Science (80-)*. **362**, eaav0725 (2018).
- Gillette, T. G., Kumar, B., Thompson, D., Slaughter, C. A. & DeMartino, G. N. Differential roles of the COOH termini of AAA subunits of PA700 (19 S regulator) in asymmetric assembly and Activation of the 26 S proteasome. *J. Biol. Chem.* **283**, 31813–31822 (2008).
- Kumar, B., Kim, Y. C. & DeMartino, G. N. The C terminus of Rpt3, an ATPase subunit of PA700 (19 S) regulatory complex, is essential for 26 S proteasome assembly but not for activation. *J. Biol. Chem.* **285**, 39523–39535 (2010).
- Opoku-Nsiah, K. A. & Gestwicki, J. E. Aim for the core: suitability of the ubiquitin-independent 20S proteasome as a drug target in neurodegeneration. *Transl. Res.* **198**, 48–57 (2018).
- Trader, D. J., Simanski, S., Dickson, P. & Kodadek, T. Establishment of a suite of assays that support the discovery of proteasome stimulators. *Biochim. Biophys. Acta* **1861**, 892–899 (2017).
- Gizyńska, M. et al. Proline- and Arginine-Rich Peptides as Flexible Allosteric Modulators of Human Proteasome Activity. *J. Med. Chem.* **62**, 359–370 (2019).
- Glickman, M. H., Rubin, D. M., Fried, V. A. & Finley, D. The Regulatory Particle of the *Saccharomyces cerevisiae* Proteasome. *Mol. Cell Biol.* **18**, 3149–3162 (1998).
- Stadtmueller, B. M. et al. Structure of a Proteasome Pba1-Pba2 Complex Implications for Proteasome Assembly, Activation, and Biological Function. *J. Biol. Chem.* **287**, 37371–37382 (2012).
- Ganesh, A. N., Donders, E. N., Shoichet, B. K. & Shoichet, M. S. Colloidal aggregation: From screening nuisance to formulation nuance. *Nano Today* **19**, 188–200 (2018).
- Fujiwara, K., Toda, H. & Ikeguchi, M. Dependence of α -helical and β -sheet amino acid propensities on the overall protein fold type. *BMC Struct. Biol.* **12**, 6–15 (2012).
- Stadtmueller, B. M. et al. Structural models for interactions between the 20S proteasome and its PAN/19S activators. *J. Biol. Chem.* **285**, 13–17 (2010).
- Smith, D. M. et al. ATP binding to PAN or the 26S ATPases causes association with the 20S proteasome, gate opening, and translocation of unfolded proteins. *Mol. Cell* **20**, 687–698 (2005).
- Yu, Z. et al. Allosteric coupling between α -rings of the 20S proteasome. *Nat. Commun.* **1–11**, 832113 (2019).
- Thibaudeau, T. A., Anderson, R. T. & Smith, D. M. A common mechanism of proteasome impairment by neurodegenerative disease-associated oligomers. *Nat. Commun.* **9**, 1097 (2018).
- Harshbarger, W., Miller, C., Diedrich, C. & Sacchettini, J. Crystal Structure of the Human 20S Proteasome in Complex with Carfilzomib. *Structure* **23**, 418–424 (2015).
- Park, S. et al. Reconfiguration of the proteasome during chaperone-mediated assembly. *Nature* **497**, 512–516 (2013).
- Huang, X., Luan, B., Wu, J. & Shi, Y. An atomic structure of the human 26S proteasome. *Nat. Struct. Mol. Biol.* **23**, 778–785 (2016).

37. Chelli, R., Gervasio, F. L., Procacci, P. & Schettino, V. Stacking and T-shape competition in aromatic-aromatic amino acid interactions. *J. Am. Chem. Soc.* **124**, 6133–6143 (2002).
38. Knowlton, J. R. et al. Structure of the proteasome activator REGalpa (PA28alpha). *Nature* **390**, 639–643 (1997).
39. Barthelme, D. & Sauer, R. T. Bipartite determinants mediate an evolutionarily conserved interaction between Cdc48 and the 20S peptidase. *Proc. Natl Acad. Sci. U.S.A.* **110**, 3327–3332 (2013).
40. Kiessling, L. L., Gestwicki, J. E. & Strong, L. E. Synthetic multivalent ligands as probes of signal transduction. *Angew. Chem.—Int. Ed.* **45**, 2348–2368 (2006).
41. Toste Rêgo, A. & da Fonseca, P. C. A. Characterization of Fully Recombinant Human 20S and 20S-PA200 Proteasome Complexes. *Mol. Cell* **76**, 138–147.e5 (2019).
42. Li, X., Thompson, D., Kumar, B. & DeMartino, G. N. Molecular and cellular roles of PI31 (PSMF1) protein in regulation of proteasome function. *J. Biol. Chem.* **289**, 17392–17405 (2014).
43. Ramachandran, K. V. & Margolis, S. S. A mammalian nervous-system-specific plasma membrane proteasome complex that modulates neuronal function. *Nat. Struct. Mol. Biol.* **24**, 419–430 (2017).
44. Hirano, Y. et al. A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. *Nature* **437**, 1381–1385 (2005).
45. Schmidt, M. & Finley, D. Regulation of proteasome activity in health and disease. *Biochim. Biophys. Acta* **1843**, 13–25 (2014).
46. Li, J., Powell, S. R. & Wang, X. Enhancement of proteasome function by PA28 overexpression protects against oxidative stress. *FASEB J.* **25**, 883–893 (2011).
47. Qu, J. et al. Specific Knockdown of α -Synuclein by Peptide-Directed Proteasome Degradation Rescued Its Associated Neurotoxicity. *Cell Chem. Biol.* **27**, 751–762 (2020).
48. Jones, C. L., Njomen, E., Sjögren, B., Dexheimer, T. S. & Tepe, J. J. Small Molecule Enhancement of 20S Proteasome Activity Targets Intrinsically Disordered Proteins. *ACS Chem. Biol.* **12**, 2240–2247 (2017).
49. Coleman, R. A. & Trader, D. J. Development and Application of a Sensitive Peptide Reporter to Discover 20S Proteasome Stimulators. *ACS Comb. Sci.* **20**, 269–276 (2018).
50. Crooks, G., Hon, G., Chandonia, J. & Brenner, S. WebLogo: A Sequence Logo Generator. *Genome Res* **14**, 1188–1190 (2004).
51. Suloway, C. et al. Automated molecular microscopy: The new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
52. Lander, G. C. et al. Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* **166**, 95–102 (2009).
53. Zheng, S. Q. et al. MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
54. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
55. Scheres, S. H. W. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
56. Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUS in RELION-2. *Elife* **5**, 1–21 (2016).
57. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **68**, 352–367 (2012).

Acknowledgements

This work is dedicated to L. L. Kiessling on the occasion of her 60th birthday. This work was supported by grants from the Tau Consortium and NIH AG053619 (to J.E.G.), NIH

AG061697 (to G.C.L.), and NIH/NIGMS R01GM083960 (to A.S.). Additional support included an NSF GRFP fellowship and an HHMI Gilliam fellowship (to K.A.O.-N.) and American Cancer Society fellowship (A.H.P.). All cryo-EM data were collected at the Scripps Research electron microscopy facility and we thank B. Anderson for his microscope support. We thank J.C. Ducom at the Scripps Research High Performance Computing for computational support. Computational analyses of EM data were performed using shared instrumentation funded by NIH S10OD021634 to G.C.L. The authors thank the laboratory of C.S. Craik (University of California San Francisco) for assistance with peptide synthesis, the laboratory of P. Coffino (The Rockefeller University) for the PA26 construct, as well as A.K. Mapp, M.A. Ravalin and L. Pope for helpful suggestions and comments on the manuscript.

Author contributions

K.A.O.-N. and J.E.G. designed the studies and wrote the manuscript. All authors edited the manuscript. K.A.O.-N., S.K.W., and N.C. conducted biochemistry experiments, performed data analysis, and generated necessary reagents. A.H.P. performed cryo-EM sample preparation, data collection, and data processing. J.E.G., K.A.O.-N., A.S., and G.C.L. provided funding.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-28864-x>.

Correspondence and requests for materials should be addressed to Jason E. Gestwicki.

Peer review information *Nature Communications* thanks Christopher Hill, Jetze Tepe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022