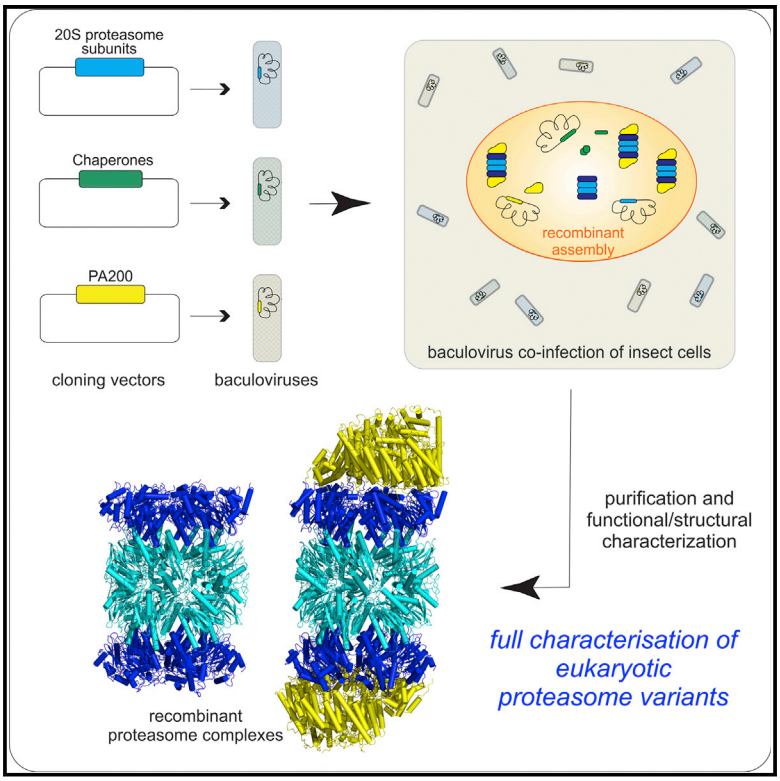
Article

Molecular Cell

Characterization of Fully Recombinant Human 20S and 20S-PA200 Proteasome Complexes

Graphical Abstract



Highlights

- Recombinant human 20S proteasomes and 20S-PA200 complexes are characterized
- Cryo-EM reveals intricate 20S-PA200 interactions and PA200-bound cofactors
- PA200 binding is allosterically communicated to the proteolytic active sites
- Basis to fully characterize the function and regulation of proteasome variants

Authors

Ana Toste Rêgo, Paula C.A. da Fonseca

Correspondence

pauladf@mrc-lmb.cam.ac.uk

In Brief

We report the characterization of recombinant human 20S and 20S-PA200 proteasome complexes, confirming their structural and functional integrity. In particular, the cryo-EM structure of 20S-PA200 reveals unique features, including unusual 20S proteasome conformational rearrangements, with functional consequences.

Toste Rêgo & da Fonseca, 2019, Molecular Cell *76*, 138–147
October 3, 2019
https://doi.org/10.1016/j.molcel.2019.07.014

CellPress

CellPress

Molecular Cell
Article

Characterization of Fully Recombinant Human 20S and 20S-PA200 Proteasome Complexes

Ana Toste Rêgo[1] and Paula C.A. da Fonseca[1,2,*]
[1]MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK
[2]Lead Contact
*Correspondence: pauladf@mrc-lmb.cam.ac.uk
https://doi.org/10.1016/j.molcel.2019.07.014

SUMMARY

Proteasomes are essential in all eukaryotic cells. However, their function and regulation remain considerably elusive, particularly those of less abundant variants. We demonstrate the human 20S proteasome recombinant assembly and confirmed the recombinant complex integrity biochemically and with a 2.6 Å resolution cryo-EM map. To assess its competence to form higher-order assemblies, we prepared and analyzed recombinant human 20S-PA200, a poorly characterized nuclear complex. Its 3.0 Å resolution cryo-EM structure reveals the PA200 unique architecture; the details of its intricate interactions with the proteasome, resulting in unparalleled proteasome α ring rearrangements; and the molecular basis for PA200 allosteric modulation of the proteasome active sites. Non-protein cryo-EM densities could be assigned to PA200-bound inositol phosphates, and we speculate regarding their functional role. Here we open extensive opportunities to study the fundamental properties of the diverse and distinct eukaryotic proteasome variants and to improve proteasome targeting under different therapeutic conditions.

INTRODUCTION

The ubiquitin-proteasome pathway, ubiquitous to all eukaryotes, is critical for proteostasis and for the strictly regulated proteolysis of key proteins that signal for fundamental processes, including cell cycle progression, stress responses, DNA repair, and the onset of apoptosis. The proteasome is the crucial protease complex at the downstream end of this pathway. It comprises a proteolytic core, the 20S proteasome, which, in eukaryotes, is formed by hetero-heptameric rings of α1–α7 and β1–β7 subunits that stack into $\alpha_{(1-7)}\beta_{(1-7)}\beta_{(1-7)}\alpha_{(1-7)}$ barrel-shaped assemblies (Groll et al., 1997; Löwe et al., 1995). The β1, β2, and β5 subunits are proteolytically active, each with distinct substrate binding and cleavage specificities that are caspase-, trypsin-, and chymotrypsin-like, respectively (Heinemeyer et al., 1997). The proteolytic active sites are located within the proteasome inner chamber, and their full activation requires the binding of regulatory complexes at the outer surfaces of the 20S proteasome α rings (Groll et al., 2000). The 19S regulatory particle, which associates with the core to form the 26S proteasome, is the activator that recruits, engages, and unfolds fully folded ubiquitinated proteins for degradation (Bhattacharyya et al., 2014; Collins and Goldberg, 2017; Livneh et al., 2016).

Apart from the constitutive 26S proteasome, less characterized but important complexes add to the intricacies of proteasome function and regulation. The immunoproteasome, thymoproteasome, and spermatoproteasome are core variants found in higher eukaryotes (Kniepert and Groettrup, 2014; Murata et al., 2018). Additionally, ATP-independent proteasome regulators, less characterized and alternative to the 19S regulatory particle, include 11S particles (PA28/REG) (Huber and Groll, 2017; Mao et al., 2008) and PA200, a still poorly characterized 200-kDa monomeric proteasome regulator associated with DNA repair, spermatogenesis, chromatin remodeling, and acetylation-dependent histone degradation (Blickwedehl et al., 2008; Hoffman et al., 1992; Khor et al., 2006; Mandemaker et al., 2018; Qian et al., 2013; Ustrell et al., 2002). These alternative regulators do not have the ubiquitin binding or unfolding activities of the 19S regulatory particles, but their binding to the 20S proteasome modulates its peptidase activities against small peptides and unfolded proteins. On the other hand, they can still be involved in the degradation of ubiquitin-tagged substrates in the context of hybrid assemblies by binding to 20S proteasomes singly capped with one 19S regulatory particle (Cascio et al., 2002; Kopp et al., 2001). Last, proteasome function can be further modulated by a range of transiently binding ancillary proteins and post-translational modifications.

Despite recent progress, significantly driven by cryoelectron microscopy (cryo-EM) studies (Bard et al., 2018; Collins and Goldberg, 2017; Hochstrasser, 2016; Wehmer and Sakata, 2016), the functional and tight regulation mechanisms of the human proteasome are still not fully characterized. Contributing factors include its endogenous compositional and structural variability, lack of mutagenesis studies because of the proteasome's obligatory role and consequent lethality associated with the disruption of its function, and limited availability of purified homogeneous human complexes, particularly of less abundant variants. This can in principle be overcome by recombinant assembly. Recombinant archaeal 20S proteasomes, which are simpler complexes formed by homo-heptameric rings of α and β subunits, can be assembled by co-expressing its two subunits in *E. coli* (Zwickl et al., 1992). Such an approach is not feasible for the human 20S proteasome

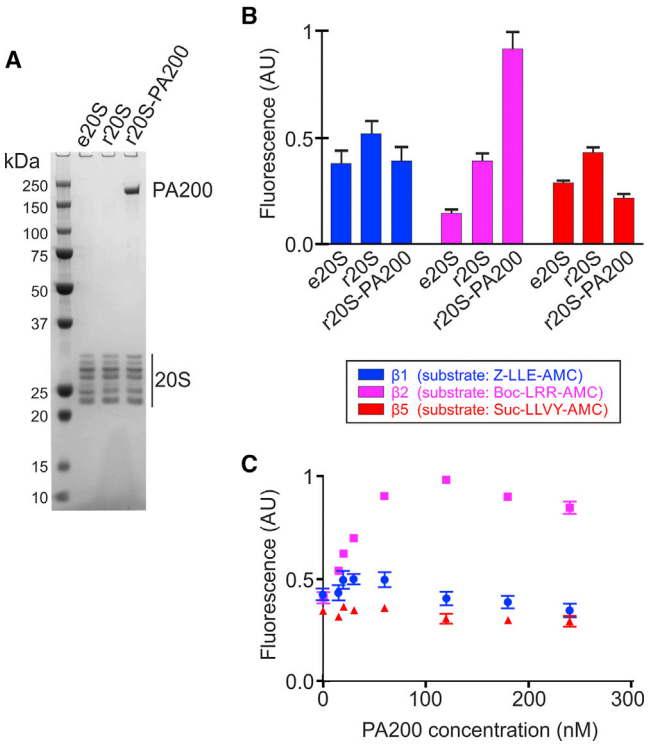


Figure 1. Biochemical Characterization of Recombinant Human 20S Proteasome and 20S-PA200 Complexes

(A) SDS-PAGE of endogenous (e20S) and recombinant (r20S) human 20S proteasomes and recombinant human 20S-PA200 complexes (r20S-PA200) (see also Figures S1A and S1B).

(B) Proteolytic activities of e20S, r20S, and r20S-PA200 against the fluorogenic substrates Z-LLE-AMC (blue), Boc-LRR-AMC (magenta), and Suc-LLVY-AMC (red), specific for the proteasome β1 caspase-like, β2 trypsin-like, and β5 chymotrypsin-like active sites, respectively. Similar 20S proteasome molarities were used in each assay, as shown in (A).

(C) Proteolytic activities of recombinant human 20S proteasomes in the presence of increasing concentrations of PA200, color coded as in (B) (see also Figure S1C).

Error bars are represented as mean ± SD.

because of its tightly regulated biogenesis pathway (Budenholzer et al., 2017; Livneh et al., 2016). This involves the step assembly of two copies of 14 closely related subunits, α1–α7 and β1–β7, which is assisted by 5 dedicated chaperones: the heterodimers PAC1-PAC2 and PAC3-PAC4 and POMP. The proteasome α rings are initially formed and serve as scaffolds for the successive and ordered incorporation of β subunits. A dimerization step follows, where two pre-assembled half-proteasomes fuse together, triggering a final maturation step that involves the self-cleavage of the β subunit pre-peptides. These are N-terminal extensions present in immature β subunits, believed to shield their proteolytic activity until formation of the full 20S proteasome, and they may also contribute as scaffolds for proteasome assembly. Here we describe the successful assembly of recombinant human 20S proteasomes in insect cells, which we expand to the functional and structural characterization of a recombinant human 20S-PA200 proteasome complex.

RESULTS

Recombinant Human 20S Proteasomes

We investigated the recombinant assembly of human 20S proteasomes in insect cells as a proof of principle for the preparation of other highly homogeneous eukaryotic proteasome complexes, particularly those difficult to obtain from endogenous sources. We used a bacoluvirus-insect cell expression system where we successfully co-expressed the 19 individual human proteins required, the 14 proteasome constitutive subunits α1–α7 and β1–β7, and the 5 dedicated proteasome assembly chaperones PAC1-PAC2, PAC3-PAC4, and POMP (Figure S1A). We observed that incorporating all 14 proteasome subunits (α1–α7/β1–β7) into a single baculovirus, ensuring their similar expression levels and, therefore, the correct proteasome subunit stoichiometry, significantly facilitates the assembly of recombinant complexes.

The recombinant proteasomes were purified by affinity and size exclusion chromatography steps (Figure S1B). SDS-PAGE shows identical protein profiles for recombinant and endogenous human proteasomes (Figure 1A). Moreover, the proteasome peptidase activities, measured against fluorogenic substrates specific for each of the three proteasome active sites, are also comparable for both complexes, albeit with slightly increased activity for the recombinant complex (Figure 1B). Finally, we determined the cryo-EM structure of the recombinant human 20S proteasome at a resolution of 2.6 Å (Figures 2A–2C, S2A, S2C, S2E, and S2G). This structure, consistent with that of the endogenous complex, unambiguously confirms recombinant proteasome structural integrity and full maturation, with the absence of all β subunit pre-peptides, which was also confirmed by mass spectrometry. We therefore demonstrate that eukaryotic proteasomes can be successfully assembled by recombinant expression, opening new perspectives for full characterization of the diverse eukaryotic proteasomes.

Recombinant Human 20S-PA200 Complexes

Preparation of recombinant proteasomes is particularly relevant for the study of less abundant and still poorly characterized variants. In this context, we prepared recombinant human 20S-PA200 complexes by co-expression in insect cells. The 20S-PA200 assembly further supports the recombinant human 20S proteasome structure integrity. 20S proteasomes are fully activated by the binding of regulators at their α ring outer surfaces. Any divergence of these surfaces from their endogenous state should result in capping impairment. To confirm that the recombinant human 20S proteasome retains its competence to be capped, we co-expressed it with PA200. SDS-PAGE clearly reveals co-purification of all proteasome subunits and PA200 (Figure 1A), with the formation of 20S-PA200 complexes confirmed by cryo-EM imaging (Figure S2B). The recombinant 20S-PA200 complex, free from other human proteasome variant contaminants, is highly suitable for a detailed biochemical characterization. PA200 is known to modulate the peptidase activities of the 20S proteasome. We find that, in the 20S-PA200 complex, trypsin-like activity, associated with the β2 active site, is enhanced compared with the uncapped recombinant 20S proteasome, whereas there is slight inhibition of the β1

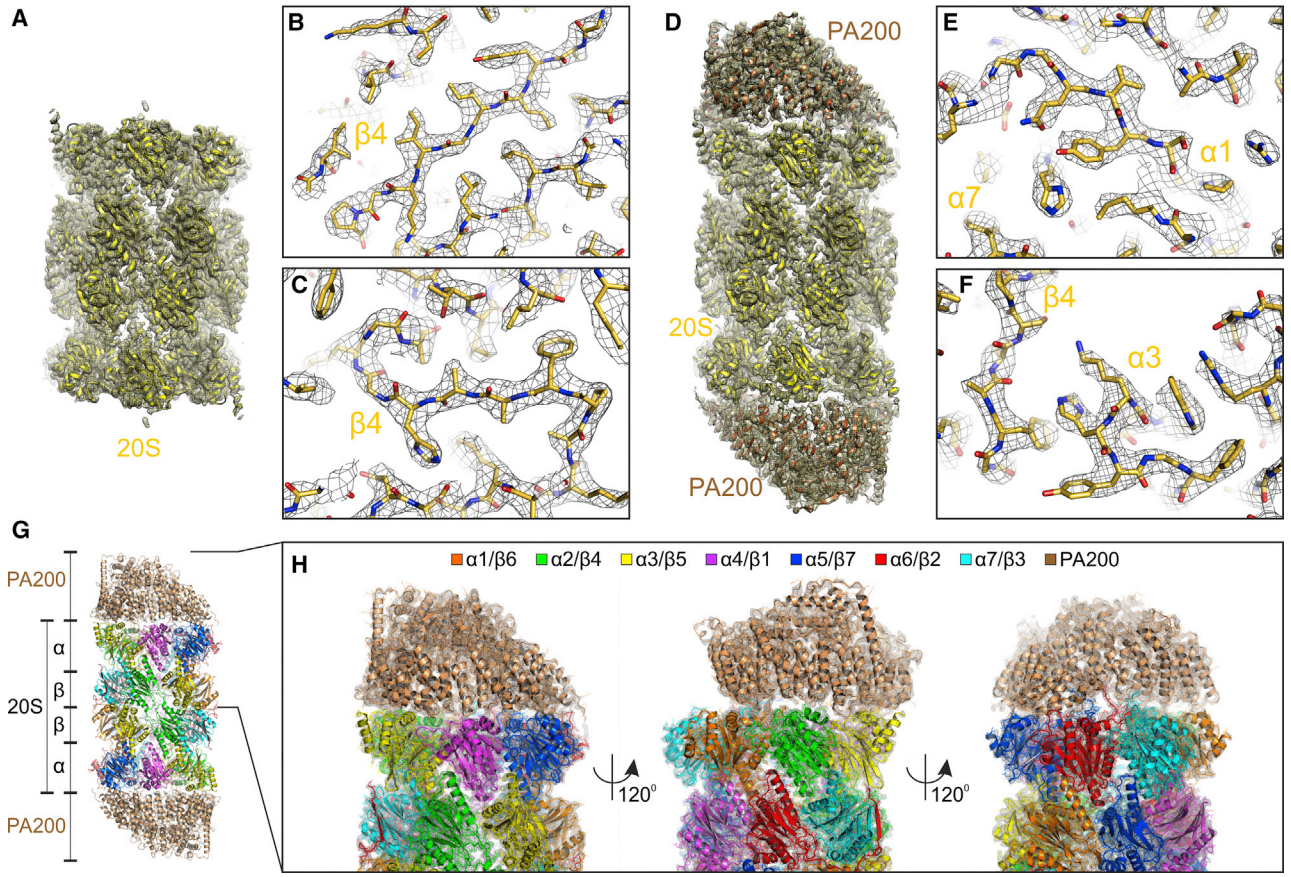


Figure 2. Cryo-EM Structures of the Recombinant Human 20S Proteasome and 20S-PA200 Complex

(A) Cryo-EM structure of the recombinant human 20S proteasome with a fitted atomic model (see also Figure S2).

(B and C) Two close-up views of the structure shown in (A), which has well-resolved side chains throughout.

(D) Cryo-EM structure of the recombinant human 20S-PA200 complex with a fitted atomic model (see also Figures S2 and S3).

(E and F) Two close-up views of the structure shown in (D), which has well-resolved side chains throughout.

(G) Overall view of the 20S-PA200 atomic model, with major domains indicated.

(H) Close-up views of the 20S-PA200 cryo-EM map (gray mesh) with a fitted atomic model (cartoon representation), with each subunit color-coded as indicated at the top.

In (A)–(F) The cryo-EM maps are shown as mesh and the protein models as cartoons (A and D) or sticks (B, C, E, and F).

and β5 active sites (Figure 1B). However, such a pattern of proteolytic activity modulation differs from that described previously by adding endogenous bovine testis PA200 containing fractions to purified human 20S proteasomes, which resulted mainly in activation of proteasome caspase-like activity (Ustrell et al., 2002). Therefore, to confirm our results, we expressed and purified human PA200 on its own and added it to purified 20S proteasomes. Upon addition of increasing concentrations of PA200, both endogenous and recombinant proteasome peptidase activities show modulation identical to the recombinant 20S-PA200 complex, validating its functional integrity (Figures 1C and S1C).

Structural information currently available on endogenous 20S-PA200 is limited to a low-resolution (23 Å) cryo-EM map of a bovine complex (Ortega et al., 2005). Taking advantage of our recombinant human 20S-PA200 and the recent advances in cryo-EM, we determined the cryo-EM structure of the double-capped complex at a resolution of 3 Å (Figures 2D–2H, S2B, S2D, S2F, S2H, S3A, and S3B). The quality of the map allows unambiguous building of an atomic model of the full complex *ab initio* for PA200. PA200 is mostly formed by helical repeats spiraling around a dome above the proteasome α ring (Figures 2H, S3A, and S3B). Although its overall arrangement resembles that of the yeast proteasome-bound Blm10, a yeast ortholog of PA200 whose structure was solved by X-ray crystallography at 3.4 Å resolution (Sadre-Bazzaz et al., 2010), because of significant differences between the two complexes, the structure of Blm10 could not be used as a direct template for PA200 modeling.

Relevant and Unique 20S-PA200 Structural Features

Our 20S-PA200 structure reveals extensive interactions between PA200 and the surface of all α subunits, except for α7, where the interactions occur only at its N-terminal tail. There are two main anchor regions among the PA200 and 20S proteasome interaction network (Figures 3, S4A, and S4B). One

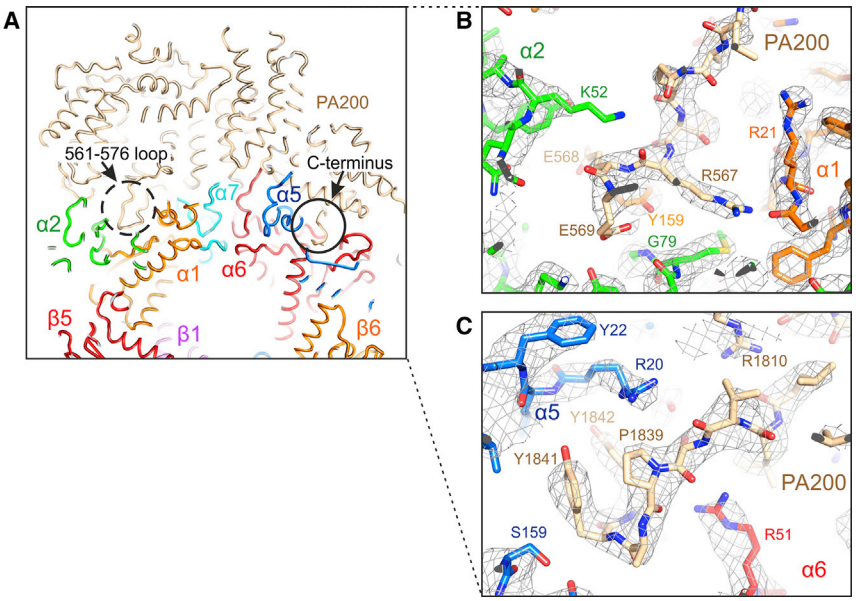


Figure 3. PA200 Main Docking Sites on the Proteasome α Rings

(A) Cartoon representation of the PA200 structure, showing the two major anchor regions of PA200 at the proteasome α ring, one involving the PA200 loop formed by residues 561–576 (dashed circle, left) and the other involving the PA200 C terminus (solid circle, right).

(B) Close-up view of interactions between the PA200 loop, residues 561–576, and the proteasome subunits α1 (orange) and α2 (green).

(C) Close-up view of interactions between the PA200 C terminus and the proteasome subunits α5 (blue) and α6 (red).

In (A)–(C) the 20S-PA200 structure is oriented to best depict the interactions highlighted. In (B) and (C), the cryo-EM maps are shown as gray mesh and the atomic models as sticks. The protein-protein interaction network involving these two PA200 anchor regions and the 20S proteasome are represented in Figures S4A and S4B. See also Figure S5.

involves the PA200 C-terminal ^{1838}SPCYYA^{1843} residues, which include and extend upstream of the HbYX motif (hydrophobic-tyrosine-other) common to other 20S proteasome binding proteins. These residues, except for Pro1839, form a tightly packed interaction network with the HbYX motif-interacting pocket at the proteasome α5-α6 interface (Figures 3A, 3C and S4B). The other main anchor region comprises the analogous binding pocket at the α1-α2 interface, where an extended PA200 loop, formed by residues 561–576, docks (Figures 3A, 3B, and S4A). Although this loop tightly binds to a known HbYX-interacting pocket, its sequence differs from that of this motif.

The extensive interactions between PA200 and the proteasome result in significant α ring conformational rearrangements. In 20S-PA200, there is α subunit radial displacement, away from the proteasome long axis, particularly evident for α3 and α4 (Figure 4). The N-terminal tails of α5–α7 are fully ordered and relocated from the proteasome axis into grooves at the inner surfaces of the PA200 dome (Figures 5 and S4C–S4E), whereas, for the remaining α subunits, these tails are disordered and not recovered in our map. Remarkably, for α3, the H0 helix is also disordered, resulting in an unusually wide α ring opening (Figures 4C and S3C), as discussed below. Despite the proteasome α subunit rearrangements, the location of PA200 immediately above the α ring opening, rather than aligned with the 20S proteasome central axis (Figures 2H, S3A, and S3B), results in the proteasome inner chamber remaining considerably enclosed, restricting proteolytic active site access to small peptides and unfolded proteins, as reported previously (Ustrell et al., 2002).

Interestingly, among the unique features of the 20S-PA200 structure are two positively charged grooves at the distal outer surface of PA200 (Figure 6A). Channels at the center of these grooves appear to be significantly obstructed by well-defined densities, not accounted for by the protein model, consistent with cofactors coordinated by dense clusters of PA200 lysine and arginine side-chains (Figures 6, S5, and S6). We find that one of these densities is compatible with a molecule of inositol hexakisphosphate (InsP6) (Figures 6A–6D and S5A), where only one of the InsP6 phosphate groups does not interact directly with PA200; consequently, it is somewhat less well recovered in our map. The density obstructing the channel of the second PA200 groove appears to be less recovered, but its shape and coordination by a cluster of positively charged residues led us to tentatively assign it to (5,6)-bisdiphosphoinositol tetrakisphosphate (5,6[PP]$_2$-InsP4) (Figures 6A, 6E–6G, S5, and S6B).

Allosteric Modulation of the 20S Proteasome Active Sites by PA200

It is well established that the 20S proteasome peptidase activities are modulated by binding of regulators to the α rings, but direct structural evidence of the associated alterations in the active sites has been missing. Our 20S-PA200 cryo-EM structure shows that the proteasome α ring conformational rearrangements, induced by PA200 binding, result in global allosteric structure adjustments that extend to the proteasome β subunits. A close comparison of the active sites in our 20S proteasome and 20S-PA200 complex atomic models reveals slight differences in the lining of the three proteasome substrate binding pockets (Figure 7), which explain the proteolytic activity modulation observed upon PA200 binding (Figures 1B and 1C). The caspase-, trypsin-, and chymotrypsin-like proteasome activities, associated with the β1, β2, and β5 active sites, were probed using fluorogenic substrates with glutamate, arginine, and tyrosine residues at position P1, respectively. Our 20S-PA200 structure shows that there is widening of the β2 S1 pocket in 20S-PA200 that may facilitate arginine binding, consistent with the observed trypsin-like activation. On the other hand, lesser narrowing at the β1 and β5 sites can reduce their accessibility to glutamate and tyrosine residues, respectively, explaining the observed slight inhibition of the caspase- and chymotryptic-like activities

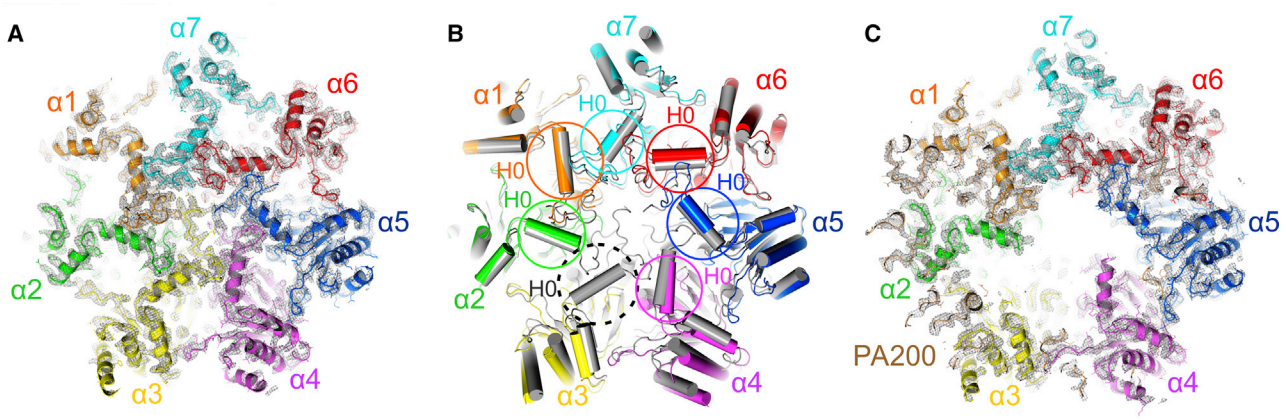


Figure 4. PA200 Induced Conformational Changes in the 20S Proteasome α Rings

(A) The closed α ring outer surface of the 20S proteasome.

(B) Superimposition of the α ring outer-surface atomic models of the recombinant human 20S proteasome (gray cartoon) and 20S-PA200 (cartoon with subunits color coded). The H0 helices of the 20S proteasome α1–α2 and α4–α7 are indicated by solid circles color coded as in 20S-PA200, whereas the H0 helix of α3, which is disordered in 20S-PA200, is encircled by a black dashed circle.

(C) The open α ring outer surface of the 20S-PA200 complex.

In (A) and (C), the cryo-EM maps are shown as gray mesh, and the atomic models are represented as cartoons.

observed in the 20S-PA200 complex. Moreover, the apparent negative charge lining the β5 S1 pocket in the 20S-PA200 complex, compared with the hydrophobic pocket found in the 20S proteasome, may also facilitate binding of arginine. This could also contribute to the raised tryptic activity we observed for the 20S-PA200 complex, even when, in the apo 20S proteasome, the tryptic activity is normally assigned only to the β2 active site.

DISCUSSION

We demonstrate that eukaryotic proteasomes can be successfully assembled by recombinant expression. This important new tool in proteasome research can be exploited in the preparation of highly homogeneous samples suitable for fundamental functional and structural studies that can be supported by mutagenesis studies and for therapeutic drug discovery and development. Three proteasome inhibitors are already in clinical use against cancer, primarily multiple myeloma, and new compounds are being developed for higher efficacy and for the treatment of other varied pathological conditions, including inflammatory diseases and parasitic infections (Ettari et al., 2016; Khare et al., 2016; Li et al., 2016; Park et al., 2018).

Recombinant expression of proteasome complexes can be particularly important for studying less abundant proteasome variants and complexes difficult to purify in high quantities from endogenous sources, which may include those from infectious parasites (Khare et al., 2016; Li et al., 2016). Another relevant example, where the availability of recombinant proteasomes could have a significant effect, is the specific targeting of the human immunoproteasome, a variant related to the constitutive 20S proteasome, where the proteolytic active subunits are replaced by interferon-γ-induced iβ1, iβ2, and iβ5 counterparts. The nature of the human immunoproteasome impairs its purification from endogenous sources, with a detrimental effect on its characterization and drug design and development. In one remarkable effort to overcome this limitation, a yeast strain was modified to incorporate in its 20S proteasomes chimeric yeast and human β5 subunits, designed to mimic the proteolytic activity of the endogenous human iβ5 subunit while incorporating the yeast elements required for proteasome assembly and maturation (Huber et al., 2016). However, although these humanized proteasomes are already contributing to drug development, they are not physiological and only mimic iβ5 without providing information regarding ligand binding at the iβ1or iβ2 active sites and no information regarding possible ligand-induced allosteric effects on the human immunoproteasome.

Recombinant complexes can significantly contribute to the full characterization of proteasomes bound to ATP-independent regulators, which are less abundant and significantly less studied than the canonical 26S proteasome. Structures of 11S complexes have been determined by X-ray crystallography; namely, a human PA28α_7 homoheptamer (Knowlton et al., 1997) and mouse PA28α_7, PA28β_7, and PA28$\alpha_4\beta_3$ (Huber and Groll, 2017). Structural information regarding the interaction of 11S complexes with 20S proteasomes was obtained with the X-ray structures of chimeric complexes formed by yeast and archaeal 20S proteasomes and *Trypanosoma brucei* PA26, a PA28 homolog (Förster et al., 2003, 2005; Whitby et al., 2000). These structures reveal displacement of the N-terminal loops of the proteasome α subunits with opening of a proteasome axial channel without significant additional proteasome conformational changes (Figures S3C and S3D). Prior to the work described here, the structure of mammalian 20S-PA200 complexes have been restricted to a 23 Å resolution cryo-EM map that also showed opening of a proteasome axial channel but without providing information regarding detailed molecular interactions or rearrangements because of resolution constraints (Ortega et al., 2005). On the other hand, from the crystal structure of a yeast 20S proteasome bound to Blm10, a PA200 homolog, it has been suggested that both Blm10 and PA200 would

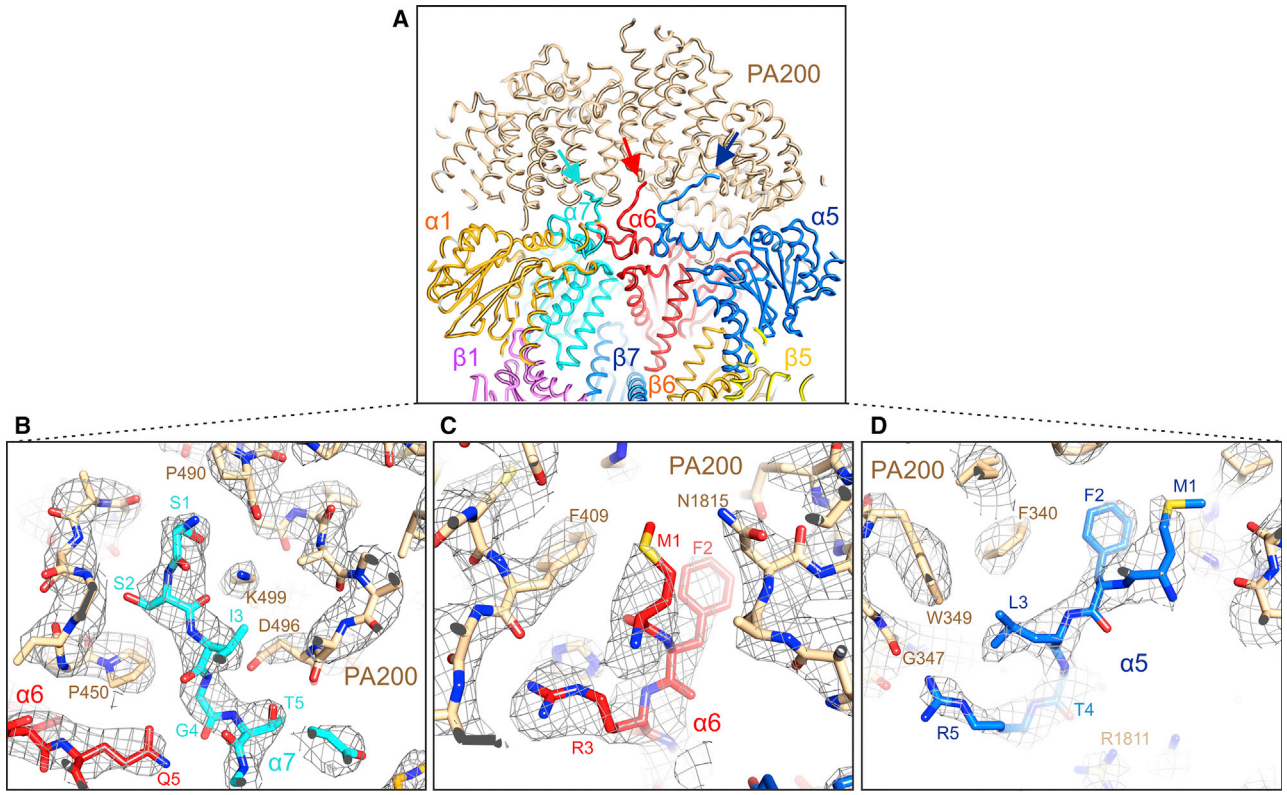


Figure 5. Interaction between the N-Terminal Loops of the Proteasome α Subunits with PA200

(A) Cartoon representation of the N-terminal tails of the proteasome α5–α7 subunits, indicated by arrows, at the PA200 dome inner surface.

(B–D) Close-up views of the N termini of the proteasome subunits α7 (B), α6 (C), and α5 (D). The cryo-EM maps are shown as gray mesh and the atomic models as sticks.

In (A)–(D), the 20S-PA200 structure is oriented to best depict the highlighted interactions. The protein-protein interaction network involving the N-terminal tails of the proteasome α5–α7 subunits and PA200 are represented in Figures S4C–S4E.

rearrange the N-terminal loops of the 20S proteasome α subunits in a way somewhat similar to that proposed for 11S regulators, with only a localized effect at the α subunit N-terminal tails at the proteasome axis (Sadre-Bazzaz et al., 2010). In contrast, we show an intricate interaction network between PA200 and the proteasome that results in proteasome conformational changes with functional implications.

Analysis of the recombinant human 20S proteasome and 20S-PA200 complexes provides significant and unexpected new insights, with uniquely large proteasome subunit conformational changes induced by binding of PA200. These include axial displacements and significant rearrangements of the proteasome α subunits, involving not only their N-terminal loops but, significantly, also the H0 helix of α3, resulting in an unusually wide proteasome α ring opening (Figures 4, S3C, and S3D). In comparison, in the yeast 20S-Blm10 structure, all α subunit H0 helices are ordered, and, consequently, the α ring opening appears much less prominent (Sadre-Bazzaz et al., 2010; Figures S3C and S3D). Notably, the α subunit rearrangements observed so far in proteasome open-state conformations (Chen et al., 2016; Wehmer et al., 2017) are also significantly subtler than in the 20S-PA200 complex (Figures S3C and S3D). Nevertheless, the extensive rearrangements induced by PA200 binding, as we describe, reflect a high degree of conformational plasticity of the proteasome α ring, which may well be required for the processing of at least some proteasome substrates and may not be restricted to the 20S-PA200 complex.

Previous studies showed strong evidence of allosteric communication between proteasome regulators binding at the outer surface of the α rings and the proteolytic active sites within the proteasome inner chamber (Bech-Otschir et al., 2009; Chu-Ping et al., 1994; Haselbach et al., 2017; Kleijnen et al., 2007; Ma et al., 1992; Ruschak and Kay, 2012). In this context, it is well established that binding of regulators at the 20S proteasome α rings modulates the proteolytic activities of β1, β2, and β5. However, the detailed structural rearrangements of the proteasome active sites associated with such allosteric modulation have not been described previously for any proteasome regulator. Earlier studies attributed this modulation to the opening of a proteasome axial gate, characterized by displacements of the α subunit N-terminal loops away from the proteasome long axis (Chen et al., 2016; Sadre-Bazzaz et al., 2010; Wehmer et al., 2017; Whitby et al., 2000). It has been proposed that such an axial gate can facilitate access of small peptides to the proteasome inner chamber, where the proteolytic active sites are located, resulting in enhanced peptidase activities. However, this simplified model suggests that all

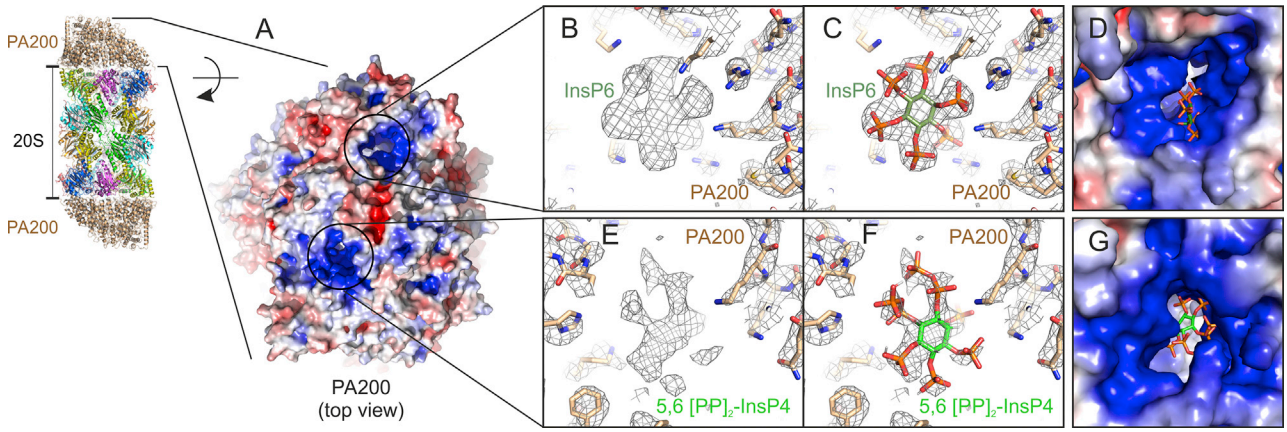


Figure 6. Relevant Features in the Structure of the Human PA200

(A) van der Waals surface representation, colored by charge, of the PA200 distal outer surface (oriented as indicated on the left), showing two prominent positively charged grooves, indicated by solid rings.

(B) Close-up view of the non-protein density blocking the channel on the positively charged groove indicated in (A).

(C) The densities for the cofactor in (B) are fitted with a model of InsP6. Only one of the InsP6 phosphate groups does not interact directly with PA200 and, consequently, is somewhat less well recovered in our map.

(D) The location of InsP6 (B and C) within its PA200 groove.

(E) Close-up view of the non-protein density blocking the channel of the positively charged groove indicated in (A).

(F) The densities for the cofactor in (E) are fitted with a model of 5,6[PP]$_2$-InsP4.

(G) The location of 5,6[PP]$_2$-InsP4 within its PA200 groove.

See also Figures S5 and S6.

20S proteasome peptidase activities would be equally activated upon gate opening and, therefore, cannot account for the more complex modulation of each of the individual proteolytic active sites that is observed.

Our data show that, despite the unusually wide opening at the proteasome α ring induced by PA200 binding (Figures 4, S3C, and S3D), there is only significant activation of proteasome trypsin-like activity (Figure 1). Such modulation is not consistent with the previously reported activation of mainly the human proteasome caspase-like activity by addition of a bovine testis PA200-enriched fraction (Ustrell et al., 2002), but this discrepancy is likely due to the different experimental approach used. Here we confirmed the integrity of the peptidase activities of our 20S-PA200 complex by titrating purified endogenous and recombinant human 20S proteasomes with homogeneously purified PA200, which resulted in identical proteolytic activity patterns (Figures 1 and S1C). Nevertheless, if the effects on proteolytic activity were due to a simple axial gate opening, then an equal effect should be expected for all active sites, which is not observed. Indeed, our 20S-PA200 structure shows that the rearrangements observed at the α subunits are propagated allosterically into the β subunits, resulting in the different conformational changes at each of the proteolytic active sites that we describe in Results. These structural readjustments explain, at the molecular level, how PA200 binding to the 20S proteasome differentially modulates each of the individual proteasome proteolytic activities (Figures 1 and 7); namely, the selective activation of its trypsin-like activity and slight inhibition of its caspase- and chymotryptic-like activities. Although the detailed molecular basis of allosteric modulation of the proteasome active sites has not yet been described for any other proteasome, regulator-related effects should be expected because it is known that different proteasome regulators induce different effects on each of the three proteasome proteolytic activities.

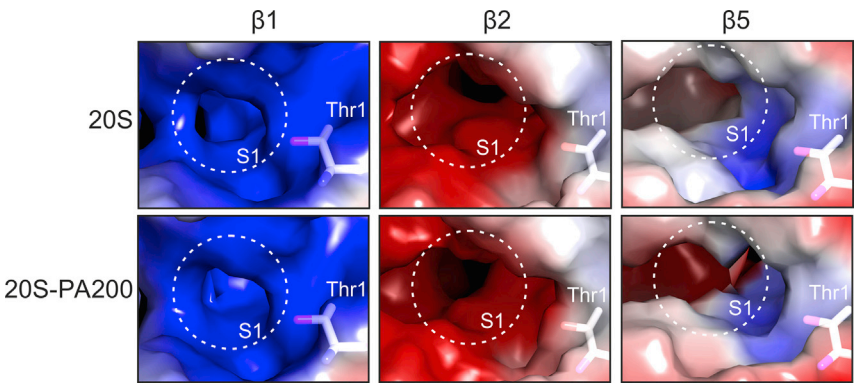


Figure 7. Comparison of the Proteolytic Active Sites in the Human 20S Proteasome and 20S-PA200 Complexes

Shown are van der Waals surface representations, colored by charge, of the three proteasome active sites (β1, β2, and β5) of the recombinant 20S proteasome (top row) and 20S-PA200 complexes (bottom row), viewed from the proteasome inner cavity. White dashed circles indicate the S1 pocket of each active site.

Our structural data allow some speculation regarding the PA200 functional mechanisms. PA200 is involved in acetylation-dependent degradation of histones (Mandemaker et al., 2018; Qian et al., 2013), and in this context, it was predicted that its residues 1650–1738 fold into an acetyl-lysine binding bromodomain, with a hydrophobic pocket formed by residues Phe1676, Asn1716, and Phe1717 (Qian et al., 2013). However, our 20S-PA200 structure reveals that, although these residues are indeed neighbors, the side chain of Phe1676 is buried in the protein structure and oriented away from those of Asn1716 and Phe1717. Although we do not identify bromodomains in our PA200 structure, it is noteworthy that we assigned non-protein densities in our 20S-PA200 cryo-EM map to inositol phosphates bound to PA200 (Figures 6 and S5). Inositol phosphates have been shown to bind and regulate the catalytic activity of class I histone deacetylases (Millard et al., 2013; Watson et al., 2012, 2016). Therefore, it seems reasonable to suggest that PA200-bound inositol phosphates may be correlated with the PA200's role in the degradation of acetylated histones, although this needs further investigation.

The results presented here serve as a framework for further studies ultimately aiming at a complete characterization of the function and regulation of the different eukaryotic proteasomes and their variants. Our results open very substantial new research opportunities that can be explored not just in therapeutic drug development but that can also serve as a basis for fully understanding the critical role of proteasomes within all eukaryotic cells.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Lines
  - Bacterial Strains
- METHOD DETAILS
  - Cloning and generation of baculovirus
  - Expression of proteasome complexes
  - Purification proteasome complexes
  - Expression and purification of PA200
  - Proteasome activity assay
  - Cryo-electron microscopy
  - Single particle analysis
  - Molecular modeling
  - Graphic representations
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2019.07.014.

ACKNOWLEDGMENTS

We thank E. Rajendra for help and advice regarding the Multibac and Baculovirus expression systems, J. García-Nafría for advice regarding IVA cloning, J. Shi for Baculovirus facility management, M. Skehel for mass spectrometry, and J. Grimmett and T. Darling for computing assistance. This study was supported by the MRC-LMB Electron Microscopy Facility. We are grateful to T. Grant for advice regarding the use of cisTEM, E. Morris for discussions and help with the preparation of cryo-EM grids, and M. Kišonaitė, P. Afanasyev, J. Tafilaku, E. Morris, and D. Barford for critical reading of the manuscript. This work was supported by Medical Research Council grant MC_UP_1201/5 (to P.C.A.d.F.).

AUTHOR CONTRIBUTIONS

A.T.R. designed, cloned, expressed, and established purification protocols for the preparation of recombinant complexes and performed the biochemical assays. A.T.R. assessed the purified complexes by electron microscopy of negatively stained samples. P.C.A.d.F. prepared the cryo-EM grids. A.T.R and P.C.A.d.F. collected cryo-EM image datasets. P.C.A.d.F. performed image analysis, and A.T.R. and P.C.A.d.F. performed atomic model building. A.T.R. and P.C.A.d.F. interpreted the results and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 28, 2019
Revised: May 27, 2019
Accepted: July 10, 2019
Published: August 28, 2019

REFERENCES

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. *68*, 352–367.

Bard, J.A.M., Goodall, E.A., Greene, E.R., Jonsson, E., Dong, K.C., and Martin, A. (2018). Structure and function of the 26S proteasome. Annu. Rev. Biochem. *87*, 697–724.

Bech-Otschir, D., Helfrich, A., Enenkel, C., Consiglieri, G., Seeger, M., Holzhütter, H.-G., Dahlmann, B., and Kloetzel, P.-M. (2009). Polyubiquitin substrates allosterically activate their own degradation by the 26S proteasome. Nat. Struct. Mol. Biol. *16*, 219–225.

Bhattacharyya, S., Yu, H., Mim, C., and Matouschek, A. (2014). Regulated protein turnover: snapshots of the proteasome in action. Nat. Rev. Mol. Cell Biol. *15*, 122–133.

Blickwedehl, J., Agarwal, M., Seong, C., Pandita, R.K., Melendy, T., Sung, P., Pandita, T.K., and Bangia, N. (2008). Role for proteasome activator PA200 and postglutamyl proteasome activity in genomic stability. Proc. Natl. Acad. Sci. USA *105*, 16165–16170.

Budenholzer, L., Cheng, C.L., Li, Y., and Hochstrasser, M. (2017). Proteasome structure and assembly. J. Mol. Biol. *429*, 3500–3524.

Cascio, P., Call, M., Petre, B.M., Walz, T., and Goldberg, A.L. (2002). Properties of the hybrid form of the 26S proteasome containing both 19S and PA28 complexes. EMBO J. *21*, 2636–2645.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Chen, S., Wu, J., Lu, Y., Ma, Y.B., Lee, B.H., Yu, Z., Ouyang, Q., Finley, D.J., Kirschner, M.W., and Mao, Y. (2016). Structural basis for dynamic regulation of the human 26S proteasome. Proc. Natl. Acad. Sci. USA *113*, 12991–12996.

Chu-Ping, M., Vu, J.H., Proske, R.J., Slaughter, C.A., and DeMartino, G.N. (1994). Identification, purification, and characterization of a high molecular weight, ATP-dependent activator (PA700) of the 20 S proteasome. J. Biol. Chem. *269*, 3539–3547.

Collins, G.A., and Goldberg, A.L. (2017). The logic of the 26S proteasome. Cell *169*, 792–806.

da Fonseca, P.C., and Morris, E.P. (2015). Cryo-EM reveals the conformation of a substrate analogue in the human 20S proteasome core. Nat. Commun. *6*, 7573.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Ettari, R., Previti, S., Bitto, A., Grasso, S., and Zappalà, M. (2016). Immunoproteasome-selective inhibitors: a promising strategy to treat hematologic malignancies, autoimmune and inflammatory diseases. Curr. Med. Chem. *23*, 1217–1238.

Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., and Berger, I. (2006). Protein complex expression by using multigene baculoviral vectors. Nat. Methods *3*, 1021–1032.

Förster, A., Whitby, F.G., and Hill, C.P. (2003). The pore of activated 20S proteasomes has an ordered 7-fold symmetric conformation. EMBO J. *22*, 4356–4364.

Förster, A., Masters, E.I., Whitby, F.G., Robinson, H., and Hill, C.P. (2005). The 1.9 A structure of a proteasome-11S activator complex and implications for proteasome-PAN/PA700 interactions. Mol. Cell *18*, 589–599.

García-Nafría, J., Watson, J.F., and Greger, I.H. (2016). IVA cloning: A single-tube universal cloning system exploiting bacterial *In Vivo* Assembly. Sci. Rep. *6*, 27459.

Grant, T., Rohou, A., and Grigorieff, N. (2018). *cis*TEM, user-friendly software for single-particle image processing. eLife *7*, e35383.

Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H.D., and Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 A resolution. Nature *386*, 463–471.

Groll, M., Bajorek, M., Köhler, A., Moroder, L., Rubin, D.M., Huber, R., Glickman, M.H., and Finley, D. (2000). A gated channel into the proteasome core particle. Nat. Struct. Biol. *7*, 1062–1067.

Haselbach, D., Schrader, J., Lambrecht, F., Henneberg, F., Chari, A., and Stark, H. (2017). Long-range allosteric regulation of the human 26S proteasome by 20S proteasome-targeting cancer drugs. Nat. Commun. *8*, 15578.

Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U., and Wolf, D.H. (1997). The active sites of the eukaryotic 20 S proteasome and their involvement in subunit precursor processing. J. Biol. Chem. *272*, 25200–25209.

Hochstrasser, M. (2016). Gyre and gimble in the proteasome. Proc. Natl. Acad. Sci. USA *113*, 12896–12898.

Hoffman, L., Pratt, G., and Rechsteiner, M. (1992). Multiple forms of the 20 S multicatalytic and the 26 S ubiquitin/ATP-dependent proteases from rabbit reticulocyte lysate. J. Biol. Chem. *267*, 22362–22368.

Huber, E.M., and Groll, M. (2017). The mammalian proteasome activator PA28 forms an asymmetric $\alpha4\beta3$ complex. Structure *25*, 1473–1480.e3.

Huber, E.M., Heinemeyer, W., de Bruin, G., Overkleeft, H.S., and Groll, M. (2016). A humanized yeast proteasome identifies unique binding modes of inhibitors for the immunosubunit $\beta5$i. EMBO J. *35*, 2602–2613.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

Khare, S., Nagle, A.S., Biggart, A., Lai, Y.H., Liang, F., Davis, L.C., Barnes, S.W., Mathison, C.J., Myburgh, E., Gao, M.Y., et al. (2016). Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. Nature *537*, 229–233.

Khor, B., Bredemeyer, A.L., Huang, C.Y., Turnbull, I.R., Evans, R., Maggi, L.B., Jr., White, J.M., Walker, L.M., Carnes, K., Hess, R.A., and Sleckman, B.P. (2006). Proteasome activator PA200 is required for normal spermatogenesis. Mol. Cell. Biol. *26*, 2999–3007.

Kleijnen, M.F., Roelofs, J., Park, S., Hathaway, N.A., Glickman, M., King, R.W., and Finley, D. (2007). Stability of the proteasome can be regulated allosterically through engagement of its proteolytic active sites. Nat. Struct. Mol. Biol. *14*, 1180–1188.

Kniepert, A., and Groettrup, M. (2014). The unique functions of tissue-specific proteasomes. Trends Biochem. Sci. *39*, 17–24.

Knowlton, J.R., Johnston, S.C., Whitby, F.G., Realini, C., Zhang, Z., Rechsteiner, M., and Hill, C.P. (1997). Structure of the proteasome activator REGalpha (PA28alpha). Nature *390*, 639–643.

Kopp, F., Dahlmann, B., and Kuehn, L. (2001). Reconstitution of hybrid proteasomes from purified PA700-20 S complexes and PA28alphabeta activator: ultrastructure and peptidase activities. J. Mol. Biol. *313*, 465–471.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nat. Methods *11*, 63–65.

Laskowski, R.A., and Swindells, M.B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model. *51*, 2778–2786.

Li, X., Kusmierczyk, A.R., Wong, P., Emili, A., and Hochstrasser, M. (2007). β-Subunit appendages promote 20S proteasome assembly by overcoming an Ump1-dependent checkpoint. EMBO J. *26*, 2339–2349.

Li, H., O'Donoghue, A.J., van der Linden, W.A., Xie, S.C., Yoo, E., Foe, I.T., Tilley, L., Craik, C.S., da Fonseca, P.C., and Bogyo, M. (2016). Structure- and function-based design of *Plasmodium*-selective proteasome inhibitors. Nature *530*, 233–236.

Livneh, I., Cohen-Kaplan, V., Cohen-Rosenzweig, C., Avni, N., and Ciechanover, A. (2016). The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. Cell Res. *26*, 869–885.

Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 A resolution. Science *268*, 533–539.

Ma, C.P., Slaughter, C.A., and DeMartino, G.N. (1992). Identification, purification, and characterization of a protein activator (PA28) of the 20 S proteasome (macropain). J. Biol. Chem. *267*, 10515–10523.

Mandemaker, I.K., Geijer, M.E., Kik, I., Bezstarosti, K., Rijkers, E., Raams, A., Janssens, R.C., Lans, H., Hoeijmakers, J.H., Demmers, J.A., et al. (2018). DNA damage-induced replication stress results in PA200-proteasome-mediated degradation of acetylated histones. EMBO Rep. *19*, e45566.

Mao, I., Liu, J., Li, X., and Luo, H. (2008). REGgamma, a proteasome activator and beyond? Cell. Mol. Life Sci. *65*, 3971–3980.

Marques, A.J., Glanemann, C., Ramos, P.C., and Dohmen, R.J. (2007). The C-terminal extension of the $\beta7$ subunit and activator complexes stabilize nascent 20 S proteasomes and promote their maturation. J. Biol. Chem. *282*, 34869–34876.

Millard, C.J., Watson, P.J., Celardo, I., Gordiyenko, Y., Cowley, S.M., Robinson, C.V., Fairall, L., and Schwabe, J.W. (2013). Class I HDACs share a common mechanism of regulation by inositol phosphates. Mol. Cell *51*, 57–67.

Morris, E.P., and da Fonseca, P.C.A. (2017). High-resolution cryo-EM proteasome structures in drug development. Acta Crystallogr. D Struct. Biol. *73*, 522–533.

Murata, S., Takahama, Y., Kasahara, M., and Tanaka, K. (2018). The immunoproteasome and thymoproteasome: functions, evolution and human disease. Nat. Immunol. *19*, 923–931.

Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta. Crystallogr. D. Biol. Crystallogr. *68*, 352–367.

O'Reilly, D.R., Miller, L.K., and Luckow, V.A. (1994). Baculovirus Expression Vectors: A Laboratory Manual (Oxford University Press).

Ortega, J., Heymann, J.B., Kajava, A.V., Ustrell, V., Rechsteiner, M., and Steven, A.C. (2005). The axial channel of the 20S proteasome opens upon binding of the PA200 activator. J. Mol. Biol. *346*, 1221–1227.

Park, J.E., Miller, Z., Jun, Y., Lee, W., and Kim, K.B. (2018). Next-generation proteasome inhibitors for cancer therapy. Transl. Res. *198*, 1–16.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Qian, M.X., Pang, Y., Liu, C.H., Haratake, K., Du, B.Y., Ji, D.Y., Wang, G.F., Zhu, Q.Q., Song, W., Yu, Y., et al. (2013). Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. Cell *153*, 1012–1024.

Ruschak, A.M., and Kay, L.E. (2012). Proteasome allostery as a population shift between interchanging conformers. Proc. Natl. Acad. Sci. USA *109*, E3454–E3462.

Sadre-Bazzaz, K., Whitby, F.G., Robinson, H., Formosa, T., and Hill, C.P. (2010). Structure of a Blm10 complex reveals common mechanisms for proteasome binding and gate opening. Mol. Cell *37*, 728–735.

Schrader, J., Henneberg, F., Mata, R.A., Tittmann, K., Schneider, T.R., Stark, H., Bourenkov, G., and Chari, A. (2016). The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. Science *353*, 594–598.

Ustrell, V., Hoffman, L., Pratt, G., and Rechsteiner, M. (2002). PA200, a nuclear proteasome activator involved in DNA repair. EMBO J. *21*, 3516–3525.

Watson, P.J., Fairall, L., Santos, G.M., and Schwabe, J.W. (2012). Structure of HDAC3 bound to co-repressor and inositol tetraphosphate. Nature *481*, 335–340.

Watson, P.J., Millard, C.J., Riley, A.M., Robertson, N.S., Wright, L.C., Godage, H.Y., Cowley, S.M., Jamieson, A.G., Potter, B.V., and Schwabe, J.W. (2016). Insights into the activation mechanism of class I HDAC complexes by inositol phosphates. Nat. Commun. *7*, 11262.

Wehmer, M., and Sakata, E. (2016). Recent advances in the structural biology of the 26S proteasome. Int. J. Biochem. Cell Biol. *79*, 437–442.

Wehmer, M., Rudack, T., Beck, F., Aufderheide, A., Pfeifer, G., Plitzko, J.M., Förster, F., Schulten, K., Baumeister, W., and Sakata, E. (2017). Structural insights into the functional cycle of the ATPase module of the 26S proteasome. Proc. Natl. Acad. Sci. USA *114*, 1305–1310.

Whitby, F.G., Masters, E.I., Kramer, L., Knowlton, J.R., Yao, Y., Wang, C.C., and Hill, C.P. (2000). Structural basis for the activation of 20S proteasomes by 11S regulators. Nature *408*, 115–120.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. Nat. Methods *12*, 7–8.

Zwickl, P., Lottspeich, F., and Baumeister, W. (1992). Expression of functional *Thermoplasma acidophilum* proteasomes in *Escherichia coli*. FEBS Lett. *312*, 157–160.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
DH10EmbacY cells	Geneva Biotech	N/A
Pir1 competent cells	ThermoFisher scientific	*Cat# C101010*
Top10 competent cells	ThermoFisher scientific	*Cat# C404003*
Chemicals, Peptides, and Recombinant Proteins		
FuGENE® HD Transfection Reagent	Promega	Cat# E2311
Human 20S Proteasome	Enzo	Cat# BML-PW8720-0050
Superose 6 increase 10/300 GL	GE Healthcare	Cat# 29-0915-96
Streptactin Superflow plus 5ml cartridge	QIAGEN	Cat# 30060
Suc-LLVY-AMC	Boston Biochem	Cat# S-280
Boc-LRR-AMC	Boston Biochem	Cat# S-300
Z-LLE-AMC	Boston Biochem	Cat# S-230
Tris-HCl	Sigma-Aldrich	N/A
Glycerol	VWR	Cat# 56-81-5
NaCl	Sigma-Aldrich	
TEV protease	N/A	N/A
DTT	Thermo Scientific	Cat# R0862
EDTA	Sigma-Aldrich	
Cre recombinase	NEB	Cat# M0298S
Qiaprep Spin Miniprep kit reagents	QIAGEN	Cat# 27106
InsectXPRESS media	Lonza	Cat# BE12-730Q
FBS	GIBCO	*Cat# 10270106*
30 kDa cut-off Vivaspin 20 concentrators		
Superdex 200 Increase 10/300 GL	GE Healthcare	Cat# 28-9909-44
HEPES	Sigma-Aldrich	Cat# H3375
DMSO	Sigma-Aldrich	Cat# BP231-100
Deposited Data		
r20S EM map	This study	EMD-4877
r20S atomic coordinates	This study	PDB: 6RGQ
r20S-PA200 EM map	This study	EMD-4860
r20S-PA200 atomic coordinates	This study	PDB: 6REY
Experimental Models: Cell Lines		
Sf9 insect cells	N/A	N/A
Recombinant DNA		
cDNAs for all subunits (Gene synthesis)	Epoch	N/A
pACEBAc1	Geneva Biotech	N/A
pIDC	Geneva Biotech	N/A
pIDS	Geneva Biotech	N/A
Software and Algorithms		
EPU Automated Data Acquisition Software for Single Particle Analysis v1.9.1	ThermoFisher	https://www.fei.com/software/epu/
Prism 8 software	GraphPad	https://www.graphpad.com/scientific-software/prism/
ResMap1.1.4	Kucukelbir et al., 2014	http://resmap.sourceforge.net

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cisTEM	Grant et al., 2018	RRID:SCR_016502; https://cistem.org
Coot v0.8.9	Emsley et al., 2010	RRID:SCR_014222; https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
PHENIX v1.13	Mustyakimov et al., 2012	RRID:SCR_014224; http://www.phenix-online.org
UCSF Chimera	Pettersen et al., 2004	RRID:SCR_004097 ; https://www.cgl.ucsf.edu/chimera/
Phyre2	Kelley et al., 2015	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index
MolProbity	Chen et al., 2010	RRID:SCR_014226 ; http://molprobity.biochem.duke.edu/index.php?MolProbSID=s5mqsj74op6r8o3u1j5jkeasr2&eventID=2
Pymol Molecular Graphics System	Schrödinger, LLC	RRID:SCR_000305; https://pymol.org/2/
LigPlot+	Laskowski and Swindells, 2011	https://www.ebi.ac.uk/thornton-srv/software/LigPlus/
I-Tasser	Yang et al., 2015	RRID:SCR_014627 ; https://zhanglab.ccmb.med.umich.edu/I-TASSER/
Other		
QUANTIFOIL R1.2/1.3 grids 300 mesh	Quantifoil Micro Tools	R1.2/1.3

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Paula da Fonseca (pauladf@mrc-lmb.cam.ac.uk).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Lines

Sf9 insect cells were cultivated in InsectXPRESS media (Lonza), at 27°C.

Bacterial Strains

Constructs were transformed and amplified in either *Escherichia coli* TOP10 or Pir1 chemically competent cells (Invitrogen).

The DH10EMBacY *Escherichia coli* strain (Geneva Biotech) containing a baculovirus shuttle vector (bacmid) and a constitutively expressing YFP expression cassette was used to create all expression bacmids containing the cloned genes of interest.

METHOD DETAILS

Cloning and generation of baculovirus

cDNAs encoding each of the human 20S proteasome subunits, proteasome assembly chaperones and PA200 were synthesized as codon-optimized genes for expression in *E. coli* (Epoch). These genes were subsequently cloned in Multibac Turbo vectors pACEBac1, pIDS and/or pIDC using a strategy adapted from the IVA method (García-Nafría et al., 2016). All proteasome α subunits were cloned into a pACEBac1 vector, the β subunits into a pIDS, the chaperones in both pIDC and pACEBac1, and PA200 in pACEBac1. The proteasome β subunits were cloned with their N-terminal pre-peptides, which are cleaved at the final stages of proteasome assembly and maturation resulting in proteolytic active complexes (Budenholzer et al., 2017). We added a TwinStrep-tag at the C terminus of β7 to assist on the purification of mature homogeneous proteasome complexes, as this is known to be the last subunit to be incorporated during the proteasome assembly pathway (Li et al., 2007; Marques et al., 2007). We also created untagged and N-terminal Twin-Strep tagged versions of PA200, to be co-expressed with the 20S proteasome or expressed on its own, respectively. The pACEBac1 vector, containing the α subunits, and the pIDC vector, containing the β subunits, were Cre-Loxed (Fitzgerald et al., 2006) to create a single vector (20S proteasome plasmid). Attempts were made to Cre-Lox the chaperones containing pIDC vector to either the protesome α or β subunit vectors, but due to the size of the plasmids, the high number of genes and gene repetitions many random recombinant events occur, and for that reason we decided to clone the proteasome assembly chaperones in a separate pACEBac1 vector (Figure S1A).

All assembled plasmids were transposed to DH10EmBacY cells (Geneva Biotech), and Bacmid DNA purified following published protocols (O'Reilly et al., 1994). Bacmids were transfected (P1 virus), after which Sf9 cells, at 1.5x10^6 cells/ml in InsectXPRESS media (Lonza), were infected and incubated for approximately 72 hours until viability was around 80%. The P2 virus was then filtered and 2% Hi-FBS serum added. This was kept as a stock at 4°C, protected from light.

Expression of proteasome complexes

Each individual baculovirus was initially amplified (P3 amplification) by infecting Sf9 cells at 1.5 x10^6 cells/ml with 1:100 of P2 virus, followed by a 72 hour incubation at 27°C with shaking, or until cell viability was about 80%. The P3 virus were then filtered and used to co-infect Sf9 cells with a density of roughly 2x10^6 cells/ml (12 mL of each P3 virus for each 500 mL of cells). Expression of the human 20S proteasome was accomplished by co-infection of Sf9 cells with two baculoviruses, one containing the 20S proteasome the other the assembly chaperones. The 20S-PA200 complex was expressed by triple co-infection with baculoviruses containing (1) the 20S proteasome, (2) the chaperones and (3) the PA200. Interestingly, we observed that human 20S proteasomes can assemble when expressing only its 14 subunits, but with more than a 3-fold decrease in yield of mature complexes compared with that obtained with chaperone co-expression (data not shown). This suggests that endogenous insect proteasome assembly chaperones can, at least partially, assist in the assembly of the human complex. After infection, the Sf9 cells were incubated at 27°C for 48 hours and harvested by centrifugation at 3,000 *g*, 20 min, 4°C. Cell pellets were washed with cold PBS, frozen in LN2 and stored at −80°C.

Purification proteasome complexes

The purification strategy was the same for both reconstituted 20S proteasome and 20S-PA200 complexes. Briefly, thawed cells were resuspended in 3-4 times volume of buffer W (50 mM Tris pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, 1 mM EDTA), lysed by sonication and the lysate cleared by centrifugation at 48,400 *g* for 30 minutes. The clear lysate was then passed through a 5 μm filter and loaded onto tandem Streptactin Superflow Plus columns (QIAGEN), equilibrated in buffer W, and eluted in buffer E (50 mM Tris pH7.5, 150 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, 1mM EDTA, 2.5 mM d-desthiobiotin). Proteasome containing fractions were pooled together and the TwinStrep-tag cleaved by overnight dialysis against buffer W containing TEV protease, at a ratio of 1:50 (TEV to protein). The sample was filtered again with a 0.4 μm filter and loaded onto Streptactin Superflow Plus columns, where the flow-through was collected. The protein was concentrated using 30 kDa cut-off Vivaspin 20 concentrators (Sartorius) and loaded onto a Superose 6 Increase 10/300 gel filtration column (GE Healthcare) equilibrated with 50 mM Tris pH 7.4, 100 mM NaCl, 1 mM EDTA (Figure S1B). Typical yields for the purification of reconstituted 20S proteasomes and 20S-PA200 complexes are 2-3 mg and 1 mg of protein purified from 1 L of Sf9 cultures, respectively.

Expression and purification of PA200

Expression of the human PA200 was accomplished by infection of Sf9 cells with the TwinStrep-Tev-PA200 containing baculovirus. The purification protocol was the same as described for the recombinant 20S proteasome, except that the last size exclusion chromatography step was performed in a Superdex 200 Increase 10/300 (GE Healthcare).

Proteasome activity assay

20S proteasome proteolytic activities were measured by fluorescence spectroscopy using the β1, β2 and β5 specific substrates (from Boston Biochem) carboxybenzyl-Leu-Leu-Glu-7-amino-4-methylcoumarin (Z-LLE-AMC), tert-butyloxycarbonyl-Leu-Arg-Arg-7-amino-4-methylcoumarin (Boc-LRR-AMC) and N-succinyl-Leu-Leu-Val-Tyr-7-amino-4-methylcoumarin (Suc-LLVY-AMC), respectively.

Recombinant human 20S proteasomes, endogenous human 20S proteasomes (Enzo) and recombinant human 20S-PA200 complexes, all at 7.5 nM in 25 mM HEPES, pH 7.5 and 0.5 mM EDTA, were incubated with 50 μM substrate (stock at 5 mM in DMSO) for 30 minutes at room temperature. Fluorescence intensity, generated by the release of AMC fluorophores, was measured in triplicate with a Pherastar (BMG Labtech) with $\lambda_{excitation}$ = 350 nm and $\lambda_{emission}$ = 450 nm.

PA200 titration experiments were done by mixing 60 nM of either endogenous or recombinant 20S proteasomes with increasing concentrations of recombinant PA200 (0 nM, 15 nM, 20 nM, 30 nM, 60 nM, 120 nM, 180 nM and 240 nM), followed by incubation for 1 hour at room temperature. The samples were then diluted 20 times and incubated with 50 μM of each of the three proteasome fluorogenic substrates. Measurements were done in Pherastar (BMG Labtech) after incubation at room temperature for 30 minutes. All experiments were done in triplicate.

Cryo-electron microscopy

Quantifoil R1.2/1.3 electron microscope grids, with a 300 gold mesh, were coated with a thin layer of carbon freshly floated from mica, following the procedures we previously optimized for the preparation of grids with endogenous human proteasomes (da Fonseca and Morris, 2015; Morris and da Fonseca, 2017). We find that the continuous carbon layer favors an even particle distribution, while its electron scattering facilitates the assessment of the information in the recorded images, as judged by the recovery of Thon rings to high resolution in their power spectra, which also contributes to an accurate defocus estimation for the correction

of the contrast transfer function associated effects. The grids used for imaging the recombinant human 20S proteasome were glow-discharged in the presence of pentylamine, as we previously described for the endogenous complex (da Fonseca and Morris, 2015; Morris and da Fonseca, 2017). The grids used for the 20S-PA200 complex were glow-discharged in atmospheric air. 3 μL of sample were loaded on the grids and flash frozen by plunging the grids into liquid ethane using a Vitrobot Mark IV (FEI), operated at 22°C, 95% humidity, 20 s waiting and 5 s blotting times. The grids were transferred into a FEI Titan Krios electron microscope, operated at 300 keV and a nominal magnification of x95,000, resulting in a calibrated sampling of 0.81 Å per pixel at the image level. Images were recorded with EPU software using a Falcon III direct electron detector operating in counting mode, at an electron dose per pixel of ~0.5 e^-/s, with 60 s exposures saved as 75 frame movies with evenly distributed electron dose (Figures S2A and S2B).

Single particle analysis

The cryo-EM images were processed using the *cis*TEM software (Grant et al., 2018). For the analysis of the recombinant human 20S proteasome, the frames of each of the 483 movies recorded were aligned and summed into single images. Upon inspection, 411 images were selected for further analysis, based on the ice quality, image contrast and the recovery of isotropic Thon rings to high resolution. The selected images had a defocus range from −0.8 μm to −3.4 μm. A total of 58,281 particles were picked from the selected images and subsequently classified. The resulting 2D class averages reveal a very homogeneous sample and were used to exclude particle picking false positives. No proteasome top-views, corresponding to projections along the proteasome long axis, were selected since the proteasome side-views show a tomographic distribution of Euler angles, evenly distributed around one of the main axis, providing a complete and even information recovery in Fourier space. The resulting selection of 50,885 particles were used for Auto Refine, with C2 symmetry imposed, using our previous 3.5 Å cryo-EM map of the endogenous human 20S proteasome (EMD-2981) (da Fonseca and Morris, 2015) as starting reference. The overall image processing workflow is shown in Figure S7A. The resolution of the final map was estimated at 2.6 Å by Fourier shell correlation, as implemented in cisTEM (Figure S2C).

The analysis of the cryo-EM images of the recombinant human 20S-PA200 followed the same procedures as described for the recombinant 20S proteasome. From the 1,114 movies recorded, 717 were selected for further analysis and had a defocus range from −1.2 μm to −2.9 μm. 29,434 particles were picked from the selected images and classified. The resulting class averages revealed the presence of some uncapped 20S proteasomes, which were excluded from the particle data-set together with particle picking false positives. The 17,356 particles from the selected class-averages were subjected to another round of 2D classification into 10 classes that showed no evidence for significant heterogeneity. Therefore, all 17,356 particles were used for 3D Auto Refine, with C2 symmetry imposed, using our cryo-EM map of the recombinant human 20S proteasome as starting reference, followed by a round of local refinement. An additional 3D classification into 2 classes, using C1 symmetry, indicated that 28% of the particles in the final map are from single capped complexes. The overall image processing workflow is shown in Figure S7B.The resolution of the final map was estimated at 3.0 Å by Fourier shell correlation, as implemented in cisTEM (Figure S2D).

Molecular modeling

The model of the human recombinant 20S proteasome was built based on the X-ray crystal structure of the endogenous human 20S proteasome (PDB: 5LE5) (Schrader et al., 2016) using real-space refinement in Coot (Emsley et al., 2010) and Phenix (Afonine et al., 2012). The model of the recombinant human 20S-PA200 complex was built based on our cryo-EM model of the recombinant human 20S proteasome and models for the human PA200 derived from Phyre2 (Kelley et al., 2015) and I-Tasser (Yang et al., 2015). These PA200 models were used as initial guides for the assignment of secondary structure using real-space refinement in Coot (Emsley et al., 2010) and Phenix (Afonine et al., 2012), followed by correction of the sequence register using the amino acid side chain densities clearly resolved in our cryo-EM map. Connecting loops and all other regions were built *ab initio* from the cryo-EM density. The final atomic models of the recombinant human 20S proteasome and 20S-PA200 complexes were validated using MolProbity (Chen et al., 2010) (Figures S2G and S2H).

Graphic representations

The representations of our structures as shown in Figures 2A, 2D, S2E, and S2F were created using UCSF Chimera (Pettersen et al., 2004), while all the remaining structure representations in the manuscript were created using the Pymol Molecular Graphics System, Schrödinger, LLC. The cryo-EM maps shown in Figures S2E and S2F were colored according to local resolution as estimated using ResMap (Kucukelbir et al., 2014). The diagrams of the protein-protein interaction networks shown in Figure S4 were created with LigPlot+ (Laskowski and Swindells, 2011).

QUANTIFICATION AND STATISTICAL ANALYSIS

Cell density was measured in a Countess II (Life technology).

Protein concentrations were measured in a Nanodrop 2000c (Life technologies).

Proteasome activity was measured in Pherastar a (BMG Labtech) with $\lambda_{excitation}$ = 350 nm and $\lambda_{emission}$ = 450 nm. Data analysis, including the calculation of Error bars as mean with standard deviation derived from the triplicate measurements represented in Figures 1 and S1, was done in Prism 8 software (GraphPad).

DATA AND CODE AVAILABILITY

The cryo-EM density maps of 20S and 20S-PA200 have been deposited into the EMDataBank with accession codes EMD-4877 and EMD-4860, respectively. The corresponding atomic coordinates have been deposited in the Protein Data Bank with accession codes PDB: 6RGQ and PDB: 6REY.