

Structural basis of human 20S proteasome biogenesis

Received: 21 September 2023

Hanxiao Zhang , Chenyu Zhou, Zarith Mohammad & Jianhua Zhao  

Accepted: 11 September 2024

Published online: 18 September 2024

 Check for updates

New proteasomes are produced to accommodate increases in cellular catabolic demand and prevent the accumulation of cytotoxic proteins. Formation of the proteasomal 20S core complex relies on the function of the five chaperones PAC1-4 and POMP. Here, to understand how these chaperones facilitate proteasome assembly, we tagged the endogenous chaperones using CRISPR/Cas gene editing and examined the chaperone-bound complexes by cryo-EM. We observe an early α -ring intermediate subcomplex that is stabilized by PAC1-4, which transitions to β -ring assembly upon dissociation of PAC3/PAC4 and rearrangement of the PAC1 N-terminal tail. Completion of the β -ring and dimerization of half-proteasomes repositions critical lysine K33 to trigger cleavage of the β pro-peptides, leading to the concerted dissociation of POMP and PAC1/PAC2 to yield mature 20S proteasomes. This study reveals structural insights into critical points along the assembly pathway of the human proteasome and provides a molecular blueprint for 20S biogenesis.

The proteasome is the primary protease that is responsible for degrading the majority of proteins in eukaryotic cells and plays a key role in metabolism and protein quality control mechanisms^{1–8}. Proteasome function is crucial for maintaining nutrient balance and removing cytotoxic polypeptides that can cause cellular stress and dysfunction. Consequently, decreases in proteasome levels and activity are associated with aging and neurodegenerative diseases such as Alzheimer's disease and Parkinson's disease^{9–12}. To maintain homeostasis, cells must replenish existing proteasomes as well as produce new proteasomes in response to elevated levels of damaged and misfolded proteins^{5,6,13,14}.

The 20S proteasome is the core complex that degrades proteins into short peptides. It consists of four heptameric rings of α and β subunits arranged in a $\alpha(1-7); \beta(1-7); \beta(1-7); \alpha(1-7)$ configuration. The α subunits form gates at both ends of the barrel-shaped complex that control substrate entry, while the catalytic β subunits cleave substrates inside the complex. There are three different catalytic β subunits in eukaryotic 20S proteasomes: $\beta 1$ has caspase-like activity, $\beta 2$ has trypsin-like activity, and $\beta 5$ has chymotrypsin-like activity¹⁵. The β subunits are expressed with pro-peptides that protect the critical N-terminal threonine of the catalytic β subunits¹⁶. These pro-peptides also aid in 20S assembly and are cleaved in the mature 20S complex^{16–20}.

The production of new proteasomes is an intricate process that requires the precise assembly of many different components and involves the action of multiple chaperones. Early in the biogenesis pathway, proteasome assembly chaperones (PAC) 1-4 coordinate the arrangement of 7 different α subunits ($\alpha 1-7$) into a hetero-heptameric α -ring^{21–23}. Biochemical studies indicate that the PAC1/PAC2 and PAC3/PAC4 heterodimers bind initially to $\alpha 4-7$, which provides a subcomplex that seeds the addition of $\alpha 1-3$. Structural studies have provided important insights into how PAC1/PAC2 engages with the α subunits^{24–26}, which both contain C-terminal HbYX motifs that help to anchor PAC1/PAC2 to the α -ring^{27,28}. However, there is a lack of structural data on how PAC3/PAC4 engages with the α subunits to facilitate α -ring formation. Furthermore, it has been postulated that the N-terminal tail of PAC1 may be involved in α -ring formation²⁴, but how this may play a role remains unclear.

Once the α -ring is formed, the proteasome maturation protein (POMP) chaperone assembles the 7 different β subunits onto the α -ring^{29–32} with concomitant unbinding of the PAC3/PAC4 heterodimer. While PAC3/PAC4 helps to assemble the α -ring, they block the completion of the β -ring. Binding of POMP and dissociation of PAC3/PAC4 marks the transition from α -ring formation to β -ring formation, but how this transition occurs remains unclear. The different β subunits

are added in a stepwise fashion, with the addition of $\beta 7$ incorporating last to complete the β -ring. The α -ring/ β -ring heterocomplex (half-proteasome) dimerizes to form the pre-20S complex, which activates the autocatalytic cleavage of the β pro-peptides^{17,18}. How dimerization of the half-proteasomes induces cleavage of the β pro-peptides remains unclear. Finally, POMP and the PAC1/PAC2 heterodimer dissociate to yield the mature 20S proteasome^{16,21,29}. The sequential order of these final steps and their structural basis remains largely unexplored.

Here, we provide a broad structural analysis of 20S biogenesis, which clarifies the role of each chaperone across the assembly process and reveals molecular insights into PAC3/PAC4 function and late-stage maturation of the pre-20S complex. The complexity of proteasome assembly has made it challenging to replicate the process *in vitro*. Knock-down and mutation of some subunits in yeast and mammalian cells have been useful strategies in trapping certain proteasomal intermediates for biochemical and structural analyses^{16,24,29,33}. However, decreases in cellular fitness and changes in complex stoichiometry associated with traditional approaches have complicated the in-depth structural characterization of the proteasome biogenesis pathway^{21,22,34–37}. We have previously developed efficient strategies of tagging and purifying endogenous proteins for cryo-EM analysis^{38,39}, which we have employed in this study to dissect the stepwise assembly pathway of the human 20S proteasome.

Results

Overall assembly of the 20S proteasome

Assembly of the proteasomal 20S core complex is orchestrated by five dedicated chaperones. To study their function, we individually tagged three of the five 20S chaperones in Expi293F cells using CRISPR/Cas gene editing^{38,40} and separately purified the endogenous chaperone-

associated complexes. Analysis of the samples by single particle cryogenic electron microscopy (cryo-EM) enabled structure determination of intermediate complexes at each major stage of 20S assembly (Fig. 1a, Supplementary Figs. 1–3, Supplementary Table 1). Overall, each α and β subunit interacts with at least one of the 20S chaperones except for $\beta 7$ (Fig. 1b). Evidence of protein dynamics can be observed throughout the assembly process. In particular, the α subunits undergo substantial movements that appear to correlate with binding of the β subunits (Fig. 1c). On the other hand, the chaperones mainly undergo local structural rearrangements with limited changes in overall architecture (Fig. 1d). In the following, we describe mechanistic insights into α -ring formation, the transition to β -ring assembly, and dimerization and maturation into 20S complexes.

α -ring formation mediated by PAC1-4

To gain molecular insights into the function of PAC3/PAC4 in α -ring formation, we tagged PAC3 (PSMG3) with the Twinstrep tag in Expi293F cells and purified the PAC3-associated complexes (Supplementary Fig. 1e, f). Cryo-EM analysis of the sample resulted in a structure of the α -ring complex bound by PAC1-4 at ~ 3.0 Å resolution (Fig. 2a, Supplementary Figs. 2a, 3, and 10). While the PAC1/PAC2 and PAC3/PAC4 heterodimers bind on different sides of the α -ring, these two heterodimers interact with overlapping sets of α subunits. $\alpha 5$ and $\alpha 6$ make the most extensive contacts with PAC1-4, with the N termini of $\alpha 5$ and $\alpha 6$ inserting into a hydrophobic crevice formed between PAC1 and PAC2 (Supplementary Fig. 4a), which was also observed in yeast^{24,25}. Meanwhile, PAC3 and PAC4 are wedged into one side of the α -ring next to $\alpha 5/\alpha 6$ and clamp onto one of the α -helices in $\alpha 5$, explaining the tight binding of PAC3 and PAC4 to $\alpha 5$ (Fig. 2a, Supplementary Fig. 4b)^{35,41,42}. PAC3 also interacts with $\alpha 6$ (Supplementary Fig. 4c). In addition, the C terminus of PAC1 inserts into the HbYX-

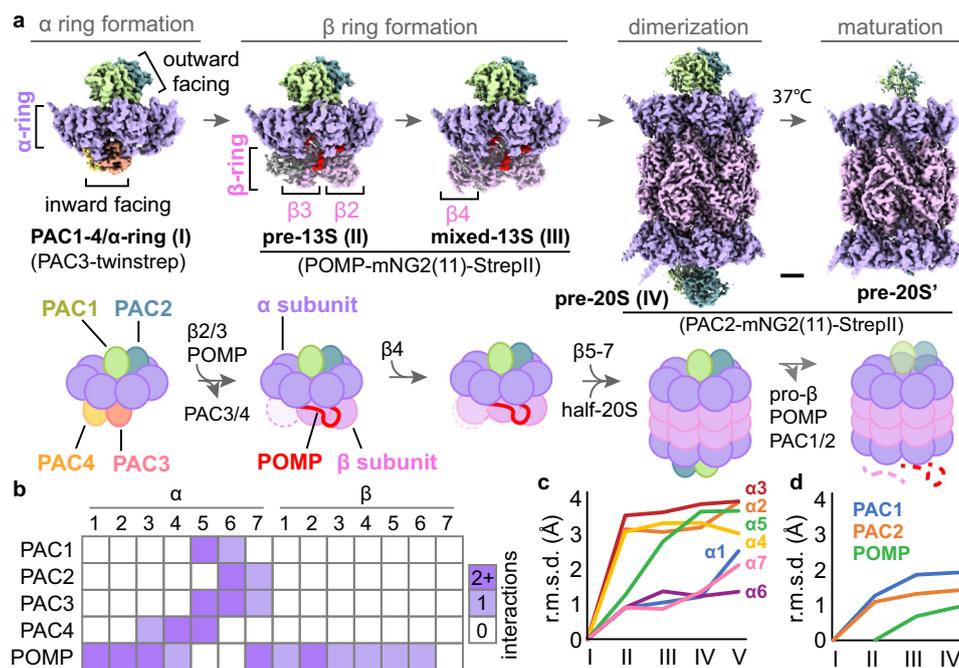


Fig. 1 | Stepwise assembly of the 20S proteasome is mediated by multiple chaperone interactions. **a** Cryo-EM structures of human 20S intermediate complexes depicting the stepwise assembly and maturation of the proteasome. Assembly begins with formation of the α -ring, aided by PAC1-4. The β -ring is formed on the α -ring, aided by POMP. Dimerization of α -ring/ β -ring half-proteasomes triggers cleavage of the β pro-peptides (pro- β) and dissociation of POMP and PAC1/PAC2 to yield the mature 20S proteasome. Scale bar: 20 Å. **b** Interaction map between proteasome chaperones and 20S subunits. Each chaperone makes

extensive contacts with multiple 20S subunits, with PAC1-4 interacting primarily with the $\alpha 4$ -7 subunits while POMP interactions are spread out across many α and β subunits. **c** Root mean square deviation (r.m.s.d.) values of the α subunit atomic positions across the PAC1-4/ α -ring (I), pre-13S (II), mixed-13S (III), pre-20S (IV), and mature 20S (V, PDB 7NAN) structures. **d** R.m.s.d. values of proteasome chaperone atomic positions across the PAC1-4/ α -ring (I), pre-13S (II), mixed-13S (III), and pre-20S (IV) structures.

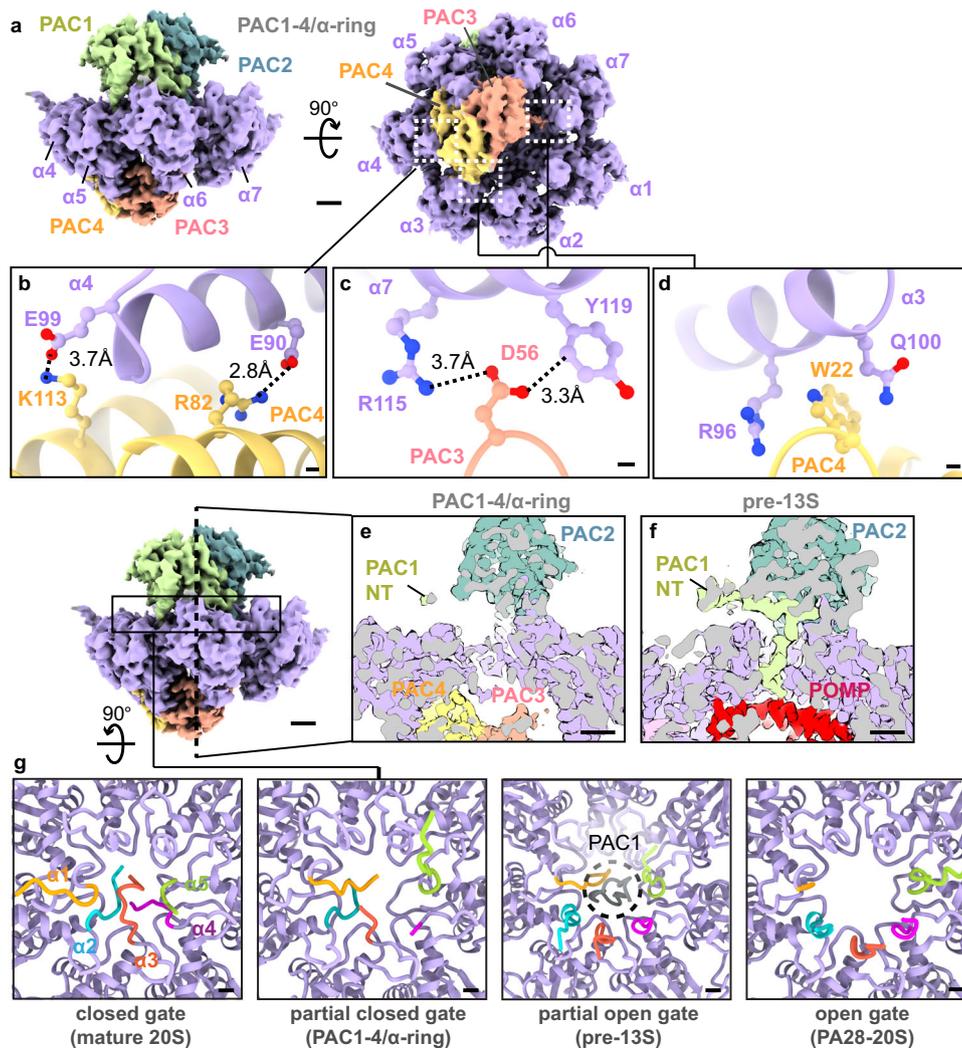


Fig. 2 | PAC1-4 helps to stabilize the α 4-7 subcomplex to initiate proteasome assembly. **a** Cryo-EM structure of the PAC1-4/ α -ring complex showing the distinct binding sites of PAC1/PAC2 and PAC3/PAC4. PAC1/PAC2 bind to α 5-7 on the outward facing side of the α -ring while PAC3/PAC4 bind to α 4-6 on the concave inward facing side of the α -ring. Scale bar: 15 Å. **b** PAC4 K113 and R82 form salt bridges with α 4 E99 and E90, respectively. Scale bar: 5 Å. **c** PAC3 D56 forms a salt bridge with α 7 R115 and coordinates with α 7 Y119. Scale bar: 5 Å. **d** PAC4 W22 fits into a pocket formed by α 3 R96 and Q100. Scale bar: 5 Å. **e** Cross-section of PAC1-4/ α -ring map

shows that the N-terminal tail of PAC1 is not inserted into the α -ring pore. Scale bar: 15 Å. **f** Cross-section of pre-13S assembly intermediate shows that the N-terminal tail of PAC1 is inserted into the α -ring pore in later stages of 20S assembly. Scale bar: 15 Å. **g** Positions of the α subunit N-terminal tails in a closed gate (mature 20S, PDB 7NAN), partially closed gate (PAC1-4/ α -ring), partially open gate (pre-13S), and fully open gate (PA28-20S, PDB 7NAO). The N-terminal tail of PAC1 is observed in the pre-13S structure but not in the PAC1-4/ α -ring structure.

binding pocket formed between α 5 and α 6 (Supplementary Fig. 4d)^{27,28}. The extensive structural contacts between PAC1-4 and α 5-6 suggest that these components may form an initial subcomplex that catalyzes α -ring formation.

To gain further insights into α -ring formation, we carried out computational 3D classification analysis of the PAC1-4/ α -ring cryo-EM data. The analysis revealed two 3D subclasses with good density for PAC1/2 (Supplementary Fig. 2a, I and IV) and two subclasses with low-quality density for PAC1/2 (Supplementary Fig. 2a, II and III), indicating that PAC1/2 may be dissociating in these subclasses. All subclasses appear to contain PAC1/2, PAC3/4, and α 1-7. Interestingly, we observed a 3D subclass with strong map densities for α 4-7 and lower map densities for α 1-3 (Supplementary Fig. 2a, I). In contrast, similar analysis of β -ring assembly intermediates revealed consistent α subunit map densities (Supplementary Fig. 5a). This result indicates that PAC1-4/ α 4-7 likely forms a stable subcomplex, consistent with previous biochemical studies⁴³. PAC1-4/ α 4-7 is stabilized by multiple contacts

between α 4 and PAC4, including salt bridges between α 4 E90 and PAC4 R82, and between α 4 E99 and PAC4 K113 (Fig. 2b, Supplementary Fig. 10). Meanwhile, α 7 makes contacts with both PAC2 and PAC3, with the C-terminal tail of PAC2 inserting into the HbYX-binding pocket formed between α 6 and α 7 (Supplementary Fig. 4d). In addition, a loop in PAC3 extends out toward α 7 to form a salt bridge between α 7 R115 and PAC3 D56 (Fig. 2c, Supplementary Fig. 10). While PAC1-4 are crucial for stabilizing α 4-7, they do not appear to play a substantial role in the incorporation of α 1-3 with the only contact being PAC4 W22 stacking in between α 3 R96 and Q100 (Fig. 2d, Supplementary Fig. 10). α 4-7 have the highest map densities in the 3D subclass, followed by α 1, α 3, and α 2 (Supplementary Fig. 2a, 4e). This suggests that α 1-3 are incorporated after formation of a stable PAC1-4/ α 4-7 complex to complete α -ring formation.

Substrate entry into the proteasome is regulated by the N-terminal tails of the α subunits, which form a closed gate in the mature 20S complex. In structures of β -ring assembly intermediates,

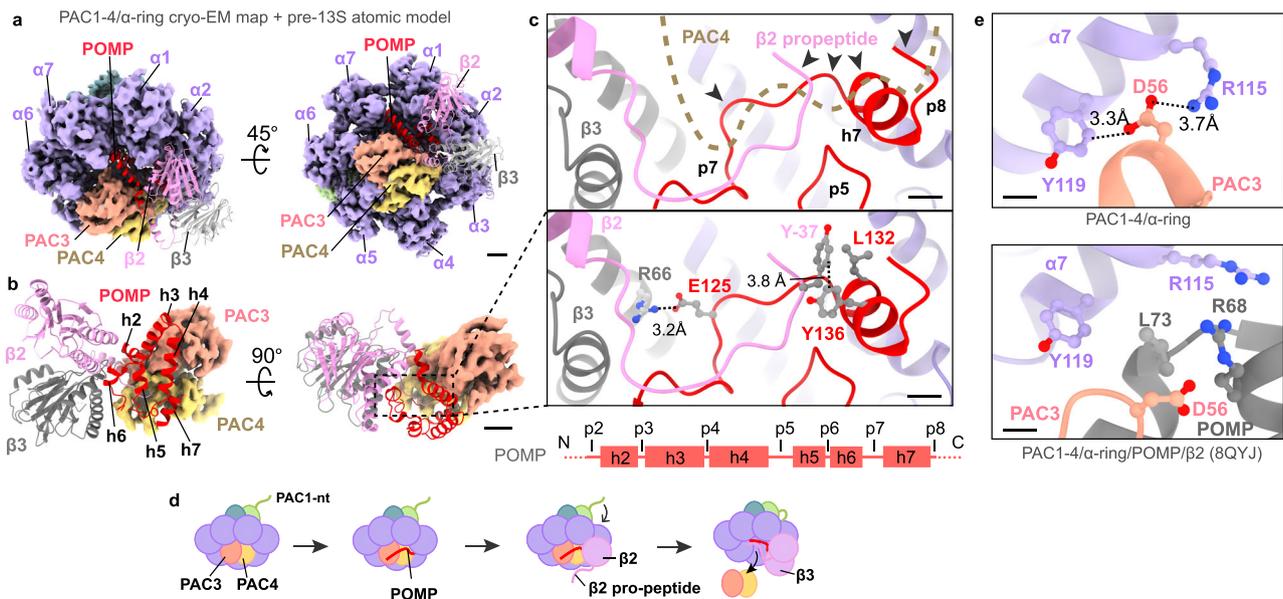


Fig. 3 | POMP and $\beta 2$ displaces PAC3/PAC4 to enable β -ring assembly. **a Cryo-EM map of PAC1-4/ α -ring overlaid with the pre-13S atomic model. POMP fits into the space between PAC3/PAC4 and $\alpha 1$ -2. $\beta 2$ and $\beta 3$ bind to the α -ring on the opposite side of PAC3/PAC4. Scale bar: 15 Å. **b** Isolated views of the overlaid structures show good fit between POMP and PAC3/PAC4 with little overlap between the chaperones. Scale bar: 15 Å. **c** The PAC4 structure clashes with the C-terminal region**

of POMP and the N-terminal tail of the $\beta 2$ pro-peptide (arrowheads), which are stabilized by interactions with $\beta 3$. Scale bar: 5 Å. **d** Schematic depicting the initiation of β -ring assembly and dissociation of PAC3/PAC4. **e** Comparison of our PAC1-4/ α -ring complex with the structure of PAC1-4/ α -ring/POMP/ $\beta 2$ (PDB 8QYJ) shows that binding of POMP disrupts the interaction between PAC3 and $\alpha 7$. Scale bar: 3 Å.

the N-terminal tail of PAC1 is inserted into the α -ring pore and blocks the open gate (Fig. 2f, g, Supplementary Fig. 5c). It has been postulated that the N-terminal tail of PAC1 may help to coordinate α -ring assembly in yeast^{24,44}. Surprisingly, the N-terminal tail of PAC1 is absent from the α -ring pore in the human PAC1-4/ α -ring structure reported here (Fig. 2e, g, Supplementary Fig. 5c). The N-terminal tails of some α -subunits in the PAC1-4/ α -ring assembly adopt unique conformations that result in a partially closed gate (Fig. 2g). Specifically, the N-terminal tails of $\alpha 2$ -4 are oriented inwards in the PAC1-4/ α -ring complex, similar to the mature 20S closed gate conformation. The unique conformations of the $\alpha 2$ and $\alpha 3$ N-terminal tails in the PAC1-4/ α -ring complex would sterically clash with the N-terminal tail of PAC1 if it was inserted in the pore (Supplementary Fig. 5c). Consequently, the N-terminal tail of PAC1 is inserted later in the assembly pathway.

Transition from α - to β -ring assembly

The assembly of the β -ring is initiated by the binding of the POMP chaperone and $\beta 2$ to the completed α -ring. Subsequently, PAC3/PAC4 must dissociate from the α -ring to make room for the incoming β subunits. To gain insight into how this occurs, we tagged POMP with an mNG2(II)-StrepII tag in Expi293F cells and purified the POMP-associated complexes (Supplementary Fig. 1e, f). Cryo-EM analysis of the sample revealed a mixture of β -ring assembly intermediates (Supplementary Fig. 2b), which were separated into four major classes using computational 3D classification of the cryo-EM images. One of the 3D subclasses lacked β subunits and PAC3/4, while another 3D subclass resembled a low-resolution structure of the 13S complex. Two 3D subclasses exhibited high-quality structural features and were refined to high-resolution. These two structures correspond to a pre-13S complex with partial binding of $\beta 4$ at ~ 3.0 Å resolution and a mixed-13S complex with partial binding of $\beta 1$, $\beta 5$, and $\beta 6$ at ~ 3.0 Å resolution (Supplementary Fig. 3)⁴⁵. Interestingly, comparison of the PAC1-4/ α -ring and pre-13S structures shows that POMP would fit into the concave surface on the α -ring next to PAC3/PAC4 with minimal overlap between the structures (Fig. 3a). In addition, $\beta 2$ binds to the α -

ring on the opposite side of where PAC3/PAC4 is bound. Biochemical studies have suggested that POMP could bind independently to the α -ring intermediate⁴⁶, which is consistent with our structural data showing that binding of $\beta 2$ would partially block access to the POMP binding pocket. Binding of POMP helps to recruit $\beta 2$, but $\beta 2$ also helps to stabilize binding of POMP¹⁶. Steric clash would occur in one small region of PAC4 with the C-terminal tail of POMP and the $\beta 2$ pro-peptide (Fig. 3b, c). The structural arrangement of the POMP C terminus and $\beta 2$ pro-peptide appears to be stabilized by interactions with $\beta 3$, which include a salt bridge between $\beta 3$ R66 with POMP E125. The release of PAC3 is coupled to the binding of $\beta 3$ ¹⁶, and our structural data provides a possible mechanism of how this may occur. POMP and $\beta 2$ bind to the PAC1-4/ α -ring complex and recruit $\beta 3$, which induces structural changes in the POMP C terminus and $\beta 2$ pro-peptide (Fig. 3d). These changes displace PAC4, which would cause the concurrent dissociation of PAC3. Comparison of our PAC1-4/ α -ring structure with the recently published PAC1-4/ α -ring/POMP/ $\beta 2$ structure⁴⁷ shows that binding of POMP disrupts the salt bridge between PAC3 D56 and $\alpha 7$ R115 (Fig. 3e). Breaking of this salt bridge likely destabilizes the binding of PAC3 to the α -ring and promotes the dissociation of PAC3.

The pre-13S and mixed-13S cryo-EM maps represent averages of multiple β -ring assembly intermediates (Supplementary Fig. 2b). Therefore, we may infer the order of β subunit incorporation from their cryo-EM map density levels. The pre-13S cryo-EM map shows strong density for $\beta 2/\beta 3$ and lower density for $\beta 4$ (Supplementary Figs. 2b, 5b), indicating that $\beta 4$ is added after $\beta 3$ as expected¹⁶. In the mixed-13S map, $\beta 2$ -4 have comparable densities that are higher than for the other β subunits (Supplementary Figs. 2b, 6). $\beta 3$ has slightly higher average map density compared to $\beta 2$ and $\beta 4$, which reflects the flexibility of the terminal β subunits of the incomplete β -ring (Supplementary Figs. 2b, 3e, 6). After $\beta 4$, $\beta 5$ is likely incorporated next as has been observed in yeast²⁴, followed by $\beta 1$ and $\beta 6$ (Supplementary Fig. 2b). There does not appear to be a clear sequential order for $\beta 1$ or $\beta 6$, but this result may be affected by the low density and potential

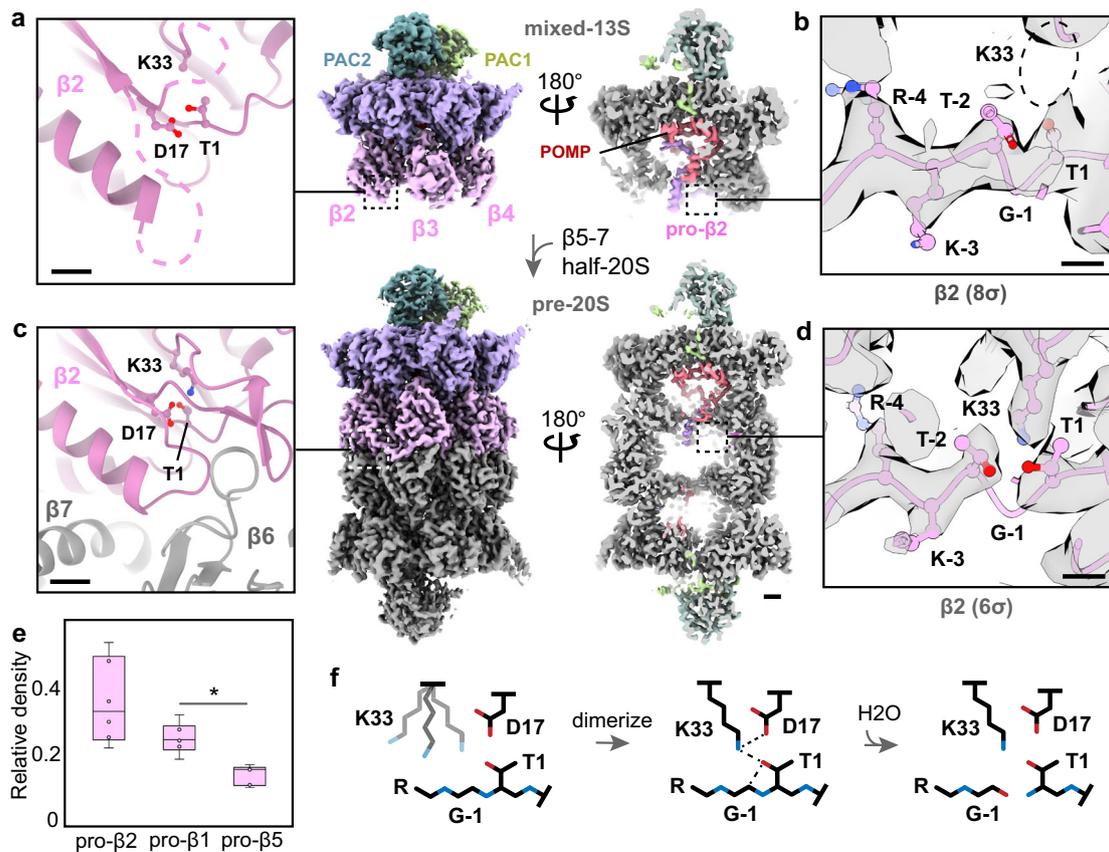


Fig. 4 | Pre-20S formation triggers autocatalysis of the β pro-peptides. **a** The cryo-EM map of mixed-13S shows poorly resolved density for the loops near the dimerization interface, indicating the region is likely flexible during β -ring assembly. Scale bar: 5 Å. **b** The sidechain of $\beta 2$ K33 is not ordered in the mixed-13S structure, preventing cleavage of the β pro-peptide. Scale bar: 3 Å. **c** Dimerization of two half-proteasomes reorders the loops near the dimer interface into a stable structural arrangement. Scale bar: 5 Å. **d** Structural rearrangements induced by half-proteasome dimerization positions the sidechain of K33 into a catalytically active conformation. A break in the cryo-EM density (gray) can be observed between $\beta 2$ T1 and T-2 at higher σ thresholds, indicating cleavage of the pro-peptide. Scale bar: 3 Å. **e** Quantification of the average cryo-EM map densities for the first six residues of the β pro-peptides after T1. Box plots indicate median (middle line), 25th, 75th percentile (box) and 5th and 95th percentile (whiskers) as well as outliers (single points). $n = 6$ residues. * p -value < 0.01, evaluated using two-tailed t-test with unequal variance. **f** Schematic depicting how structural rearrangement of K33 activates autocatalysis of the β pro-peptide. Scale bar for cryo-EM maps: 10 Å. Source data are provided as a Source data file.

flexibility of the subunits. Interestingly, no density above background is observed for $\beta 7$ in either the pre-13S or mixed-13S cryo-EM maps. This result is consistent with previous studies showing $\beta 7$ is the last subunit added to yield the half-proteasome^{16,48,49} and suggests two half-proteasomes will quickly dimerize to form the pre-20S complex.

Dimerization and autocatalytic activation

The catalytic β subunits ($\beta 1$, $\beta 2$, $\beta 5$) are expressed with N-terminal pro-peptides (Supplementary Fig. 1b, c) that are cleaved after formation of the pre-20S complex. To gain insight into how dimerization activates the catalytic β subunits, we tagged PAC2 (PSMG2) with an mNG2(11)-StrepII tag in Expi293F cells and purified the PAC2-associated complexes (Supplementary Fig. 1e, f). Cryo-EM analysis of the sample and 2D classification revealed 20S complexes mixed with α -ring and β -ring assembly intermediates (Supplementary Fig. 2c). Because samples were prepared on holey carbon grids without graphene oxide, the α -ring and β -ring assembly intermediates adopted strongly biased orientations on the grid that prevented further processing. Classes corresponding to the pre-20S proteasome were processed and resulted in a structure of the pre-20S complex at ~ 2.8 Å resolution (Fig. 4, Supplementary Figs. 2c, 3). The cryo-EM map shows clear density for PAC1/PAC2 and POMP, as well as an unknown strand of density located at the interface between the two $\beta 4$ subunits (Supplementary Fig. 7). The close proximity of this density to threonine T1 of $\beta 5$ suggests it

may be part of the $\beta 5$ pro-peptide (Supplementary Fig. 1c), which is crucial for half-proteasome dimerization and can work in trans^{16,19,20}. The recently published structures of half-proteasome and pre-20S intermediates^{47,50} lack this unassigned strand of density and show the $\beta 5$ pro-peptide in a different location, which does not support $\beta 5$ as the unassigned density and the identity of the strand remains unclear. However, we cannot rule out the possibility that the density may also represent the N terminus of POMP³⁰ or an endogenous peptide from the cell. Structures of the 20S proteasome bound by the endogenous inhibitor PI31 have been reported recently⁵¹, but the structure of PI31 is distinct from the unassigned strand of density observed in our map, indicating they are likely different molecules. Comparison of the β pro-peptide densities reveals that the different pro-peptides are processed at different rates. The $\beta 1$ pro-peptide density appears to be lower compared to the $\beta 2$ pro-peptide (Fig. 4e, Supplementary Fig. 8), although the difference is not statistically significant. Interestingly, little density for the $\beta 5$ pro-peptide can be observed near $\beta 5$ T1, indicating that the $\beta 5$ pro-peptide is mostly cleaved and has dissociated in the pre-20S cryo-EM map reported here (Supplementary Fig. 8). These data indicate that the $\beta 5$ pro-peptide is likely processed first after dimerization (Fig. 4f), followed by processing of the $\beta 1$ and $\beta 2$ pro-peptides.

Each of the catalytic β -subunits contain a catalytic threonine T1, a conserved lysine K33, and a conserved aspartic acid D17^{15,18,52,53}.

Cleavage occurs between T1 and the conserved G-1 of the pro-peptide (Supplementary Fig. 1c), which liberates the catalytic T1 residue of the β subunit. Focusing on $\beta 2$, K33 and the adjacent loop region are disordered prior to half-proteasome dimerization (Fig. 4a, b), disrupting the active site⁵⁴. Consistent with a catalytically inactive enzyme, strong density for the $\beta 2$ pro-peptide is observed in the pre-13S and mixed-13S cryo-EM maps (Fig. 4b, Supplementary Fig. 8). Upon dimerization of two half-proteasomes, K33 and the adjacent loops of $\beta 2$ are stabilized by interactions at the dimer interface (Fig. 4c, d). A recently published study identified one critical interaction as a salt bridge between a conserved arginine in the catalytic β subunit and a C-terminal aspartate in the adjacent midline spanning non-catalytic β subunit⁵⁰. This structural rearrangement places K33 into the correct position to coordinate T1 and D17, thus activating autocatalytic cleavage of the $\beta 2$ pro-peptide. Consistent with an active enzyme, a break in the pre-20S cryo-EM map density between $\beta 2$ T1 and G-1 of the pro-peptide can be observed at higher σ thresholds (Fig. 4d, Supplementary Fig. 8). Map density for the $\beta 2$ pro-peptide can still be observed at higher σ thresholds, indicating that the pro-peptide can remain bound after cleavage. While some density can be observed for $\beta 1$ and $\beta 5$ in the mixed-13S cryo-EM map, the quality of the $\beta 1$ and $\beta 5$ density is not sufficient to carry out a similar analysis as was done for $\beta 2$. However,

the conserved nature of the autocatalytic mechanism⁵² of these β subunits leads us to speculate that stabilization of K33 is a key feature of autocatalysis activation for $\beta 1$ and $\beta 5$ as well.

Maturation of pre-20S to 20S

The final steps of 20S assembly involve the removal of the β pro-peptides and dissociation of the bound chaperones, but how this occurs remains unclear. Examination of the pre-13S and mixed-13S cryo-EM maps show that the chaperones and $\beta 2$ pro-peptide are bound stoichiometrically relative to $\beta 2$ (Fig. 5, Supplementary Fig. 9). In the pre-20S cryo-EM map, both the chaperones and β pro-peptides have lower densities compared to the α and β subunits, indicating that the pre-20S cryo-EM map likely represents a mixture of maturing 20S complexes. The $\beta 2$ pro-peptide and the N-terminal region of POMP have substantially lower density in the pre-20S map determined in this study. Interestingly, a recently published structure of the pre-20S complex shows stronger $\beta 2$ pro-peptide density and an extended POMP N-terminal region that interacts with the $\beta 2$ pro-peptide⁴⁷ (Supplementary Fig. 11). Our structure of the pre-20S complex likely represents a later assembly stage compared to the recently published model, which together show that $\beta 2$ pro-peptide processing is linked to disengagement of the POMP N-terminal region (Fig. 5c, Supplementary

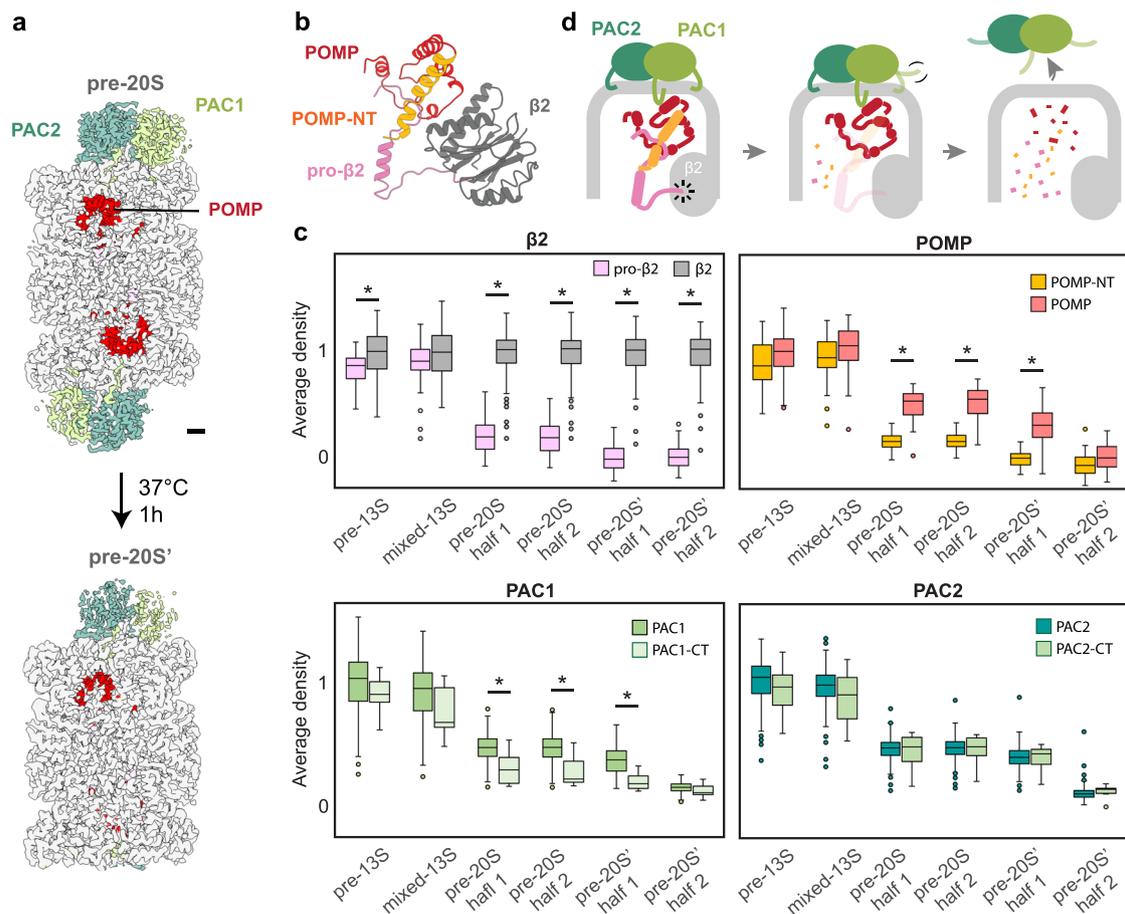


Fig. 5 | 20S maturation proceeds in three stages. a Cryo-EM maps of pre-20S and pre-20S' with cutaway section to show POMP density. Scale bar: 10 Å. **b** Isolated view of the POMP interaction with $\beta 2$ from the pre-13S atomic model. The $\beta 2$ pro-peptide (pro- $\beta 2$) interacts with the N-terminal region of POMP (POMP-NT). **c** Quantification of the cryo-EM density map values corresponding to $\beta 2$, POMP, PAC1, and PAC2 shows a decrease in chaperone occupancy over the course of 20S maturation. Examination of specific regions within the proteins reveals a correlation between density changes in pro- $\beta 2$, POMP-NT, and the C-terminal tail of PAC1 (PAC1-CT). Meanwhile, density changes in the C-terminal portion of POMP are

correlated with changes in the N-terminal portion of PAC1 and PAC2. Box plots indicate median (middle line), 25th, 75th percentile (box) and 5th and 95th percentile (whiskers) as well as outliers (single points). pro- $\beta 2$, $n = 36$ residues; $\beta 2$, $n = 163$ residues; POMP-NT, $n = 21$ residues; POMP, $n = 76$ residues; PAC1, $n = 264$ residues; PAC1-CT, $n = 11$ residues; PAC2, $n = 235$ residues; PAC2-CT, $n = 10$ residues. * p -value < 0.001, evaluated using two-tailed t-test with unequal variance. **d** Cartoon depicting the 3-stage progression of 20S maturation involving β pro-peptide cleavage, pro- $\beta 2$ and POMP-NT dissociation, and POMP/PAC1/PAC2 release and degradation. Source data are provided as a Source data file.

Fig. 9). The interaction between the $\beta 2$ pro-peptide and the POMP N-terminal region likely mediates the concerted maturation of these components. The C-terminal tail of PAC1 is also significantly lower in density relative to the rest of the protein in the pre-20S map. The relative differences in density observed within each component indicate that the final maturation of the pre-20S complex occurs through multiple stages. Because the proteins were prepared at 4 °C, we reasoned that the complexes would mature further at physiological temperatures. Indeed, incubation of purified pre-20S complexes at 37 °C for 1 h prior to cryo-EM grid preparation and analysis revealed a more mature complex, which we designate pre-20S' (Fig. 5a, Supplementary Figs. 2d, 3, 9). 2D and 3D classes corresponding to α -ring and β -ring assembly intermediates were observed as well, but orientation bias hindered further analysis of these complexes (Supplementary Fig. 2d). The cryo-EM map of pre-20S' shows lower density for the $\beta 2$ pro-peptide, PAC1, PAC2, and POMP. Taken together, these results support a 3-stage model of 20S maturation (Fig. 5d). First, the catalytic β subunits are activated through the cleavage and displacement of the β pro-peptides. Second, dissociation and degradation of the $\beta 2$ pro-peptide causes displacement of the N-terminal region of POMP, which occurs concurrently with the destabilization of the PAC1 C-terminal tail. Finally, the C-terminal region of POMP dissociates and is degraded and PAC1/PAC2 is released to yield the mature 20S proteasome.

Discussion

In this study, we present the structures of the human PAC1-4/ α -ring and late-stage pre-20S complexes, which together with the β -ring assembly intermediates and previous structures of the yeast homologs^{24,26} allow us to derive a more complete understanding of the 20S proteasome assembly pathway. Four of the five chaperones (PAC1-4) help to stabilize the initial α -4-7 subcomplex through multiple contacts. On the other hand, α -1-3 do not interact meaningfully with PAC1-4 until they bind POMP, which coincides with the insertion of the PAC1 N-terminal tail into the α -ring pore. This complex interaction between PAC1, POMP, and the α subunits may serve as a checkpoint to ensure the proper formation of the α -ring. Binding of POMP also marks the start of β -ring assembly through the recruitment of $\beta 2$. Interestingly, the clash between POMP/ $\beta 2$ and PAC3/PAC4 in our structures of the human proteins is substantially less severe than what has been predicted for the yeast homologs²⁴, suggesting a potential species-level difference in proteasome assembly at this stage. β -ring assembly appears to proceed in a similar fashion between human and yeast with $\beta 7$ as the last subunit to be incorporated, triggering dimerization of the half-proteasomes. Autocatalytic cleavage of the β pro-peptides occurs prior to the release of POMP and PAC1/PAC2, which is also observed in the yeast preholoproteasome²⁶.

One outstanding question concerns the role of the PAC1 N-terminal tail in 20S assembly, which inserts into the α -ring pore²⁴. It has been suggested that the PAC1 N-terminal tail may be involved in α -ring formation. However, the N-terminal tail of PAC1 is not inserted into the α -ring pore in the PAC1-4/ α -ring structure from this study. Therefore, it is unlikely that the PAC1 N terminus plays a role in α -ring formation and its function remains unclear. The N terminus of PAC1 is inserted in the α -ring pore of the pre-13S, mixed-13S, and pre-20S structures, indicating a structural rearrangement after completion of the α -ring and initiation of β -ring formation. It is possible that insertion of the PAC1 N-terminal tail helps to stabilize the binding of PAC1/PAC2 to the correctly formed α -ring. The N-terminal tail of PAC1 also helps to block the open α -ring gate to prevent substrate entry into the immature proteasome²⁴. In addition, the N terminus of PAC1 interacts with POMP, which may be forming a signal relay network that aids in the release of the chaperones from the maturing 20S complex.

The $\beta 5$ pro-peptide is processed more quickly relative to the pro-peptides of $\beta 1$ and $\beta 2$, but the mechanistic basis for this is unclear. $\beta 5$ contains a conserved histidine at the -2 position while threonine

occupies the same position in $\beta 1$ and $\beta 2$ (Supplementary Fig. 1c). The conserved histidine may be serving as a second proton acceptor in $\beta 5$, which is a role that is fulfilled exclusively by K33 in $\beta 1$ and $\beta 2$. The two proton acceptors in $\beta 5$ (K33 and H-2) may help to boost the autocatalysis rate of the pro-peptide. Indeed, mutation of H-2 to phenylalanine reduces but does not abrogate cleavage of the $\beta 5$ pro-peptide^{19,52}. In contrast, mutation of the conserved histidine to lysine, which also functions as a proton acceptor, preserves autocatalytic activity^{19,52}. The rapid processing of the $\beta 5$ pro-peptide may help to catalyze the degradation of other β subunit pro-peptides and POMP, which would help promote the maturation of the pre-20S complex into functional 20S proteasomes. While mutations that prevent pro-peptide cleavage of $\beta 1$ and $\beta 2$ can be tolerated, cleavage of the $\beta 5$ pro-peptide is crucial^{15,20,55-58}. The critical importance of $\beta 5$ may be the reason for the enhanced efficiency and built-in redundancy in the cleavage of the $\beta 5$ pro-peptide. Even so, studies in yeast have demonstrated that 20S complexes can still assemble and retain partial function with unprocessed $\beta 5$ subunits⁵², highlighting the robustness of 20S proteasome assembly.

Methods

Primers for gRNA synthesis

ML557

TAATACGACTCACTATAG

ML558

AAAAAAAGCACCGACTCGGTGC

ML611

AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGA-

TAACGGACTAGCCTTATTTAACTTGCTATGCTGTTTCCAGCA-

TAGCTCTTAAAC

PSMG2 oligo

TAATACGACTCACTATAGCGCCATGGAAGACACGCCGTGTTAA-

GAGCTATGCTGGAA

PSMG3 oligo

TAATACGACTCACTATAGCGCCATGGAAGACACGCCGTGTTAA-

GAGCTATGCTGGAA

POMP oligo

TAATACGACTCACTATAGGAGCTGCGGAAGATGGTGAGGTTAA-

GAGCTATGCTGGAA

Single-Strand DNA for Homology-Directed Repair (HDR)

PSMG2(PAC2)-nt-mNG2(11)-StrepII

TCTTGCCAGGCGCGGTTAGTCCCTGCTGGCCACCCCACTGC-

GACCATGACCGAGCTCAACTTCAAGGAGTGGCAAAGGCCTTACC-

GATATGATGTGGTCCCACCCTCAATTCTGA-

GAAGGGGAGTGCGTTCCCTGCGGGGAGTCGGCCCCC-

GACCTTGCCGGCTTACCCTCCT

PSMG3(PAC3)-nt-twinStrep

TTTTTCCCCTTCTTTTTAACCTAAATTAAGCTGCCACTGCA-

GAGCCCCGCATGTGGTCCCACCCTCAATTCTGA-

GAAGGGAGGGGGTCCGGTGGAGGTTCTGGGGGCTCTGCTGGTCA-

CATCCGCAGTTTGAAGAAGACACGCCGTTGGTGATATCGAAGCA-

GAAGCGGAGGTGGTGTGCGGGTCCCCA

POMP-nt-mNG2(11)-StrepII

GGCGGGTCTGACTGACGGTAACGGGGCAGAGAGGCTGTTCCGA-

GAGCTGCGGAA-

GATGACCGAGCTCAACTTCAAGGAGTGGCAAAGGCCTTACCGA-

TATGATGTGGTCCCACCCTCAATTCTGA-

GAAGGTGAGTGGTACCCGGGGCTGGAGTTCCACGGGGCTCGG-

GACCGTGGCTCGGCAGAAACAGGCAGTC

Cas9 purification

The pET15-SpCas9-NLS-HIS plasmid was purchased from Addgene (#62731)⁵⁹. T7 Express cells (New England Biolabs) were transformed with the plasmid and grown in 5 ml of LB media (VWR) supplemented with 100 μ g/ml carbenicillin (final concentration) overnight at 37 °C

with shaking at 225 RPM. The next day, the cells were added to 100 ml of LB + carbenicillin (100 µg/ml) and incubated at 37 °C with shaking at 225 RPM. When OD600 reached 0.8, the 100 ml culture was added to 1 l of pre-warmed LB + carbenicillin (100 µg/ml) + 0.5% glucose and incubated at 37 °C with shaking at 225 RPM. When OD600 reached 0.4, the temperature was reduced to 16 °C. When OD600 reached 0.6, protein expression was induced with 0.2 mM IPTG (final concentration) overnight. The next day, cells were harvested by centrifugation at 4000 × *g* for 10 min and resuspended in 20 ml HBS buffer (50 mM HEPES, pH 7, 300 mM NaCl) supplemented with 25 mM imidazole, 100 µl HALT protease inhibitor cocktail (ThermoFisher), and 4 ml of 3 M NaCl. Protein purification was carried out at 4 °C. Cells were lysed by sonication and large debris were removed by centrifugation at 4000 × *g* for 5 min. The supernatant was centrifuged at -100,000 × *g* for 20 min in an ultracentrifuge (Beckman) to remove small insoluble aggregates. The supernatant was passed through a 0.22 µm syringe filter prior to adding to 0.5 ml NEBExpress Ni Resin (New England Biolabs) by gravity flow. The beads were washed with 3 × 1 ml of HBS buffer supplemented with 25 mM imidazole. Protein was eluted with 3 × 1 ml of HBS supplemented with 300 mM imidazole and loaded onto a 1 ml HiTrap SP column equilibrated with buffer A (50 mM HEPES, 400 mM NaCl, pH 7). The column was washed with 3 ml of buffer A and bound proteins were eluted with 3 ml of buffer B (50 mM HEPES, pH 7, 650 mM NaCl). The elution was diluted 3× with buffer C (50 mM HEPES, pH 7, 150 mM NaCl) and supplemented with 1 mM DTT and 10% glycerol (final concentrations). Protein was concentrated in a 100 M.W. centrifugal device (Millipore) to -13 mg/ml, aliquoted, and stored at -80 °C.

Cas9 HDR knock-ins

The gRNA IVT template was generated by PCR in 20 µl reactions containing 8 µl Milli-Q water, 1 µl ML557 + 558 (50 µM stock), 0.5 µl ML611 (2 µM stock), 0.5 µl gene-specific oligo (2 µM stock), and 10 µl Q5 2X master mix. The PCR products were purified with the DNA clean & concentrator kit (Zymo Research, D4014) and eluted with 10 µl Milli-Q water. The gRNA was synthesized using the HiScribe T7 Quick High Yield RNA Synthesis Kit (New England Biolabs, E2050S) in 20 µl reactions containing 6 µl Milli-Q water, 1 µl DNA template (50–100 ng total), 10 µl NTP buffer mix, 2 µl T7 10X master mix, and 1 µl RNasin (Promega). The reaction mixture was incubated at 37 °C for 16 h overnight and the gRNA was purified using the RNA clean & concentrator kit (Zymo Research, R1014). The gRNA was eluted in 10 µl Milli-Q water and 1 µl of 1.5 M NaCl (150 mM final concentration) was added to the RNA elution to prevent Cas9 precipitation during RNP reconstitution. The gRNA was incubated at 70 °C for 5 min and stored at -20 °C until use. The Cas9-gRNA RNP was reconstituted by adding 20 µg gRNA (2–5 µl) to 5 µl Cas9 (80 µM; 13 mg/ml) at room temperature. 1.5 µl of ssDNA HDR template (IDT, 100 µM stock) was added to the Cas9-gRNA mixture. For each knock-in, 1 × 10⁶ Expi293F cells (Thermo Fisher, A14528) were spun down at 500 × *g* for 3 min and resuspended in 15 µl of Expi293 expression medium (Gibco). 10 µl of cells were added to the Cas9 RNP and mixed well. The mixture was gently transferred to a Lonza Nucleofector X unit S 16-well strip without generating air bubbles. Program FS-100 was executed in a Lonza 4D Nucleofector and cells were immediately transferred to a 6-well plate containing 2 ml of Expi293 expression medium with 2.9 µl HDR enhancer (IDT). The plate was incubated at 37 °C, 8% CO₂ with shaking at 125 RPM. After 24 h, the media was replaced with fresh media without HDR enhancer.

FACS for PSMG2(PAC2)-mNG2(11)-StreptII

For the PSMG2(PAC2)-mNG2(11)-StreptII knock-in, cells were sorted by fluorescence assisted cell sorting (FACS) to isolate cells with successful incorporation of the tag (Supplementary Fig. 1d). A plasmid expressing the first 10 β strands of fluorescent protein mNeonGreen2, mNG2(1-10), is transfected into the knock-in cells. mNG2(1-10) combines with the 11th β strand of mNG2, mNG2(11), in successful knock-in cells to

generate a fluorescent protein that can be used to sort the cells by FACS. When cells have recovered (7–10 days after knock-in), cells were transferred to 10 ml of Expi293 expression medium in a 125 ml flask and incubated at 37 °C, 8% CO₂, with shaking at 125 RPM. When cells reached 2 × 10⁶/ml, cells were diluted to 1 × 10⁶/ml in 20 ml media. Cells were transfected with 20 µg of pcDNA5-mNG2(1-10) plasmid using Expifectamine 293 (Gibco). Two days after transfection, fluorescent cells containing successfully incorporated tags were sorted by FACS at the Sanford Burnham Prebys Flow Cytometry Shared Resource. Sorted cells were expanded, aliquoted, and stored in liquid nitrogen.

Clonal selection of knock-in cell lines

Clonal selection was carried out on PSMG3(PAC3)-twinstrep and POMP-mNG2(11)-StreptII knock-ins to isolate cells with successfully incorporated tags. Cells were seeded at -100 cells/well in a 96-well plate and incubated at 37 °C, 8% CO₂. After the cells grew to large colonies (>100 cells), individual colonies were extracted under a Leica DM IL inverted light microscope using a pipette and seeded into a new well with fresh Expi293 expression medium containing Penicillin-Streptomycin (Cytiva). The plate was incubated at 37 °C, 8% CO₂. After 10–14 days, cells were transferred to 2 ml of fresh Expi293 expression medium in a 6-well plate and incubated at 37 °C, 8% CO₂. When cells reached 1 × 10⁶/ml, 1 ml of cells were spun down at 1000 × *g* and used for sequencing. Genomic DNA was extracted from cells using the Zymo Plasmid Miniprep kit (Zymo Research, D4209) and eluted in 25 µl Milli-Q water. DNA 100–700 bp upstream and downstream of the tag incorporation site was amplified by PCR. PCR products were purified using the DNA clean & concentrator kit (Zymo Research, D4014) and sequenced by Genewiz. Positive clones were expanded, aliquoted, and stored in liquid nitrogen.

Cell culture and proteasome purification

2 × 100 ml cultures of each cell line were grown in Expi293 expression medium (Gibco) at 37 °C, 8% CO₂, and shaking at 125 RPM. When cells reached a density of 3–4 × 10⁶/ml, cells were harvested by centrifugation at 500 × *g* for 5 min. Cells were resuspended in 10 ml HBS buffer (50 mM Hepes, 150 mM NaCl, pH 7.2) with the addition of DTT (1 mM final concentration), 100 µl of 100X HALT protease inhibitor cocktail (ThermoFisher), and 200 µl BioLock (iba). Cells were lysed by sonication and insoluble debris was removed by centrifugation at 4000 × *g* for 10 min. The supernatant was centrifuged at -100,000 × *g* for 20 min at 4 °C in an ultracentrifuge. The supernatant was filtered through a 0.2 µm syringe filter and transferred to a column containing 50 µl Strep-Tactin XT Sepharose resin (Cytiva) by gravity flow. The resin was washed with 4 × 500 µl HBS buffer and the bound proteins were eluted with 1 ml HBS buffer with 50 mM biotin (final concentration). The eluted proteins were concentrated in a 50,000 MW centrifugal concentrator (Millipore).

Immunoblotting

1–2 µg of purified and concentrated (50k MW cutoff) proteins were mixed with 6X SDS sample buffer and loaded into a 10-well 4–20% TGX precast polyacrylamide gel (Bio-Rad) along with 1 µl of Dual Color Precision Plus protein ladder (Bio-Rad). The gels were run using a mini-PROTEAN Tetra cell (Bio-Rad) and proteins were transferred to a PVDF membrane using a Trans-Blot Turbo (Bio-Rad). The membrane was washed three times in 10 ml TBST and incubated in 10 ml of 5% milk/TBST for 15 min with rocking. The membrane was washed three times with 10 ml TBST and incubated with 10 ml of 5% BSA/TBST + anti-StreptII primary antibody (Novus Biologicals, NBP2-43735, 1:10,000) for 1 h at 4 °C with rocking. The membrane was washed three times using 10 ml of TBST with 5 min incubation with rocking at room temperature between washes. The membrane was incubated with 10 ml of 5% milk/TBST + anti-mouse secondary antibody (Thermo Scientific Pierce, 31446, 1:10,000) for 1 h at 4 °C with rocking. The membrane was

washed three times using 10 ml TBST with 5 min incubation with rocking at room temperature between washes. The blots were imaged by chemiluminescence using Clarity MAX (Bio-Rad) in a ChemiDoc imaging system (Bio-Rad).

Cryo-EM imaging and data collection

For PSMG2(PAC2)-tagged samples, 2.5 μ l of purified protein (1.3 mg/ml) were applied to Quantifoil Cu 300 mesh R 2/2 grids (glow-discharged in a Pelco easiGlow at 15 mA for 25 s), blotted for 10 s (4 $^{\circ}$ C, 100% humidity, blot force 0) using Whatman 1 filter paper, and plunge frozen in liquid ethane using a Vitrobot Mark IV (ThermoFisher). For PSMG3(PAC3) and POMP-tagged samples, 2.5 μ l of purified protein (0.5 mg/ml) was applied to Quantifoil Cu 300 mesh R 2/2 grids coated with graphene oxide (Sigma)⁶⁰, blotted for 1 s (4 $^{\circ}$ C, 100% humidity, blot force 0) using Whatman 1 filter paper, and plunge frozen in liquid ethane using a Vitrobot Mark IV (ThermoFisher). The vitrified samples were imaged in a Titan Krios electron microscope (FEI/ThermoFisher) in the Cryo-EM facility of the Structural Biology Shared Resource at Sanford Burnham Prebys Medical Discovery Institute. Images were collected on a Gatan K3 detector using SerialEM⁶¹ under super-resolution settings⁶² (Supplementary Table 1).

Cryo-EM image processing

Image processing of the micrographs was carried out in cryoSPARC^{63,64}. Motion correction of the movies was carried out using the Patch Motion Correction tool using default parameters, including a maximum alignment resolution of 5 Å and B-factor of 500. CTF estimation was performed using the Patch CTF tool with an amplitude contrast of 0.1, minimum resolution of 25 Å , maximum resolution of 4 Å , minimum search defocus of 10,000 Å , and maximum search defocus of 30,000 Å . Particle picking was first carried out on the PSMG2-mNG2(11)-StreptII data using the Blob Picker tool with a minimum particle diameter of 100 Å , maximum particle diameter of 150 Å , and low-pass filter of 20 Å . 2D classification was performed using 100 classes and an initial low-pass filter resolution of 30 Å . Good 2D classes were used as templates for particle picking for the PSMG2-mNG2(11)-StreptII, PSMG2-mNG2(11)-StreptII (+37C), and POMP-mNG2(11)-StreptII data using the Template Picker tool with a particle diameter of 150 Å , low-pass filter of 20 Å , and angular sampling of 5 degrees. Good 2D classes from the POMP-mNG2(11)-StreptII data were used as templates for particle picking for the PSMG3-twinStrep data using the Template Picker tool with a particle diameter of 150 Å , low-pass filter of 20 Å , and angular sampling of 5 degrees. Multiple rounds of 2D classification were carried out for each dataset in order to remove images of non-proteins, ice, and other artefacts, as well as to isolate images that align well to produce high-quality 2D classes. Particle images that produce high-quality 2D classes were used to generate initial maps using Ab-Initio Reconstruction with initial resolution of 20 Å and maximum resolution of 12 Å . The map from the Ab-Initio Reconstruction was used as an initial model for 3D refinement using Homogeneous Refinement with an initial lowpass resolution of 20 Å and gold-standard FSC split resolution of 20 Å . For POMP-mNG2(11)-StreptII, 3D Classification was performed using 4 input maps with different β subunits masked out and initial low-pass resolution of 25 Å . 3D classes were then refined using Homogeneous Refinement with initial low-pass resolution of 30 Å . For PSMG3-twinStrep, 3D Classification was performed using 4 input maps with and without PAC1/PAC2 and PAC3/PAC4 masked out and initial low-pass resolution of 25 Å . 3D classes were then refined using Homogeneous Refinement with initial low-pass resolution of 30 Å . For PSMG2-mNG2(11)-StreptII (+37C), particles images were separated using Ab-Initio Reconstruction with 5 classes and initial lowpass resolution of 20 Å and maximum resolution of 12 Å . The Ab-Initio class corresponding to the pre-20S' was subjected to Heterogeneous Refinement using as initial models a previous map of 20S and two different maps from processing of the PSMG2-mNG2(11)-

StreptII dataset, with initial low-pass resolution of 20 Å . The 3D classes corresponding to pre-20S' were combined and refined using Homogeneous Refinement with one of the 3D classes as the initial model, initial low-pass resolution of 30 Å , and gold-standard FSC split resolution of 20 Å . Resolution of the refined cryo-EM maps were estimated by gold-standard Fourier Shell Correlation at the 0.143 criterion⁶⁵. Directional FSC of the maps were calculated⁶⁶. Atomic models were built into the cryo-EM density maps using Coot^{67,68} and Phenix^{69,70}. Initial models used for model building were Protein Data Bank (PDB) ID codes 7NAN (20S), 6JPT (PAC3), and 5WTQ (PAC4). The resulting cryo-EM maps and models were visualized using UCSF Chimera and UCSF ChimeraX^{71,72}.

Density analysis and statistics

The average cryo-EM map density values corresponding to individual subunits and chaperones were calculated per residue of an atomic model using the 'Values at Atom Positions' function in Chimera and analyzed in Microsoft Excel. Only backbone atoms were considered in the analysis. Statistical significance was evaluated using a two-tailed t-test with unequal variance to calculate *p*-values. The buried surface area of the PAC1 and PAC2 C-terminal tails were calculated in ChimeraX using the *measure buriedarea* function.

Reporting summary

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

Data availability

The atomic models and cryo-EM maps were deposited in the Protein Data Bank and Electron Microscopy Data Bank under accession codes 8TM3 and 41377 (PAC1-4/ α -ring), 8TM4 and 41378 (pre-13S), 8TM5 and 41379 (mixed-13S), 8TM6 and 41380 (pre-20S), and 41381 (pre-20S'). Source data for the cryo-EM map density analysis are provided in the Source data file with this paper. Source data are provided with this paper.

References

1. Rousseau, A. & Bertolotti, A. Regulation of proteasome assembly and activity in health and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 697–712 (2018).
2. Xie, Y. & Varshavsky, A. RPN4 is a ligand, substrate, and transcriptional regulator of the 26S proteasome: a negative feedback circuit. *Proc. Natl Acad. Sci. USA* **98**, 3056–3061 (2001).
3. Ma, M. & Liu, Z. L. Comparative transcriptome profiling analyses during the lag phase uncover YAP1, PDR1, PDR3, RPN4, and HSF1 as key regulatory genes in genomic adaptation to the lignocellulose derived inhibitor HMF for *Saccharomyces cerevisiae*. *BMC Genomics* **11**, 1–19 (2010).
4. Meiners, S. et al. Inhibition of proteasome activity induces concerted expression of proteasome genes and de novo formation of mammalian proteasomes. *J. Biol. Chem.* **278**, 21517–21525 (2003).
5. Kwak, M.-K., Wakabayashi, N., Greenlaw, J. L., Yamamoto, M. & Kensler, T. W. Antioxidants enhance mammalian proteasome expression through the Keap1-Nrf2 signaling pathway. *Mol. Cell Biol.* **23**, 8786–8794 (2003).
6. Radhakrishnan, S. K. et al. Transcription factor Nrf1 mediates the proteasome recovery pathway after proteasome inhibition in mammalian cells. *Mol. Cell* **38**, 17–28 (2010).
7. Rousseau, A. & Bertolotti, A. An evolutionarily conserved pathway controls proteasome homeostasis. *Nature* **536**, 184–189 (2016).
8. Rock, K. L. et al. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* **78**, 761–771 (1994).
9. 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).

10. Jung, S. H., Jae, S. H., Chang, I. & Kim, S. Age-associated decrease in proteasome content and activities in human dermal fibroblasts: restoration of normal level of proteasome subunits reduces aging markers in fibroblasts from elderly persons. *J. Gerontology: Ser. A* **62**, 490–499 (2007).
11. Ciechanover, A. & Kwon, Y. T. Degradation of misfolded proteins in neurodegenerative diseases: therapeutic targets and strategies. *Exp. Mol. Med.* **47**, e147–e147 (2015).
12. Vilchez, D., Saez, I. & Dillin, A. The role of protein clearance mechanisms in organismal ageing and age-related diseases. *Nat. Commun.* <https://doi.org/10.1038/ncomms6659>. (2014)
13. Walter, P. & Ron, D. The unfolded protein response: from stress pathway to homeostatic regulation. *Science* **334**, 1081–1086 (2011).
14. Hanssum, A. et al. An inducible chaperone adapts proteasome assembly to stress. *Mol. Cell* **55**, 566–577 (2014).
15. Arendt, C. S. & Hochstrasser, M. Identification of the yeast 20S proteasome catalytic centers and subunit interactions required for active-site formation. *Proc. Natl Acad. Sci. USA* **94**, 7156–7161 (1997).
16. Hirano, Y. et al. Dissecting β -ring assembly pathway of the mammalian 20S proteasome. *EMBO J.* **27**, 2204–2213 (2008).
17. Seemuller, E., Lupas, A. & Baumeister, W. Autocatalytic processing of the 20S proteasome. *Nature* **382**, 468–470 (1996).
18. Schmidtke, G. et al. Analysis of mammalian 20S proteasome biogenesis: the maturation of beta-subunits is an ordered two-step mechanism involving autocatalysis. *EMBO J.* **15**, 6887–6898 (1996).
19. Li, X., Li, Y., Arendt, C. S. & Hochstrasser, M. Distinct elements in the proteasomal $\beta 5$ subunit propeptide required for autocatalytic processing and proteasome assembly. *J. Biol. Chem.* **291**, 1991–2003 (2016).
20. Chen, P. & Hochstrasser, M. Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell* **86**, 961–972 (1996).
21. Hirano, Y. et al. A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. *Nature* **437**, 1381–1385 (2005).
22. Hirano, Y. et al. Cooperation of multiple chaperones required for the assembly of mammalian 20S proteasomes. *Mol. Cell* **24**, 977–984 (2006).
23. Le Tallec, B. et al. 20S Proteasome assembly is orchestrated by two distinct pairs of chaperones in yeast and in mammals. *Mol. Cell* **27**, 660–674 (2007).
24. Schnell, H. M. et al. Structures of chaperone-associated assembly intermediates reveal coordinated mechanisms of proteasome biogenesis. *Nat. Struct. Mol. Biol.* **28**, 418–425 (2021).
25. Stadtmueller, B. M. et al. Structure of a proteasome Pba1-Pba2 complex. *J. Biol. Chem.* **287**, 37371–37382 (2012).
26. Walsh, R. M. et al. Structure of the preholoproteasome reveals late steps in proteasome core particle biogenesis. *Nat. Struct. Mol. Biol.* **30**, 1516–1524 (2023).
27. Smith, D. M. et al. Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's α ring opens the gate for substrate entry. *Mol. Cell* **27**, 731–44 (2007).
28. Kusmierczyk, A. R., Kunjappu, M. J., Kim, R. Y. & Hochstrasser, M. A conserved 20S proteasome assembly factor requires a C-terminal HbYX motif for proteasomal precursor binding. *Nat. Struct. Mol. Biol.* **18**, 622–629 (2011).
29. Ramos, P. C., Höckendorff, J., Johnson, E. S., Varshavsky, A. & Dohmen, R. J. Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell* **92**, 489–499 (1998).
30. Burri, L. et al. Identification and characterization of a mammalian protein interacting with 20S proteasome precursors. *Proc. Natl Acad. Sci. USA* **97**, 10348–10353 (2000).
31. Witt, E. et al. Characterisation of the newly identified human Ump1p homologue POMP and analysis of LMP7($\beta 5$) incorporation into 20S proteasomes. *J. Mol. Biol.* **301**, 1–9 (2000).
32. Griffin, T. A., Slack, J. P., McCluskey, T. S., Monaco, J. J. & Colbert, R. A. Identification of proteasemblin, a mammalian homologue of the yeast protein, Ump1p, that is required for normal proteasome assembly. *Mol. Cell Biol. Res. Commun.* **3**, 212–217 (2000).
33. Kock, M. et al. Proteasome assembly from 15S precursors involves major conformational changes and recycling of the Pba1-Pba2 chaperone. *Nat. Commun.* **6**, 1–10 (2015).
34. Kusmierczyk, A. R., Kunjappu, M. J., Funakoshi, M. & Hochstrasser, M. A multimeric assembly factor controls the formation of alternative 20S proteasomes. *Nat. Struct. Mol. Biol.* **15**, 237–244 (2008).
35. Yashiroda, H. et al. Crystal structure of a chaperone complex that contributes to the assembly of yeast 20S proteasomes. *Nat. Struct. Mol. Biol.* **15**, 228–236 (2008).
36. Gerards, W. L. H. et al. The human α -type proteasomal subunit HsC8 forms a double ringlike structure, but does not assemble into proteasome-like particles with the β -type subunits HsDelta or HsBPROS26. *J. Biol. Chem.* **272**, 10080–10086 (1997).
37. Velichutina, I., Connerly, P. L., Arendt, C. S., Li, X. & Hochstrasser, M. Plasticity in eucaryotic 20S proteasome ring assembly revealed by a subunit deletion in yeast. *EMBO J.* **23**, 500–510 (2004).
38. Zhao, J. et al. Structural insights into the human PA28–20S proteasome enabled by efficient tagging and purification of endogenous proteins. *Proc. Natl Acad. Sci. USA* **119**, e2207200119 (2022).
39. Anzalone, A. V., Koblan, L. W. & Liu, D. R. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844 (2020).
40. Leonetti, M. D., Sekine, S., Kamiyama, D., Weissman, J. S. & Huang, B. A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc. Natl Acad. Sci. USA* **113**, E3501–E3508 (2016).
41. Satoh, T. et al. Molecular and structural basis of the proteasome α subunit assembly mechanism mediated by the proteasome-assembling chaperone PAC3-PAC4 heterodimer. *Int. J. Mol. Sci.* **20**, 2231 (2019).
42. Takagi, K. et al. Pba3-Pba4 heterodimer acts as a molecular matchmaker in proteasome α -ring formation. *Biochem. Biophys. Res. Commun.* **450**, 1110–1114 (2014).
43. Wu, W. et al. PAC1-PAC2 proteasome assembly chaperone retains the core $\alpha 4$ – $\alpha 7$ assembly intermediates in the cytoplasm. *Genes Cells* **23**, 839–848 (2018).
44. Schnell, H. M. et al. Mechanism of proteasome gate modulation by assembly chaperones Pba1 and Pba2. *J. Biol. Chem.* **298**, 101906 (2022).
45. Frentzel, S., Pesold-Hurt, B., Seelig, A. & Kloetzel, P. M. 20S proteasomes are assembled via distinct precursor complexes processing of LMP2 and LMP7 proproteins takes place in 13-16S preproteasome complexes. *J. Mol. Biol.* **236**, 975–981 (1994).
46. Fricke, B., Heink, S., Steffen, J., Kloetzel, P. M. & Krüger, E. The proteasome maturation protein POMP facilitates major steps of 20S proteasome formation at the endoplasmic reticulum. *EMBO Rep.* **8**, 1170–1175 (2007).
47. Adolf, F. et al. Visualizing chaperone-mediated multistep assembly of the human 20S proteasome. *Nat. Struct. Mol. Biol.* **30**, 1176–1188 (2024).
48. Li, X., Kusmierczyk, A. R., Wong, P., Emili, A. & Hochstrasser, M. β -Subunit appendages promote 20S proteasome assembly by overcoming an Ump1-dependent checkpoint. *EMBO J.* **26**, 2339–2349 (2007).
49. Marques, A. J., Glanemann, C., Ramos, P. C. & Dohmen, R. J. The C-terminal extension of the $\beta 7$ subunit and activator complexes

- stabilize nascent 20 S proteasomes and promote their maturation. *J. Biol. Chem.* **282**, 34869–34876 (2007).
50. Velez, B. et al. Mechanism of autocatalytic activation during proteasome assembly. *Nat. Struct. Mol. Biol.* **31**, 1167–1175 (2024).
 51. Rawson, S. et al. Yeast PI31 inhibits the proteasome by a direct multisite mechanism. *Nat. Struct. Mol. Biol.* **29**, 791–800 (2022).
 52. Huber, E. M. et al. A unified mechanism for proteolysis and autocatalytic activation in the 20S proteasome. *Nat. Commun.* **7**, 1–10 (2016).
 53. Ditzel, L. et al. Conformational constraints for protein self-cleavage in the proteasome. *J. Mol. Biol.* **279**, 1187–1191 (1998).
 54. Witt, S. et al. Proteasome assembly triggers a switch required for active-site maturation. *Structure* **14**, 1179–1188 (2006).
 55. Groll, M. et al. The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study. *Proc. Natl Acad. Sci. USA* **96**, 10976–10983 (1999).
 56. Dick, T. P. et al. Contribution of proteasomal β -subunits to the cleavage of peptide substrates analyzed with yeast mutants. *J. Biol. Chem.* **273**, 25637–25646 (1998).
 57. Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U. & Wolf, D. H. The active sites of the eukaryotic 20 S proteasome and their involvement in subunit precursor processing. *J. Biol. Chem.* **272**, 25200–25209 (1997).
 58. Jäger, S., Groll, M., Huber, R., Wolf, D. H. & Heinemeyer, W. Proteasome β -type subunits: unequal roles of propeptides in core particle maturation and a hierarchy of active site function. *J. Mol. Biol.* **291**, 997–1013 (1999).
 59. D'Astolfo, D. S. et al. Efficient intracellular delivery of native proteins. *Cell* **161**, 674–690 (2015).
 60. Pantelic, R. S., Meyer, J. C., Kaiser, U., Baumeister, W. & Plitzko, J. M. Graphene oxide: a substrate for optimizing preparations of frozen-hydrated samples. *J. Struct. Biol.* **170**, 152–156 (2010).
 61. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
 62. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
 63. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
 64. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCr* **7**, 253–267 (2020).
 65. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
 66. Dang, S. et al. Cryo-EM structures of the TMEM16A calcium-activated chloride channel. *Nature* <https://doi.org/10.1038/nature25024> (2017).
 67. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr D. Biol. Crystallogr* **60**, 2126–2132 (2004).
 68. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D. Biol. Crystallogr* **66**, 486–501 (2010).
 69. Adams, P. D. et al. The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).
 70. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D. Biol. Crystallogr* **66**, 213–221 (2010).
 71. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
 72. Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

Acknowledgements

We would like to thank Laura Koepping from the Sanford Burnham Prebys Medical Discovery Institute (SBP) Cryo-EM facility of the Structural Biology Shared Resource and Amy Cortez from the SBP Flow Cytometry Shared Resource for their amazing support in the cryo-EM and cell sorting work, respectively. We would also like to thank Yifan Cheng (University of California San Francisco) and Elena Pasquale (SBP) for critical reading of the manuscript. Finally, we thank Kelly Kersten (SBP) for the help with FACS analysis. This work is supported by National Institutes of Health (NIH) grant R35 GM147487 to J.Z. This work is also supported by National Cancer Institute Cancer Center Support Grant P30 CA030199 and NIH instrumentation grant S10 OD026926. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Author contributions

H.Z. and J.Z. conceived the project and designed the experiments. H.Z. and C.Z. generated and maintained the cell lines. H.Z. and Z.M. purified the proteins and prepared the samples. H.Z. collected and processed the data. H.Z. and J.Z. analyzed the data. H.Z., C.Z., and J.Z. prepared the figures and wrote the manuscript.

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/s41467-024-52513-0>.

Correspondence and requests for materials should be addressed to Jianhua Zhao.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024