SUPPLEMENTARY INFORMATION

Structural analysis of a heptameric ATP-independent proteasome activator reveals a novel mechanism for 20S proteasome binding and activation.

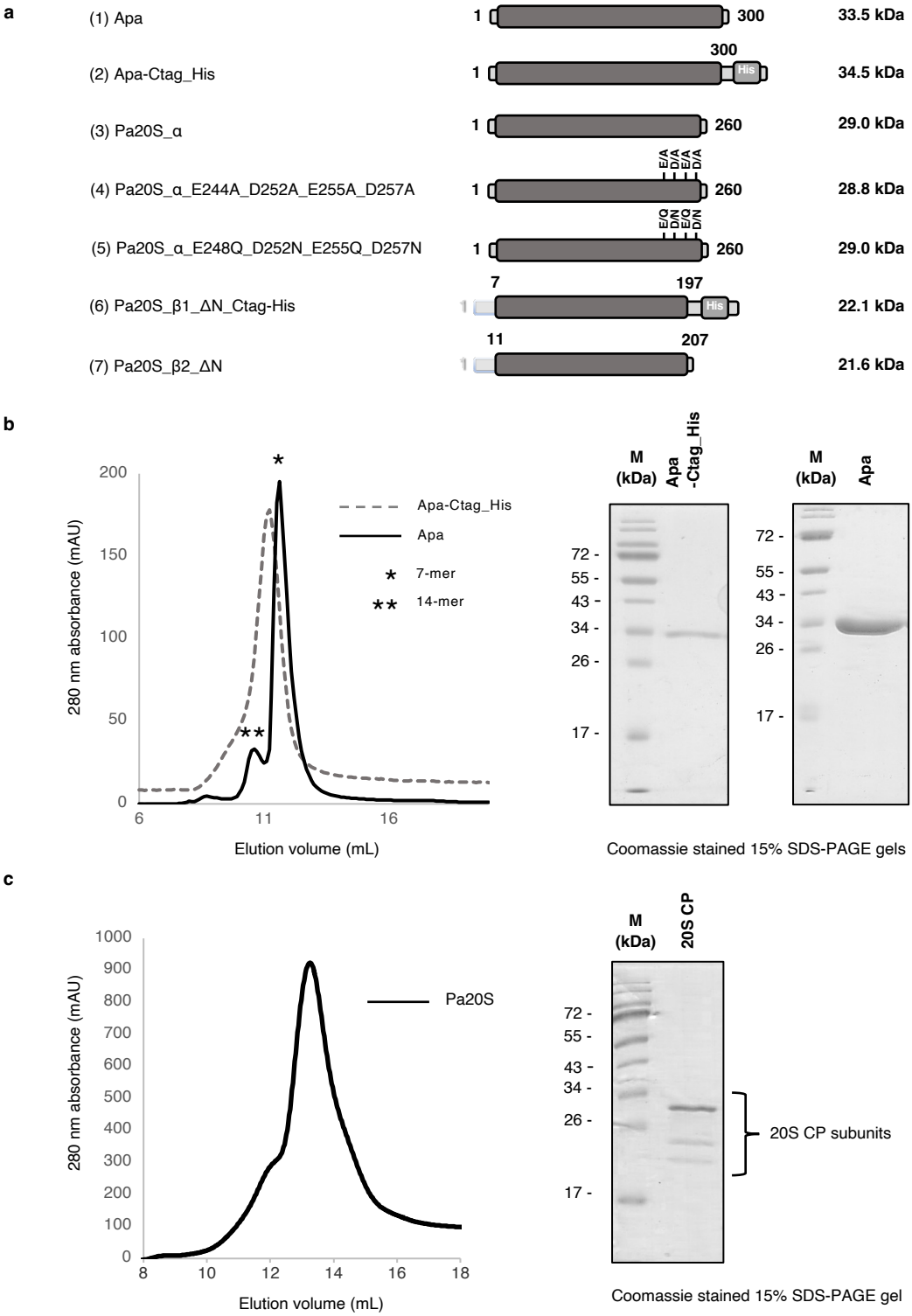
Laura Mariño Puertas[1#†], Federico Musso[1#], Céline Brochier-Armanet[2], Frank Gabel[1†], Florian Chenavier[1 †], Didier Flament[4], Felix Weis[1], Eric Girard[1], Guy Schoehn[1] & Bruno Franzetti[1*]

[1]Institut de Biologie Structurale, Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble.
[2]Laboratoire de Biométrie et Biologie Évolutive, Université Lyon 1, Lyon, France.
[3]Laboratoire de Microbiologie des Environnements Extrêmes, Ifremer-CNRS-Université de Bretagne Occidentale, Brest, France.

[#]These authors contributed equally.

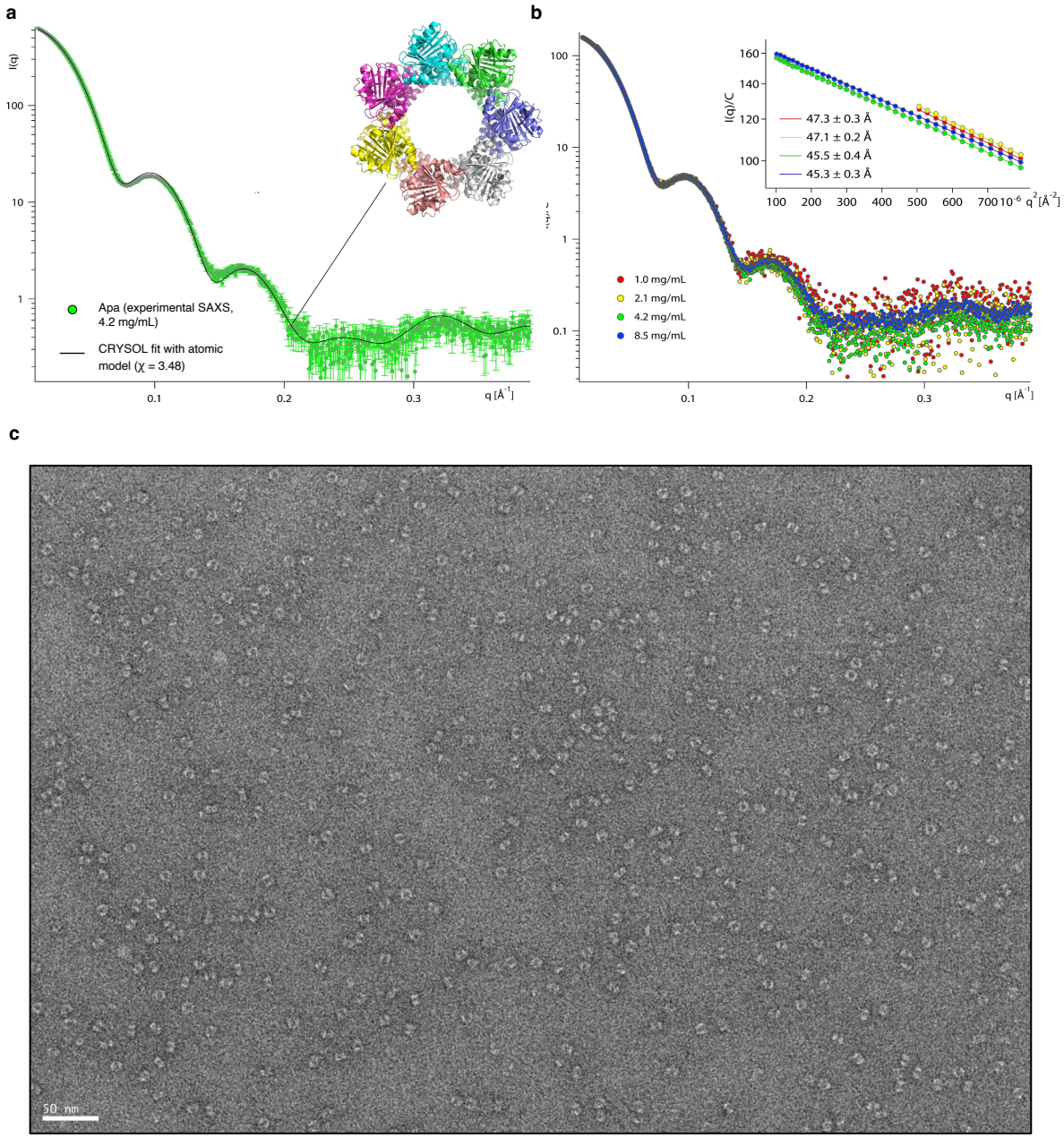
[†]Present address: LMP: Molecular Biology Institute of Barcelona, IBMB-CSIC, Barcelona, Spain. FG: Institut Laue-Langevin, Grenoble, France. FC: Chromatin Structure & Mobile DNA Laboratory, The Francis Crick Institute, London, UK

*Corresponding author. Email: bruno.franzetti@ibs.fr



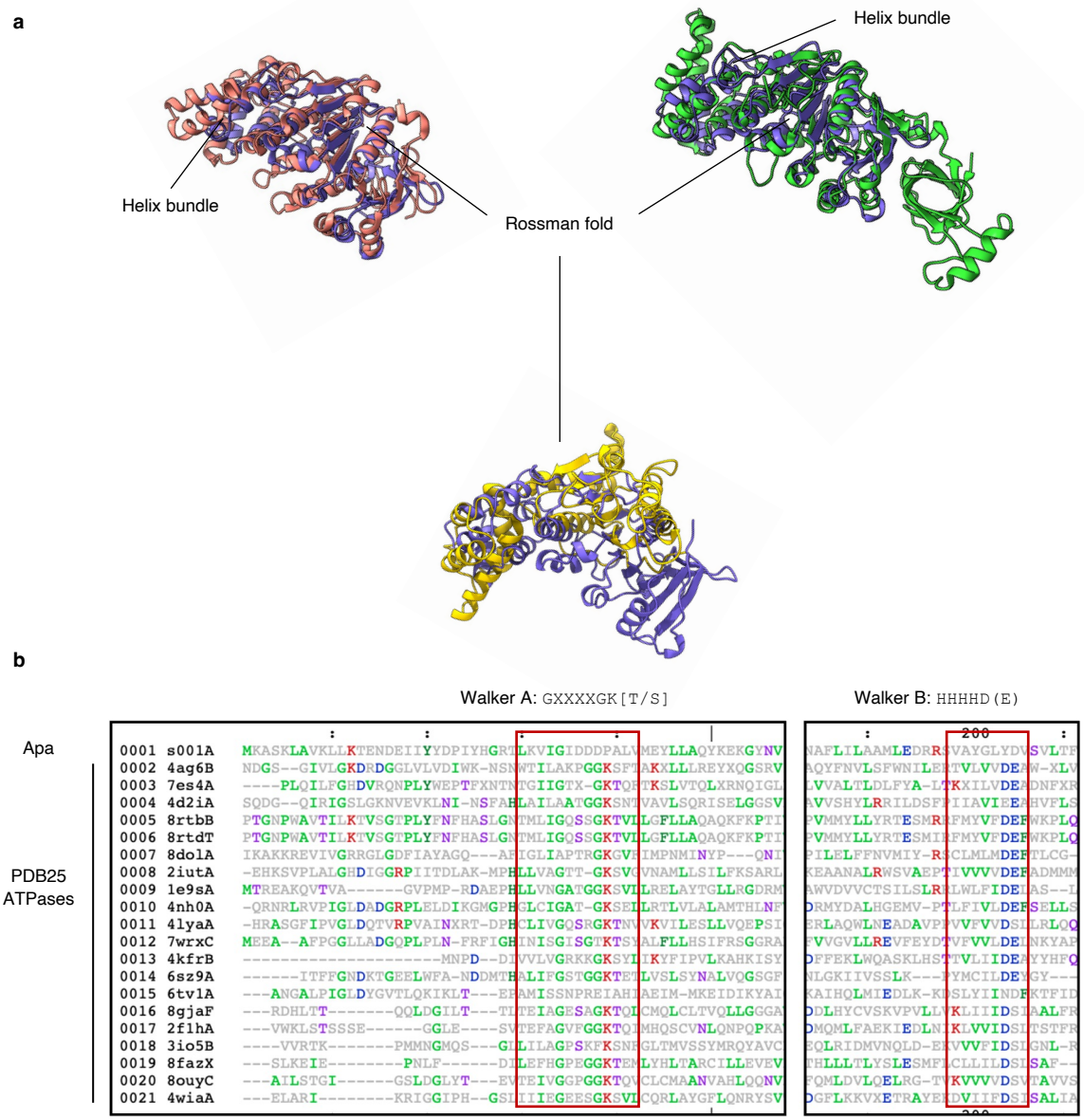
Supplementary Figure 1: Overview of the recombinant proteins studied.

a, Schematic diagram of Apa constructs and of 20S CP subunits used for the *in vitro* assembly of the *P. abyssi* 20S CP in *E. coli*. The construct (1) is the untagged full-length Apa protein, the construct (2) was modified with the addition of a 6x-Histidine tag at the C-terminal end. The construct (6) is the mature β1 20S CP subunit truncated for its 6 first amino acids, where the original Gly^6 was replaced by Met. In this construct a 6xHistidine tag was added at the C-terminal end. The construct (7) corresponds to the mature β2 20S CP subunit truncated for its 10 first amino acids, where the original Gly^{10} was also replaced by Met. **b,** Left: size exclusion chromatography profile of Apa recombinant variants eluted from a Superdex 200 10/300 column (GE Heathcare). The 14-mer form tended to evolve into the 7-mer species while the low molecular weight peak remained stable in 7-mer. The 7-mer species (estimated molecular weight of 234.5 kDa) was used for further characterization. Right: SDS-PAGE gels of purified Apa recombinant proteins stained by Coomassie Blue. **c,** Left: size exclusion chromatography profile of the 20S CP eluted from a Superose 6 10/300 column (GE Heathcare) on the left panel; the shoulder corresponds to aggregate forms. Right: SDS-PAGE gel of purified and assembled 20S CP stained by Coomassie blue.

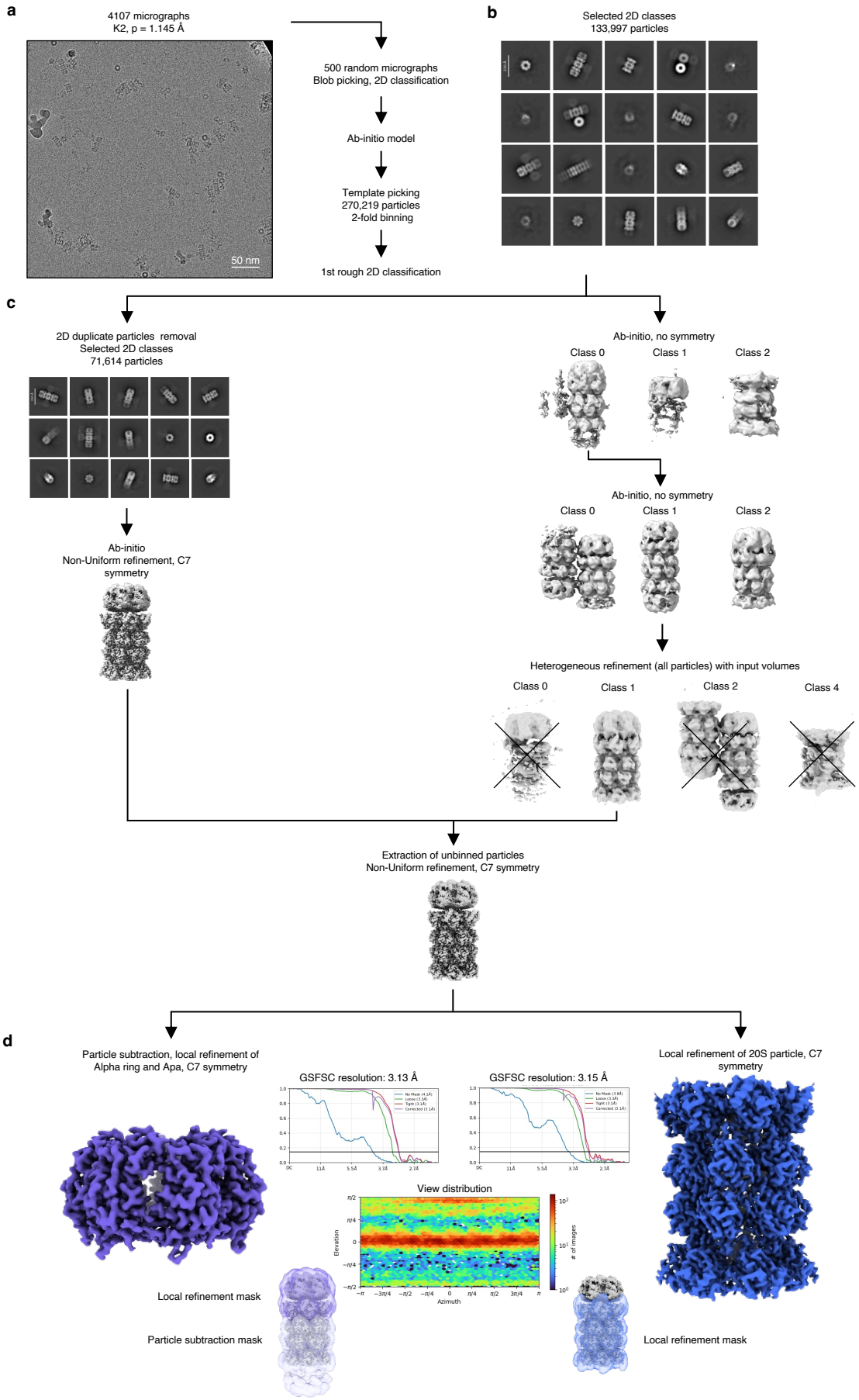


Supplementary Figure 2: SAXS and negative staining TEM characterisation of Apa.

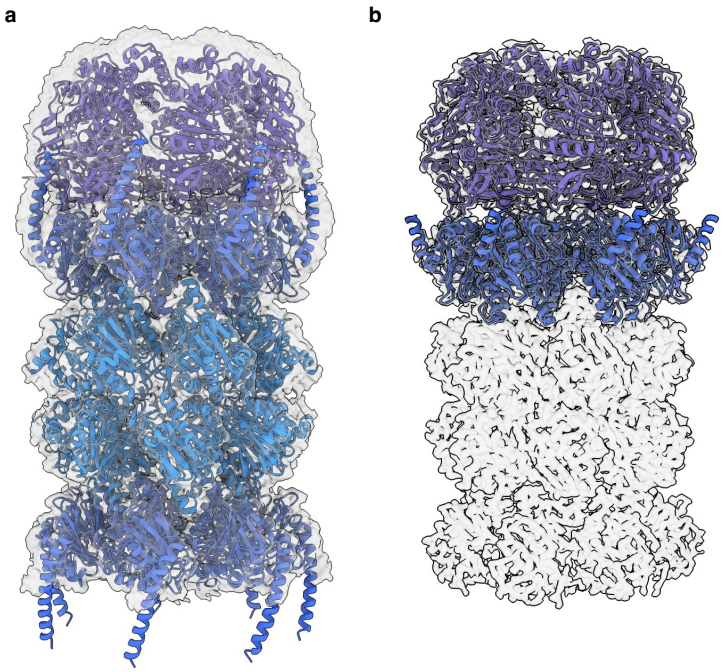
a, CRYSOL fits of the Apa X-ray atomic model with the experimental SAXS data at 4.2 mg/mL (no background adjustment applied). The numerical value of χ, as well as a visual inspection, confirm the oligomeric state of the atomic model and the conformations and arrangements of the individual subunits. **b,** Experimental Apa SAXS data at four different concentrations, normalized by concentration. *Inset*: Guinier fits and radii of gyration from the four data sets. The smallest q-values from the two lowest concentrations have been cut off, since aggregation was present in these two datasets. **c,** Negative staining micrograph of the untagged form of Apa stained with sodium silicotungstate at 30000X magnification.



Supplementary Figure 3: Structural comparison between Apa and representative hits from the DALI web server.
a, Comparison of Apa (purple) and structurally similar proteins. Left: ATP-dependent protein translocase VirB4 (PDB 4AG6, salmon); Middle: the hexameric AAA+ ATPase proteasome activator PAN (PDB 6HE4, gold); Right: the hexameric AAA+ ATPase HerA (PDB 4D2I, lime green). Common structural features of AAA+ ATPases are indicated. **b,** Structure-based sequence alignment of Apa (on top) vs. high-scoring DALI hits from the PDB25 subset, ranked by significance (Z-score). Red boxes highlight the conserved ATP binding and hydrolysis motifs, which are absent in Apa. The locations of the degenerated ATP-binding sites on Apa correspond to the lower protomer-protomer interface area, thereby preventing access to a small molecule to the hypothetical binding site.

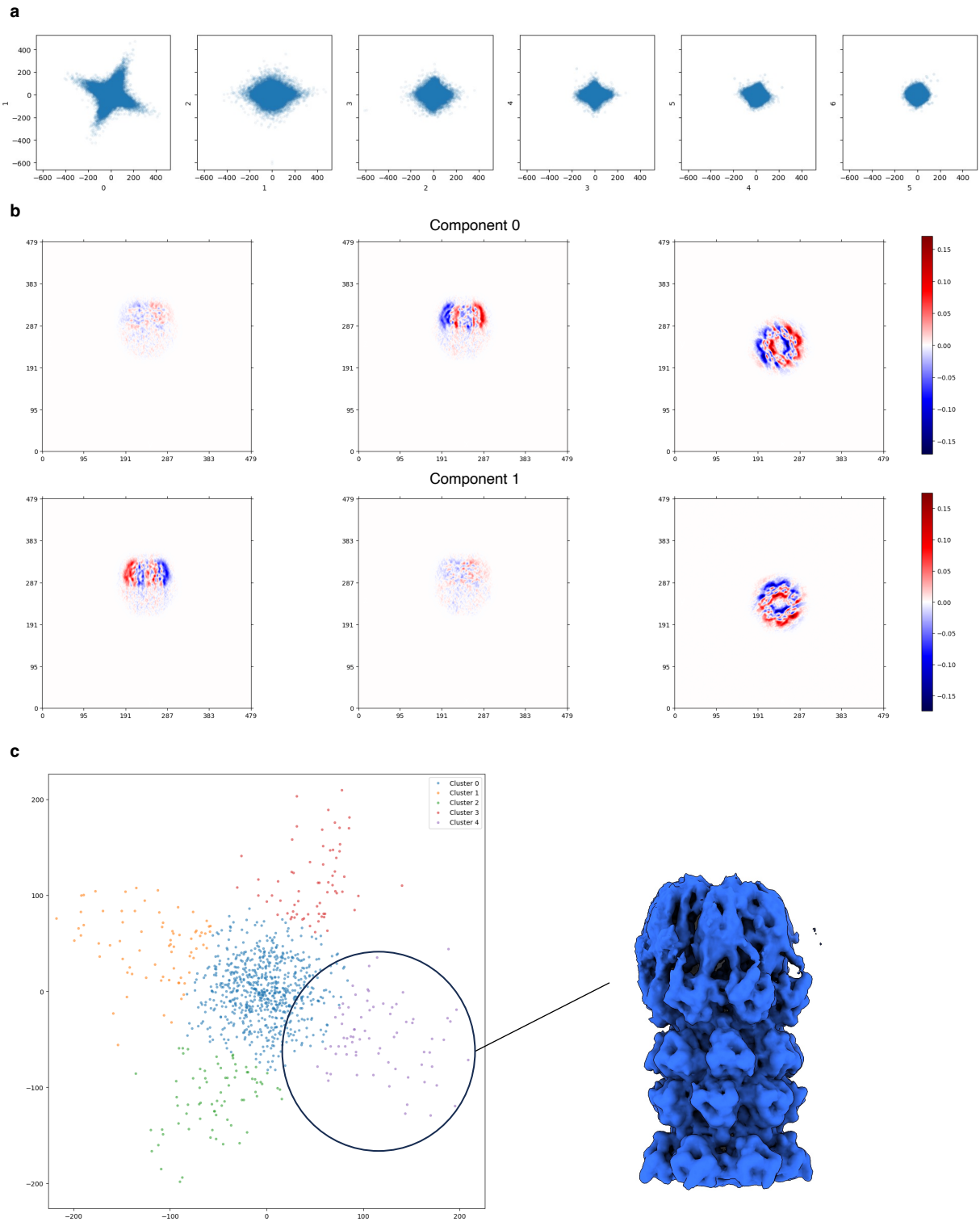


Supplementary Figure 4: Cryo-EM data processing flowchart.
a, Representative micrograph of the Apa-20S CP complex and first steps of particle curation. Uncapped proteasomes are visible along with singly or doubly capped 20S CP particles. **b,** Selected 2D classes of the first coarse round of classification. Due to the strong tendency of 20S CP particles to cluster together on CryoEM grids, some classes contain duplicated views. **c,** Outline of the strategy used to obtain the first consensus volume. **d,** Locally refined maps of Apa and the 20S CP. Middle page: GSFSC curves for each reconstruction, angular distribution of the common particle set and masks used for particle subtraction or local refinement.

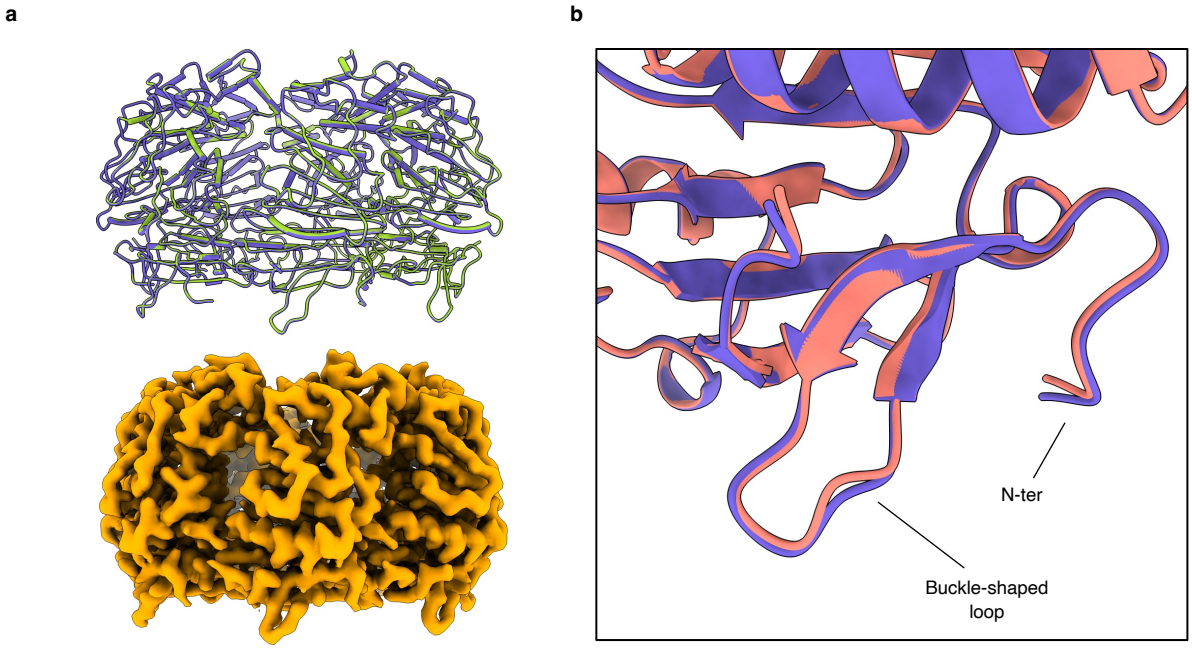


Supplementary Figure 5: Models of the Apa-20S CP complex

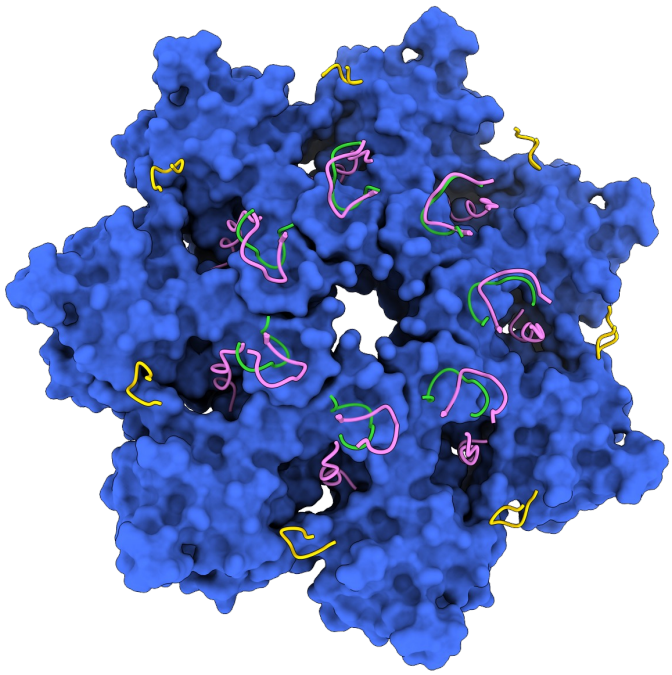
a, AlphaFold 3 predictions of the interaction between Apa (purple) and the 20S CP α ring (blue) docked into the low contour consensus map of the Apa-20S CP assembly. Additional AlphaFold 3 models for the β1 and β2 rings are represented in light blue. **b,** Atomic model of Apa (purple) and the 20S CP α ring (blue) were built inside the two focused maps, resulting in the α ring model being truncated (PDB 9S34).



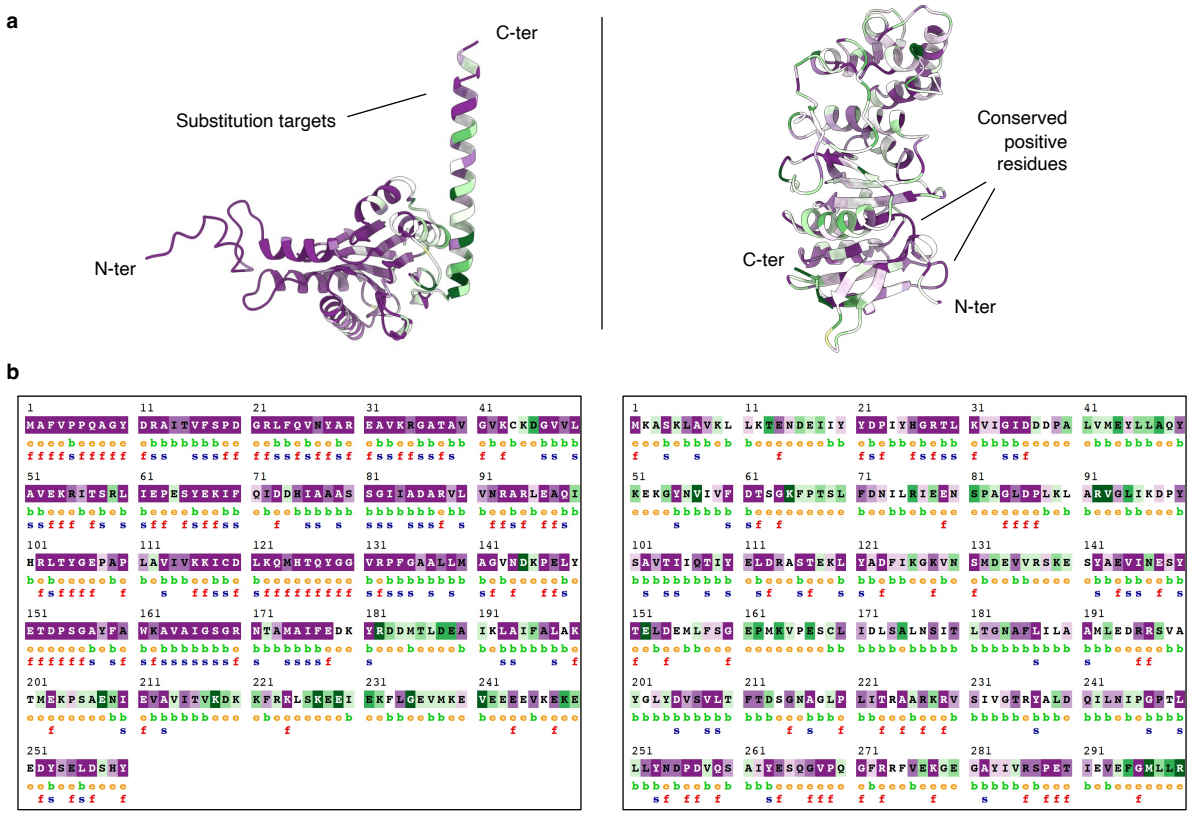
Supplementary Figure 6: 3D variability analysis of the alignments of Apa with respect to the 20S CP.
a, Distribution of particles along the first six principal components computed by the 3D Variability Analysis (3DVA) tool in CryoSPARC. The smooth distribution of particles suggests a continuous rather than discrete heterogeneity in the dataset. **b,** Detail of components 0 and 1. Differences captured by each principal component are here visualised in 2D as addition and subtraction of volume density, with red corresponding to added density and blue indicating a subtraction. **c,** The same variability along components 0,1 is here visualised in clustering mode. Clusters occupying the edges of the distribution correspond to a tilted state of Apa on top of the proteasome. The variability along components 0,1 is also summarised as a volume series in Supplementary Video 1.



Supplementary Figure 7: Structural comparison of Apa in its 20S CP-bound and free form.
a, Top: Overlay of the backbone from the cryo-EM models of Apa in its 20S CP-bound (purple) and alone in solution (lime green), showing no trace of major rearrangements (RMSD 0.1 Å between 300 Cαs). Bottom: Cryo-EM map of Apa alone in solution. **b,** Superposition detail of the unbound X-ray (salmon) vs. proteasome-bound (purple) cryo-EM Apa structure, showing the persistence of the buckle-shaped loop (RMSD 0.33 Å between 300 Cαs).

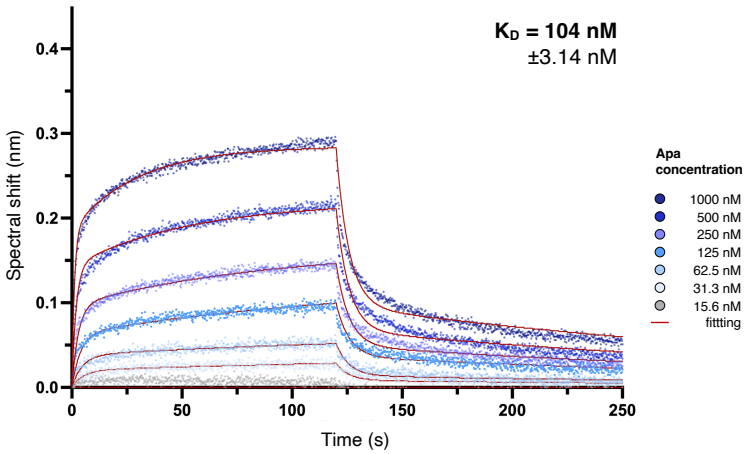


Supplementary Figure 8: Structural comparison of the activation loops used by 11S activators vs Apa.
Overlay of the buckle-shaped loops of Apa (shown as a cartoon in gold) vs. the activation loops from the eukaryotic ATP-independent proteasome activator PA28 from *Plasmodium falciparum* (PDB 6MUX, shown as a cartoon in lime green) and the combination of activation loops and docking C-termini from the PA26 activator of *Trypanosoma brucei* (PDB 1FNT, shown as a cartoon in pink). The α rings from *P. falciparum* and *T. brucei* have been superimposed to our model, which is the only one shown as a blue coloured surface (RMSD 0.91 Å over 231 Cαs for PA28 and 0.99 Å over 199 Cαs for PA26).

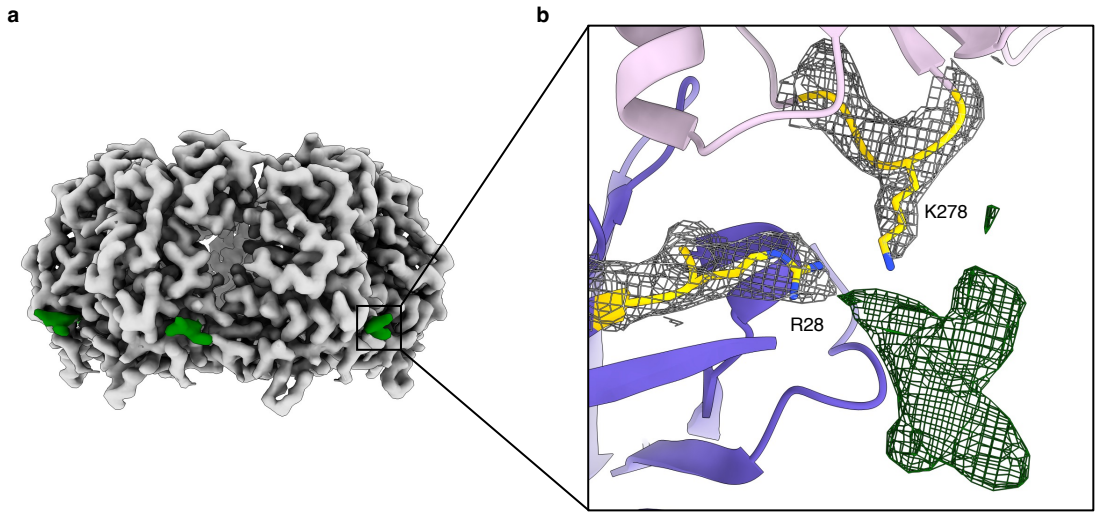


Supplementary Figure 9: Conservation of the 20S CP α subunit and Apa.

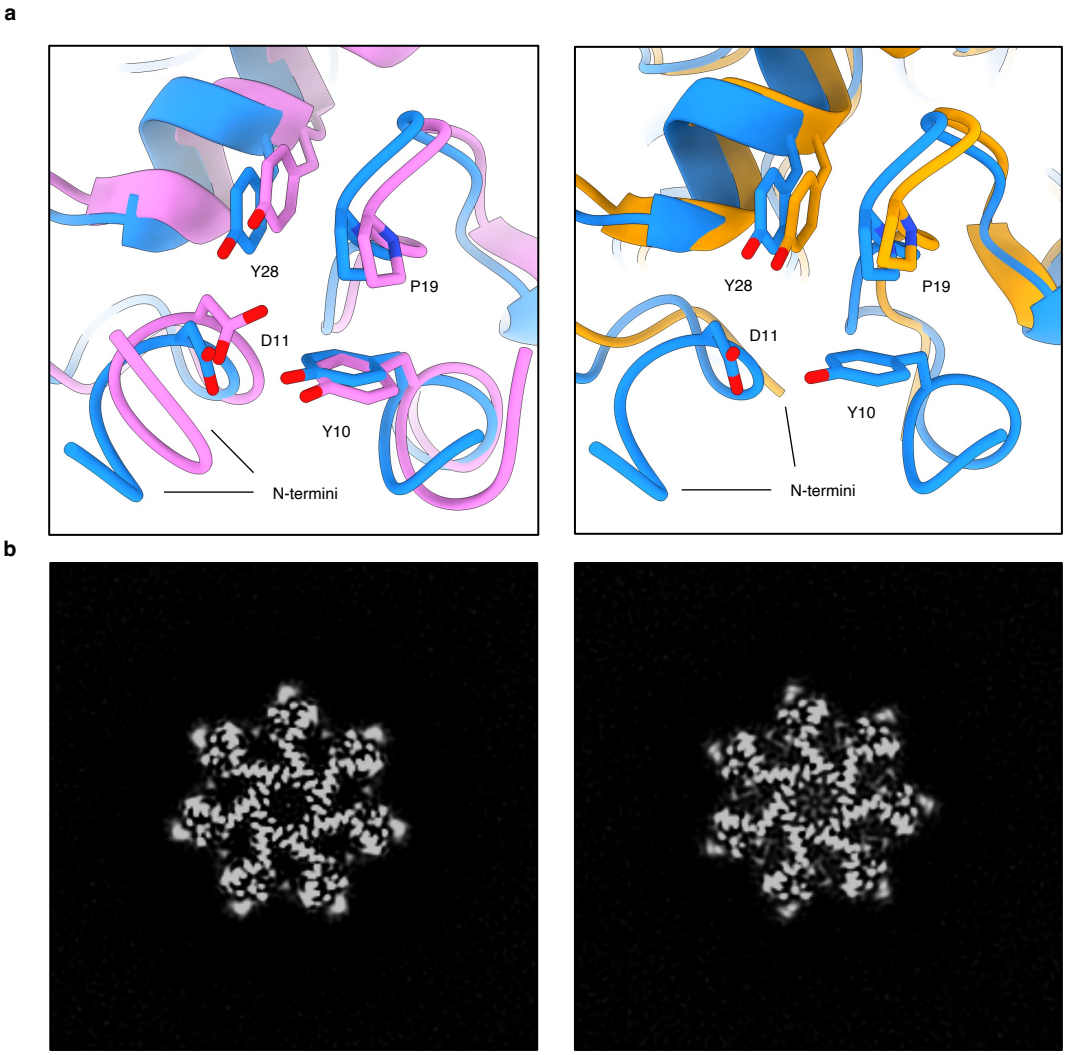
a, Left: Conservation score computed using the ConSurf webserver mapped on the AlphaFold2 model of the 20S CP α subunit, using all 100 RefSeq sequences from Archaea. The region where substitutions were performed is indicated. Conservation scores ranges from low (dark green) to medium (white) to high (dark purple). Right: conservation score mapped on our model of Apa, computed using all the 60 complete sequences within the *Thermococcales* order found by BlastP. **b,** Left: ConSurf conservation scores mapped on the sequence of the 20S CP α subunit, including predictions for buried (b) and exposed (e) residues as well as predicted functional (f) and structural (s) residues. Right: same for the Apa sequence.



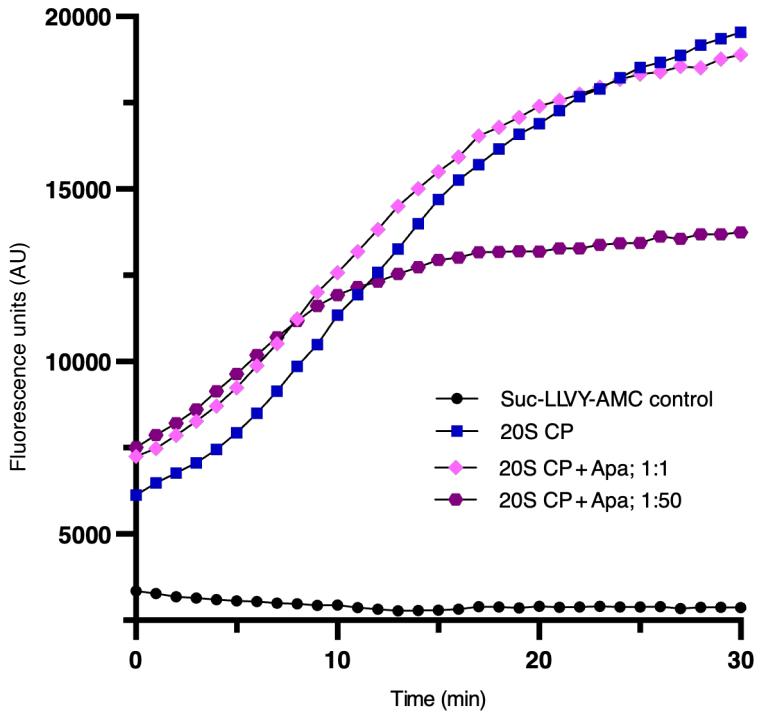
Supplementary Figure 10: BLI sensorgram of the Apa-20S CP (αE248Q,D252N,E255Q,D257N) interaction.
Disruption of the charge with a preserved side chain size causes a net decrease in affinity compared to the wild-type interaction, resulting as well in a higher dissociation rate (Ka = 2.68×10^4 1/Ms; Kdis = 2.79×10^{-3} 1/s).



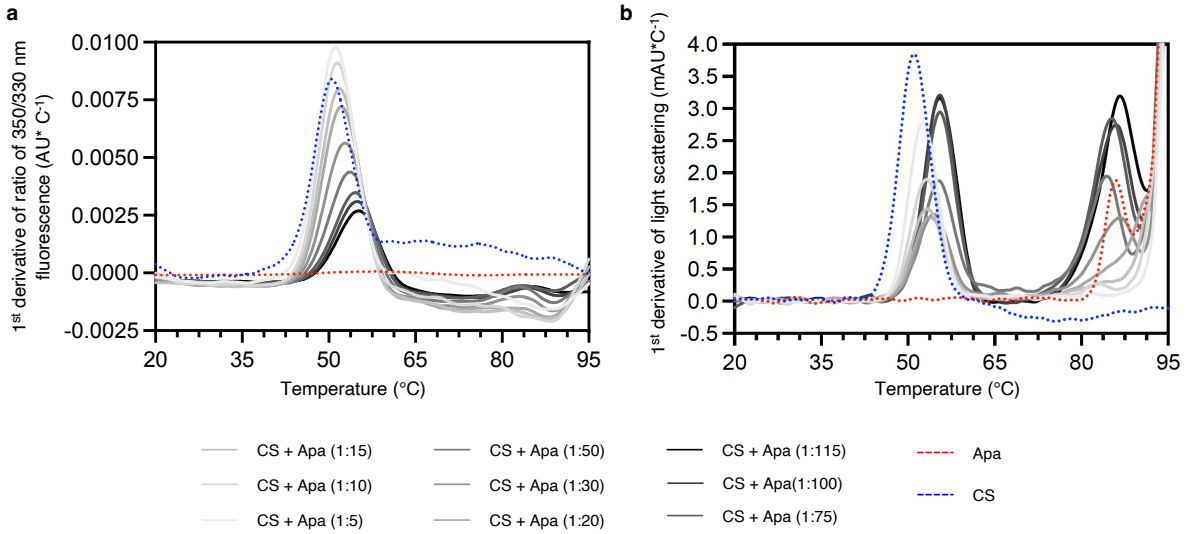
Supplementary Figure 11: Additional density observed in the focused cryo-EM map of 20S CP-bound Apa.
a, Unmodeled density is shown in green around the refined Apa map. **b,** Detail of the unmodelled density and its proximity to the conserved and positively charged residues in between two Apa protomers (coloured in purple and thistle).



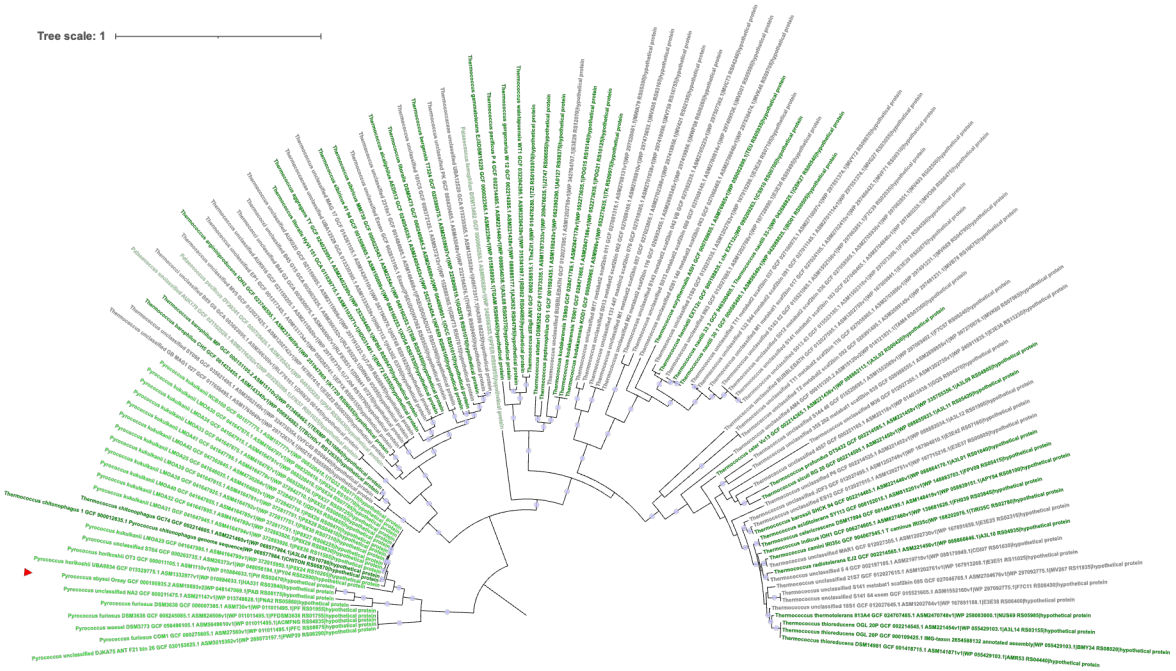
Supplementary Figure 12: α subunit gating in Apa-20S CP complexes.
a, Left: Overlay of the four interlocking residues between two α subunit protomers of the Apa-bound 20S CP (blue) with the open-gate α subunit state in the *Archaeoglobus fulgidus* PAN-20S proteasome (pink, PDB 6HE8, RMSD 0.91 Å between 208 Cαs). D11-Y10 likely establish H bonds, whereas Y28-P19 bind through CH/π interactions. The outward-facing N-termini are visible in the foreground. The first seven N-terminal residues displayed a smeared EM density, prompting us to model α gates starting from A8. Right: overlay of the same residues (blue) with the inactive closed-gate α subunit state in the *Archaeoglobus fulgidus* 20S proteasome (orange, PDB 6HE7, RMSD 0.84 Å between 201 Cαs). The N-termini of the closed-gate 20S point inwards and exhibit a smeared electron density. **b,** Left: slice through density in the Apa-bound side of the 20S proteasome showing a clear axial channel. Right: slice through density in the unbound 20S side showing a density-occupied channel.



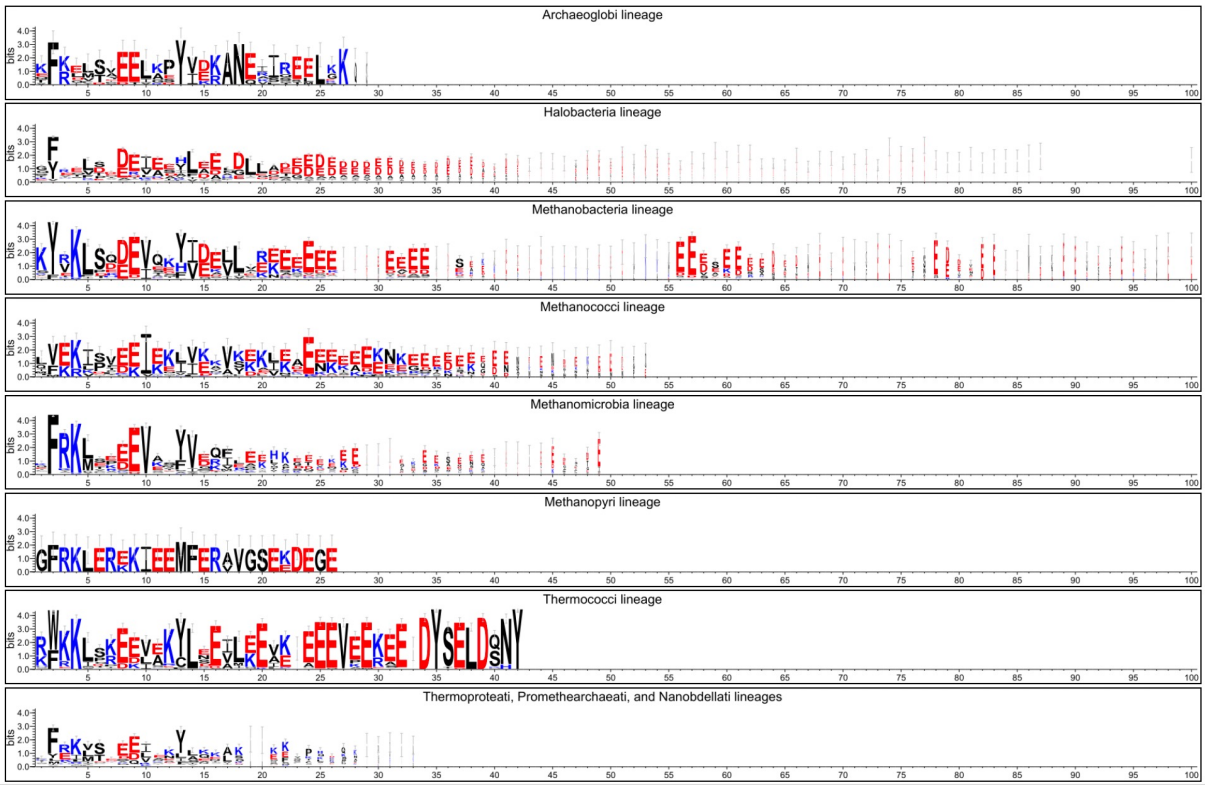
Supplementary Figure 13: Apa does not enhance the hydrolysis of Suc-LLVY-AMC by the 20S CP.
The degradation of Suc-LLVY-AMC was monitored by measuring fluorescence at 460 nm in 50 mM Tris pH 8.00 + 300 mM NaCl at 55 °C. 10 μM substrate was incubated with 70 nM 20S CP and the indicated molar ratios of Apa. Error bars are not shown for visualization purposes.



Supplementary Figure 14: Chaperone activity of Apa toward the model substrate citrate synthase.
a, First derivative of the fluorescence ratio measuring the denaturation of citrate synthase in the presence of Apa over a broad temperature range, measured with a NanoTemper nanoDSF machine. **b,** First derivative of the light scattering measurements of citrate synthase in the presence of Apa over a broad temperature range. Titration ratios of Apa over citrate synthase show that the presence of Apa slows down its thermal-dependent unfolding and aggregation.

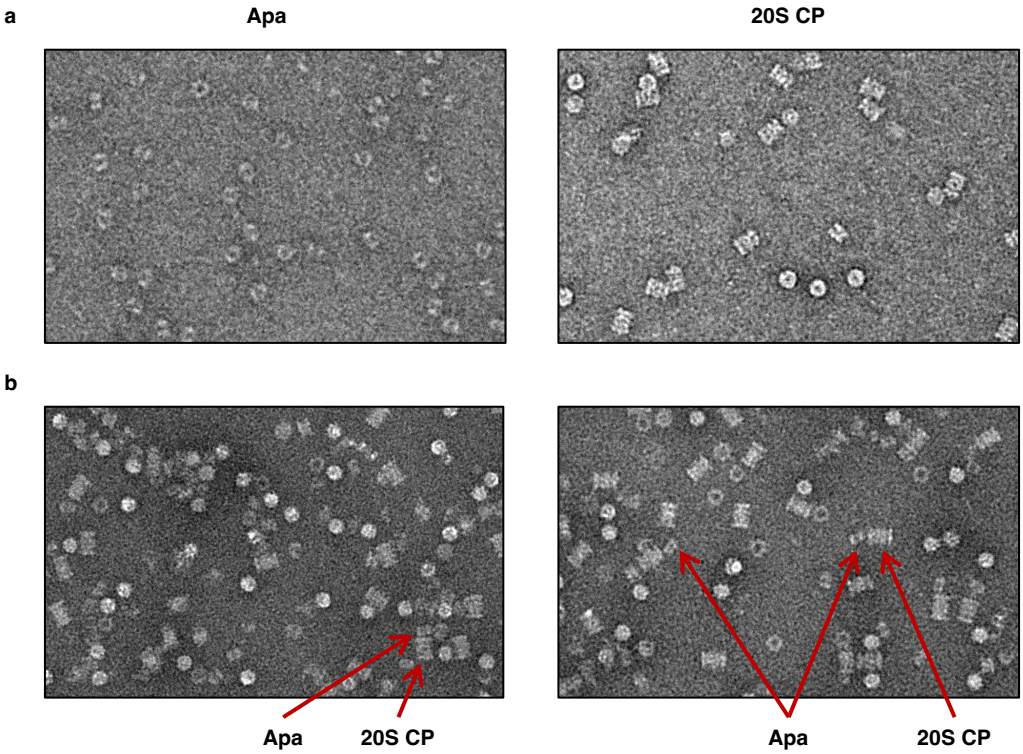


Supplementary Figure 15: Phylogeny of Apa homologue proteins.
The tree was built from 116 sequences and 294 amino acid positions using Maximum Likelihood. The tree has been rooted on the clade encompassing the *Pyrococcus* genus, in agreement with the species tree inferred from ribosomal proteins. Circles at nodes correspond to ultrafast bootstrap values greater than 90% (1,000 replicates). The scale bar represents the average number of substitutions per site.



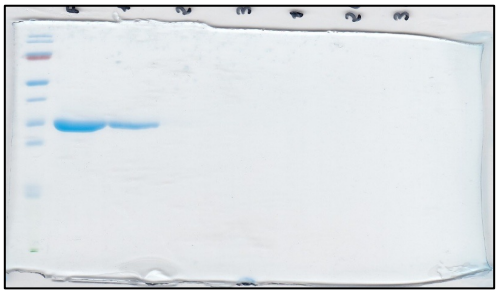
Supplementary Figure 16: Conservation of the C-terminal extension of proteasome α subunits in major archaeal lineages.

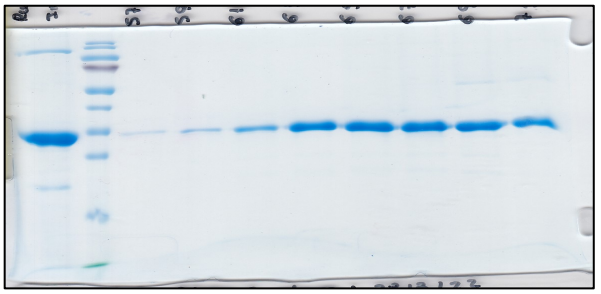
Pictures correspond to Weblogos. The height of the stack indicates the sequence conservation at that position, while the height of a letter within a given stack indicates the relative frequency of the corresponding amino acid at that position. The relative position of amino acids is conserved among logos.



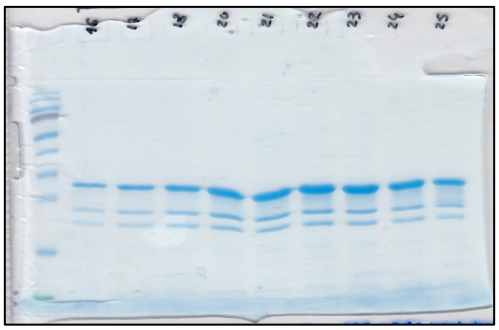
Supplementary Figure 17: Negative staining TEM of Apa-20S CP complexes.
a, Example micrographs of fractions containing only Apa or the 20S CP. **b,** Example micrographs of fractions corresponding to complexes. Tilting or partial dissociation is observable in most assemblies, with only a few views showing properly assembled complexes.

a

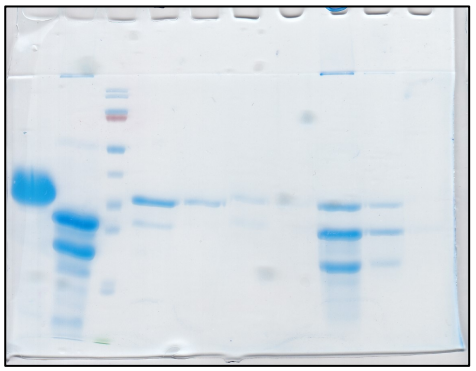




b

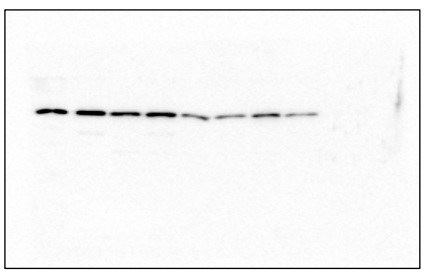


c



d

Ratio	1:0:0	1:5:0	1:0:2	1:5:2	1:5:2	1:5:2	1:5:2	1:5:2
CS	+	+	+	+	+	+	+	+
Apa untag	-	+	-	+	-	+	-	+
20S CP	-	-	+	+	+	+	+	+
Time	60	60	60	15	30	60	30	60
ATP	-	-	-	-	-	-	+	+



Supplementary Figure 18: Unprocessed gel images.

a, Left: Untagged Apa purification gel used in Figure S1. Right: His-tagged Apa purification gel used in Figure S1. **b,** 20S CP purification gel used in figure S1. **c,** Left: control pull-down gel with Apa against resin. Right: Apa-20S CP pull down gel. **d,** CS degradation assay Western-blot. Left: table with proportions of proteins. Right: corresponding bands from left to right.

Supplementary Table 1: Primers and cloning strategies for recombinant expression.

Plasmid name	Forward -primer *	Reverse-primer *	Cloning strategy
pET30a-Apa-H6	CAATCATATGAAAG CCAGCAAAC	CAATCTCGAGACGCA GCAGCATACCAAA	Synthetic gene of full-length *P. abyssi* Apa (M^{1}-R^{300}) optimized for expression in *E. coli* inserted into the pET30a (+) vector between the NdeI and XhoI restriction sites, followed by a 6x-His tag.
pET30a-Apa	CAATCATATGAAAG CCAGCAAAC	CAATCTCGAGTTAAC GCAGCAGC	Modification of pET30a-Apa-H6 plasmid with the insertion of a stop codon before the XhoI restriction site and the His tag.

* Restriction site sequences for directional cloning are underlined.

Supplementary Table 2: Sequencing primers for protein expression vectors

Primer name	Parental DNA	Primer sequence (from 5' to 3')	Comments
T7 promoter (fwd)	pET vectors	TAATACGACTCACTATAGGG	Forward primer to sequence the N-terminal end of the clonings performed and inserted into pET vectors.
T7 terminal (rev)	pET vectors	GCTAGTTATTGCTCAGCGG	Reverse primer to sequence the C-terminal end of the clonings performed and inserted into pET vectors.

Supplementary Table 3: Recombinant *Pyrococcus abyssi* proteins

Recombinant proteins	Restriction enzymes	Protein sequence *	Modifications	Extinction coefficient ($M^{-1}.cm^{-1}$)	MW (Da)
Pa20S_α	NdeI/XhoI	M^{1}+Y^{260}		20525**/20400***	28994.3
Pa20S_α_E244A_D252A_E255A_D257A	NdeI/XhoI	M^{1}+Y^{260}	E244A_D252A_E255A_D257A	20525**/20400***	28790.2
Pa20S_α_E248Q_D252N_E255Q_D257N	NdeI/XhoI	M^{1}+Y^{260}	E248Q_D252N_E255Q_D257N	20525**/20400***	28990.3
Pa20S_β1-ΔN-Ctag_His	NdeI/XhoI	M+T^{7}+L^{197}+LE+ H_{6x}	Nter deletion (1-6) 6xHis C-ter	21430	22076.5
Pa20S_β2-ΔN	NdeI/XhoI	M+T^{11}+Q^{207}	Nter deletion (1-10)	18910	21626.1
Apa	NdeI/XhoI	M^{1}-R^{300}		25330	33479.5
Apa-Ctag_His	NdeI/XhoI	M^{1}-R^{300}+LE+ H_{6x}	6xHis C-ter	25330	34544.6

* Peptide sequence of the expressed protein. Amino acids derived from the construct are in bold. See also Supplementary Fig. 1.
**assuming all pairs of Cys residues form cystines
***assuming all Cys residues are reduced

Supplementary Table 4: SEC-MALS peak analysis.

	Peak 1 (P1)				Peak 2 (P2)			
Recombinant proteins	Mw (kDa)	Mn (kDa)	Polydispersity (Mw/Mn)	Mass fraction (%)	Mw (kDa)	Mn (kDa)	Polydispersity (Mw/Mn)	Mass fraction (%)
Pa20S	745.2 (±1.3%)	744.3 (±1.3%)	1.001 (±1.9%)	73.5	1549.9 (±1.4%)	1511.3 (±1.4%)	1.026 (±1.9%)	26.5
Apa	250.8 (±1.4%)	250.5 (±1.4%)	1.001 (±2.0%)	93.4	416.8 (±1.4%)	416.5 (±1.4%)	1.001 (±2.0%)	6.6
Pa20S and Apa in complex	1033.4 (±1.4%)	1032.0 (±1.4%)	1.001 (±1.9%)	92.3	1698.6 (±1.3%)	1692.5 (±1.3%)	1.004 (±1.9%)	7.7

Supplementary Table 5: Crystallographic data collection, refinement and validation statistics.

	Native Apa	TbXo4 derivative
PDB code	30PG	N/A
Data collection and processing		
Source/Beamline	SOLEIL/Proxima1	SOLEIL/Proxima1
Spacegroup	P 3_2 1 2	P 3_2 1 2
Unit cell dimensions		
a, b, c (Å)	129.36, 129.36, 303.54	129.32, 129.32, 302.66
α, β, γ (°)	90, 90, 120	90, 90, 120
Wavelength (Å)	1.033	1.649
Resolution (Å)	49.23 – 1.95	49.16 – 2.50
(Highest resolution shell)	2.05 – 1.95	2.63 – 2.50
No. measured reflections	4421703	3991066
No. unique reflections	212021	100708
Rmerge	0.111 (1.013)	0.111 (0.594)
Rpim	0.025 (0.233)	0.018 (0.107)
CC(1/2)	1.000 (0.890)	0.999 (0.947)
I/σ(I)	22.8 (3.9)	29.8 (6.9)
Completeness (%)	99.9 (99.2)	99.9 (99.2)
Multiplicity	20.9 (19.7)	39.6 (31.1)
Refinement		
No. reflections	211063	N/A
Rwork/Rfree	0.1499/0.1806	N/A
No. of non-H atoms	18891	
Protein residues	16747	N/A
Ligand/ion	132	
Water	2012	
B factors (Å^2)	31.58	
Protein	30.57	N/A
Ligand/ion	55.41	
Water	38.48	
R.m.s. deviations		
Bond lengths (Å)	0.013	N/A
Bond angles (°)	1.210	
Clashscore	4.93	N/A
Ramachandran plot		
Favored (%)	99.38	N/A
Allowed (%)	0.62	
Disallowed (%)	0	

Supplementary Table 6: SAXS data and analysis.

Data collection and processing	
Instrument/source	beamline BM29 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France)
Wavelength (Å)	0.99
Beam energy (keV)	12.5
Beam size (μm)	~ 200 μm x 100 μm
Sample-to-detector distance (mm)	2870
q measurement range (Å^{-1})	0.044 - 0.52
Absolute scaling method	Comparison with scattering from 1 mm pure water
Normalisation	To transmitted intensity from beam stop counter
Exposure time	10 x 1 s frames
Sample temperature (°C)	25
Software employed	
SAXS data reduction	IspyB
Basic analysis (Guinier)	PRIMUS from ATSAS 2.8.3
Fitting with atomic models	CRYSOL from ATSAS 2.8.3
Mass from Vc (kDa)	4.2 mg/mL solution: 247.3 (insert ratio with theoretical mass)
Guinier analysis	
Rg (Å)	45.4 ± 0
$I(0)$ (cm^{-1})	670.6 ± 0.4
qR_g min,max	0.48, 1.23
CRYSOL (no constant subtraction)	
Crystal structure	Native Apa (PDB 30PG)
χ^2	3.48
Calculated Rg (Å)	43.9

Supplementary Table 7: Cryo-EM data collection, refinement and validation statistics.

	20S Focused map	Apa in complex with 20S Focused map	Apa alone
EMDB code	EMD-54518	EMD-54519	EMD-55799
PDB code	9S34	9S34	9TCW
Data collection and processing			
Magnification	36000	36000	36000
Voltage (kV)	200	200	200
Electron exposure (e–/Å^2)	40	40	40
Defocus range (μm)	-1 to -2.5	-1 to -2.5	-1 to -2.5
Pixel size (Å)	1.145	1.145	1.145
Symmetry imposed	C7	C7	C7
Initial particle images (no.)	534,489	534,489	1,849,407
Final particle images (no.)	50,092	50,092	1,722,565
Map resolution (Å) **0.143** FSC threshold	3.15	3.13	2.71
Map resolution range (Å)	3.15-5	3.13-5	2.71-4
Refinement			
Initial model used (PDB)	N/A	N/A	N/A
Model resolution (Å)	N/A	N/A	N/A
Model resolution range (Å)	N/A	N/A	N/A
Map sharpening *B* factor (Å^2)	-100.7	-91.7	-96.7
Model composition			
Non-hydrogen atoms	29400	29400	16492
Protein residues	3759	3759	2100
Ligands	0	0	0
B factors (Å^2)			
Protein	79.11	79.11	57.78
Ligand	N/A	N/A	N/A
R.m.s. deviations			
Bond lengths (Å)	0.007	0.007	0.006
Bond angles (°)	1.153	1.153	0.873
Validation			
MolProbity score	0.96	0.96	0.98
Clashscore	1.97	1.97	2.11
Poor rotamers (%)	0	0	0.2
Ramachandran plot			
Favored (%)	98.87	98.87	98.66
Allowed (%)	1.13	1.13	1.34
Disallowed (%)	0	0	0

Supplementary Table 8: complete data from BLI experiments.

Experiment	WT Pa20S + Apa	WT Pa20S + Apa; duplicate	Pa20S_α_E244A _D252A_E255A_ D257A + Apa	Pa20S_α_E244A _D252A_E255A_ D257A + Apa; duplicate	Pa20S_α_E248Q _D252N_E255Q_ D257N + Apa	Pa20S_α_E248Q _D252N_E255Q_ D257N + Apa; duplicate
Binding model	2:1	2:1	2:1	2:1	2:1	2:1
K_D (M)	2.59E-10	2.47E-10	1.55E-07	1.65E-07	1.04E-07	1.26E-07
K_D2	3.05E-08	5.44E-08	9.68E-08	1.38E-07	1.99E-07	2.55E-07
K_D Error	4.92E-12	3.69E-12	4.25E-09	3.13E-09	3.14E-09	2.54E-09
K_D2 Error	4.77E-09	9.12E-09	2.08E-09	2.59E-09	4.61E-09	5.50E-09
ka (1/Ms)	1.16E+06	1.13E+06	2.70E+04	1.71E+04	2.68E+04	2.97E+04
ka2	9.74E+05	6.53E+05	1.04E+06	7.71E+05	7.68E+05	6.78E+05
ka Error	1.28E+04	1.01E+04	7.24E+02	3.11E+02	7.90E+02	5.84E+02
ka2 Error	1.50E+05	1.09E+05	2.10E+04	1.37E+04	1.67E+04	1.37E+04
kdis (1/s)	2.99E-04	2.80E-04	4.19E-03	2.81E-03	2.79E-03	3.73E-03
kdis2	2.97E-02	3.55E-02	1.01E-01	1.07E-01	1.53E-01	1.73E-01
kdis Error	4.63E-06	3.36E-06	2.27E-05	1.54E-05	1.81E-05	1.79E-05
kdis2 Error	8.44E-04	7.91E-04	7.53E-04	6.45E-04	1.24E-03	1.34E-03
Full X^2	0.4077	0.2744	0.3447	0.6228	0.4144	0.3392
Full R^2	0.9968	0.998	0.9904	0.9933	0.9911	0.9940