



HAL
open science

The cryo-EM structure of the chloroplast ClpP complex

Ning Wang, Yifan Wang, Qian Zhao, Xiang Zhang, Chao Peng, Wenjuan Zhang, Yanan Liu, Olivier Vallon, Michael Schroda, Yao Cong, et al.

► **To cite this version:**

Ning Wang, Yifan Wang, Qian Zhao, Xiang Zhang, Chao Peng, et al.. The cryo-EM structure of the chloroplast ClpP complex. *Nature Plants*, 2021, 7 (11), pp.1505 - 1515. 10.1038/s41477-021-01020-x . hal-03452775

HAL Id: hal-03452775

<https://hal.science/hal-03452775>

Submitted on 14 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **The cryo-EM structure of the chloroplast ClpP complex**

2 Ning Wang^{1,2*}, Yifan Wang^{2,3,*}, Qian Zhao¹, Xiang Zhang³, Chao Peng⁴, Wenjuan
3 Zhang^{1,2}, Yanan Liu^{1,2}, Olivier Vallon⁵, Michael Schroda⁶, Yao Cong^{3,#}, Cuimin Liu^{1,2,#}

4

5 ¹ State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of
6 Genetics and Developmental Biology. The Innovative Academy of Seed Design,
7 Chinese Academy of Sciences, Beijing, 100101, China

8 ² University of Chinese Academy of Sciences, Beijing, 100101, China

9 ³ State Key Laboratory of Molecular Biology, National Center for Protein Science
10 Shanghai, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence
11 in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China;
12 Shanghai Science Research Center, Chinese Academy of Sciences, Shanghai, China
13 201210

14 ⁴ National Facility for Protein Science in Shanghai, Zhangjiang Lab, Shanghai
15 Advanced Research Institute, CAS, Shanghai 201210, China

16 ⁵ UMR7141 CNRS-Sorbonne Université, Institut de Biologie Physico-Chimique, Paris
17 75005, France

18 ⁶ Molecular Biotechnology & Systems Biology, TU Kaiserslautern, Paul-Ehrlich Straße
19 23, D-67663 Kaiserslautern, Germany

20

21 * these authors contributed equally to this work

22 # Correspondence to: cmliu@genetics.ac.cn;

23 cong@sibcb.ac.cn

24 **Keywords: Chaperone; Protease; Protein homeostasis; Photosynthesis; Rubisco;**
25 ***Chlamydomonas reinhardtii***

26

27 **Abstract**

28 Protein homeostasis in plastids is strategically regulated by the protein quality control
29 system involving multiple chaperones and proteases, among them the Clp protease.
30 Here we determined the structure of the chloroplast ClpP complex from
31 *Chlamydomonas reinhardtii* by cryo-EM. ClpP contains two heptameric catalytic rings
32 without any symmetry. The top ring contains one ClpR6, three ClpP4 and three ClpP5
33 subunits while the bottom ring is composed of three ClpP1_C subunits and one each of
34 the ClpR1-4 subunits. ClpR3, ClpR4 and ClpT4 subunits connect the two rings and
35 stabilize the complex. The chloroplast Cpn11/20/23 co-chaperonin, a co-factor of
36 Cpn60, forms a cap on the top of ClpP by protruding mobile loops into hydrophobic
37 clefts at the surface of the top ring. The co-chaperonin repressed ClpP proteolytic
38 activity *in vitro*. By regulating Cpn60 chaperone and ClpP protease activity, the co-
39 chaperonin may play a role in coordinating protein folding and degradation in the
40 chloroplast.

41

42 **INTRODUCTION**

43 Sophisticated pathways within cells control and regulate the biogenesis, trafficking
44 and degradation of proteins to ensure protein homeostasis (proteostasis). These
45 pathways belong to the protein quality control (PQC) system that involves the
46 participation of various chaperones and proteases¹⁻⁵. Molecular chaperones act as key
47 components of the PQC system. They assist protein folding when new polypeptide
48 chains emerge from ribosomes, when proteins have translocated through membranes,
49 and when proteins become misfolded. Proteases are another important component of
50 the PQC system that recognize and degrade substrate proteins which cannot fold or
51 refold to the native state.

52 The molecular chaperone chaperonin is a tetradecameric cylinder formed by two
53 rings stacked back to back, whose central cavity is used to shield a bound unfolded
54 protein from the crowded environment and to assist its folding to the native state in an
55 ATP-dependent process⁶. A co-chaperonin complex seals the central cavity of group I

56 chaperonins to form the isolated folding environment. In bacteria, this lid is formed by
57 the homo-heptameric GroES co-chaperonins. In plastids, two types of co-chaperonin
58 subunits exist, Cpn10 and Cpn20, the latter formed by two co-chaperonin domains in
59 tandem. In addition to its role as a co-chaperonin for Cpn60, Cpn20 has been shown in
60 *Arabidopsis* to mediate superoxide dismutase (FeSOD) activation independent of the
61 Cpn60 chaperonin ⁷. In addition, the expression level of the *Cpn20* gene influenced the
62 physiological function of ABA signaling during seed germination and promoted
63 stomatal closure without Cpn60 involvement ^{8,9}. These results indicate that the Cpn20
64 protein might exhibit biochemical functions in addition to that as a co-factor for Cpn60
65 in protein folding.

66 In bacterial cells or endosymbiotic organelles, proteins are degraded by many
67 types of proteases including serine-, aspartate-, and threonine-proteases ^{10,11}. The Clp
68 protease is a conserved serine protease that consists of the ClpP core complex and
69 assistant chaperones ¹². The first crystal structure of the *E. coli* ClpP core complex
70 revealed a barrel composed of two stacked heptameric rings, with one catalytic triad in
71 each subunit ¹³. The degradation of protein substrates by the Clp protease in *E. coli* is
72 assisted by two hexameric AAA+ (ATPase Associated with multiple cellular Activities)
73 chaperones, ClpX and ClpA, both of which can bind to the ClpP core complex via one
74 of its surfaces ¹⁴. Fueled by ATP hydrolysis, substrate proteins are unfolded by ClpX or
75 ClpA and threaded into the central cavity of the ClpP core complex ¹⁴, where the
76 substrate is degraded. Recent studies have explored the functional significance of the
77 symmetry mismatch between the AAA+ chaperone hexamer and the heptamer of the
78 ClpP core complex in *E. coli* ¹⁵⁻¹⁸. In particular, it has been proposed to induce a
79 rotational movement between the AAA+ ATPase and the ClpP core and to participate
80 in substrate transport ¹⁶.

81 In mitochondria, the ClpP core consists of 14 identical subunits that are similar to
82 their bacterial homologs ¹⁹. In contrast, the ClpP core complex in chloroplasts is made
83 up of three types of subunits termed ClpP, ClpR, and ClpT. ClpR subunits are
84 homologous to ClpP but cannot contribute to catalysis because they lack one or more

85 of the active residues of the Ser-His-Asp catalytic triad. ClpT proteins are completely
86 different proteins of around 20 kDa that share homology with the N-terminal domain
87 of the AAA+ chaperones. ClpT subunits have been proposed to link the two heptameric
88 rings and thus maintain the stability of the tetradecameric ClpP core²⁰⁻²³. In the green
89 alga *Chlamydomonas reinhardtii*, the subunits of the chloroplast ClpP core complex are
90 encoded by three *ClpP* genes (*clpP1*, *CLPP4*, *CLPP5*, the former chloroplast-encoded),
91 five *ClpR* genes (*CLPR1-4*, *CLPR6*) and two *ClpT* genes (*CLPT3*, *CLPT4*)^{24, 25}. The
92 plastidial *clpP1* gene produces ClpP_{1H}, which can be further processed by unknown
93 peptidases to generate ClpP_{1N}, ClpP_{1C} and ClpP_{1C'} subunits which are part of the ClpP
94 core complex²⁶. Translational attenuation of *clpP1* in *Chlamydomonas* led to the
95 stabilization of misassembled photosynthetic enzymes²⁷, while its conditional
96 repression caused serious autophagy responses and activated the protein quality control
97 system²⁸. In *Arabidopsis* chloroplasts, the ClpP complex consists of five ClpP type
98 subunits (ClpP3-6 and ClpP1), four ClpR type subunits (ClpR1-4) and two ClpT type
99 subunits (ClpT1-2)^{20, 22, 29}. Loss of *ClpP5* gene function was embryo-lethal³⁰.
100 Although the R-type subunits were considered unable to degrade substrate proteins,
101 either *ClpR2* or *ClpR4* gene knockout resulted in delayed embryogenesis suggesting
102 their functional importance³⁰. These results indicate that the Clp protease plays an
103 essential role in maintaining protein homeostasis in chloroplasts, even though some
104 subunits can be substituted by others. It was observed previously that the co-chaperonin
105 Cpn20 interacts with the ClpP complex²⁰, but how this interaction takes place and
106 whether it has functional implications are not known.

107 Recently, several structures of the Clp machinery composed of core complex and
108 AAA+ chaperone have been resolved by cryo-EM. However these studies focused on
109 bacterial Clp machineries, while there is no high-resolution structure of chloroplast Clp,
110 yet^{15-17, 31}. In this study, we solved the cryo-EM structure of the *Chlamydomonas*
111 chloroplast ClpP core complex. The complex is asymmetrical and polypeptide chains
112 could be assigned to a known Clp gene. Furthermore, the co-chaperonin complex was
113 observed to bind to the top ring of the ClpP core without inducing a major

114 conformational change. We hypothesize that co-chaperonins act as regulatory factors
115 in chloroplast proteostasis.

116

117 **RESULTS**

118 **Co-chaperonins inhibit ClpP proteolytic activity**

119 In previous work, the ClpP complex has been purified from *Chlamydomonas reinhardtii*
120 chloroplasts by Strep-tag affinity purification and the complex subunits have been
121 assigned to P-, R- and T-type by mass spectrometry³². Based on that work, we improved
122 the purification strategy by combining affinity purification with multiple
123 chromatography steps to obtain large amounts of highly purified *Chlamydomonas* ClpP
124 complexes. The purified complexes were separated on a 12%-18% SDS gel and on a
125 native gel and proteins were visualized by Coomassie staining (Figs. 1A and 1B). In the
126 native gel, the purified ClpP complex migrated as a diffuse band with a molecular mass
127 below the 820 kDa of oligomeric Cpn60 (Fig. 1B). By Asymmetric Flow Field Flow
128 Fractionation (AFFFF) we calculated a molecular mass of 549 kDa for the ClpP
129 complex, which is much larger than the 240 kDa of *E. coli* ClpP (Extended Data Fig.
130 1A). Each visible protein band in the SDS gel was cut out and analyzed by mass
131 spectrometry, and the bands identified as Clp or co-chaperonin subunits are marked
132 (Fig. 1A) (Supplementary Dataset 1). Though the migration pattern of Clp subunits was
133 slightly different from the previous report due to difference in protein preparation and
134 SDS-PAGE conditions, the same major Clp subunits were identified, except for ClpT3
135 (Supplementary Table 1)³². In addition to the Clp subunits, the three co-chaperonin
136 proteins Cpn11, Cpn20, and Cpn23 were identified with notable abundance
137 (Supplementary Table 1)(Supplementary Dataset 1). The presence of *clpP1* gene
138 products ClpP1_H, ClpP1_C and ClpP1_{C'} and of Cpn20 in the purified ClpP complex was
139 confirmed by immunoblotting using antisera against the Strep-tag and against Cpn20
140 (Fig. 1C). It is of note that the two bands detected for both ClpP1_C and ClpP1_{C'} might
141 result from the differentiated peptidase processing. To confirm the interaction of the
142 Clp complex with the co-chaperonins, we performed a co-immunoprecipitation

143 experiment using Cpn20 antiserum. As shown in Fig. 1D, Clp subunits ClpP1_C, ClpP4,
144 ClpR6, and ClpT4 were detected in the immunoprecipitate by immunoblotting. The
145 interaction between the ClpP complex and chaperonins *in vitro* was further validated
146 by size exclusion chromatography (Fig. 1E). Some of recombinantly produced Cpn20
147 co-migrated with the *Chlamydomonas* ClpP complex, and co-migration was even more
148 marked with a mixture of the three co-chaperonins Cpn11/20/23. Note that Cpn20 was
149 also detected after long time exposure in the ClpP-only fractions (CrClpP row) because
150 the co-chaperonin complex co-purifies with endogenous ClpP. No co-migration of Cpn
151 subunits was observed with the *E. coli* ClpP complex, and the *E. coli* co-chaperonin
152 GroES co-migrated neither with *Chlamydomonas* nor *E.coli* ClpP. These results
153 indicate that the interaction between the co-chaperonin and the ClpP complex can be
154 reconstituted *in vitro* and is a specific feature of the chloroplast system. This conclusion
155 is in line with the finding that *Arabidopsis* Cpn20 interacts with the *Arabidopsis* Clp
156 complex *in vivo* ²⁰.

157 Beta-casein is a commonly used model substrate for the Clp protease (Caseino-
158 Lytic Peptidase) ^{33, 34}. To analyze protease activities of purified ClpP complexes, we
159 performed protein degradation assays. ClpP of *E.coli* (EcClpP) could not degrade β -
160 casein, unless 4 μ M of the ClpP activator acyldepsipeptide (ADEP) was present in the
161 reaction (Figs. 2A, 2B and Extended Data Fig. 2A). ADEP increases the interaction
162 between ClpP monomers, competes with the Clp ATPases for their binding sites on
163 ClpP, and triggers a closed- to open-gate transition of the substrate entrance pore, which
164 is otherwise tightly closed ^{35, 36}. In contrast, purified *Chlamydomonas* ClpP was able to
165 degrade β -casein without need for ADEP. ADEP slightly accelerated protein
166 degradation by CrClpP, but only at a high concentration of 18 μ M, not of 4 or 8 μ M
167 (Figs. 2A, 2B and Extended Data Fig. 2A). Given the specific interaction between
168 *Chlamydomonas* ClpP and the co-chaperonin complex (Fig. 1), we wondered whether
169 co-chaperonins had an effect on ClpP proteolytic activity. Recombinant co-chaperonins
170 Cpn20 and Cpn11/20/23 both slowed down the proteolytic activity of CrClpP (Figs. 2C
171 and 2D), but co-chaperonin GroES had no effect (Extended Data Figs. 2A and 2B). This

172 inhibitory effect was overcome by the addition of 18 μ M ADEP (Figs. 2C and 2D).
173 CrClpP hydrolyzed β -casein into several fragments of lower molecular mass ranging
174 from 15 to 20 kDa, which were not observed with EcClpP (Extended Data Fig. 2A).
175 Hence, casein degradation by CrClpP appeared less processive than by EcClpP+ADEP.
176 Moreover, Cpn20 was still bound to the CrClpP complex after incubation with 18 μ M
177 ADEP (Extended Data Fig. 2C), indicating that either ADEP and the co-chaperonin
178 bind to different sites on the ClpP core complex or ADEP cannot expel the co-
179 chaperonin from its binding sites. As tiny amounts of AAA+ chaperones might have
180 co-purified with CrClpP, we cannot conclude that CrClpP is proteolytically active in
181 the complete absence of AAA+ chaperones. However, treatment with
182 hexokinase/glucose, which will remove all ATP from the buffer, had no effect on casein
183 degradation (Extended Data Fig. 2A), suggesting that CrClpP can degrade this substrate
184 in a completely energy-independent process.

185 **Cryo-EM structure of ClpP complex and subunit assignment**

186 To analyze structural features of the *Chlamydomonas* plastidic ClpP complex, we
187 performed cryo-EM single particle analysis on the system. Purified ClpP complexes
188 were applied to specimen grids and vitrified for cryo-EM analysis (Extended Data Fig.
189 3 and 4). After 3D classification and iterative refinement, the optimally visualized
190 particle groups were selected to generate two major groups, named ClpP-S1 and ClpP-
191 S2, at resolutions of 3.3 \AA and 3.6 \AA , respectively. The main difference between these
192 two groups was the appearance of a small cap in the latter particles (Figs. 3A-B and
193 Extended Data Figs. 3C, 4A and 4D). Focused 3D classification, particle re-extraction,
194 re-centering and refinement of the cap, which was later identified as co-chaperonin
195 Cpn11/20/23, generated particles with better structural features and more complete
196 mobile loops at 4.8 \AA resolution (Extended Data Figs. 3C and 4B). We further combined
197 the 4.8 \AA Cpn11/20/23 and ClpP-S2 maps using the *vop maximum* command in
198 Chimera and generated a ClpP-S2-composite map (ClpP-S2c)(Fig. 3B and Extended
199 Data Fig. 3C). The ClpP-S1 map displayed an obviously asymmetric structural
200 conformation with a height of about 108 \AA , a width of about 105 \AA and a length of

201 about 148 Å. The diameters of the central pore of ClpP-S1 are 28 Å and 20 Å for the
202 top and bottom rings in the cut away views, respectively (Fig. 3A). This difference in
203 pore size implied that there may exist a functional division between top and bottom ring
204 in terms of substrate admittance. The cap in the ClpP-S2 particles increased the height
205 to about 150 Å and the width to 118 Å, while the length was the same as ClpP-S1 (Figs.
206 3A and 3B). Overall, the resolution in the middle of the particles was higher than on the
207 surface (Extended Data Fig. 4D), especially for ClpP-S1 where the resolution in the
208 particle core reached about 3 Å, which allows the identification of amino acid side
209 chains.

210 To analyze the structural features of the CrClpP complex, the crystal structure of
211 the *E.coli* ClpP complex (PDB: 1TYF) was manually docked into the CrClp-S1 density
212 map (Fig. 3C)¹³. The EcClpP structure fits well into the ClpP-S1 map around the central
213 pore, with clear separation of double rings. Compared to the symmetrical EcClpP
214 structure, three prominent additional densities appeared in ClpP-S1 that were labeled
215 respectively as A1, A2 and A3. The A1 density is very large and locates adjacent to the
216 interface of the two rings like a handle, while densities A2 and A3 appear to be more
217 similar and locate to the peripheral region of the top ring.

218 In *Chlamydomonas*, three P-type (ClpP1, P4, P5), five R-type (ClpR1-4, R6) and
219 two T-type (ClpT3, T4) subunits constitute the ClpP complex. ClpP_{1H} is processed by
220 unknown peptidases to generate ClpP_{1N}, ClpP_{1C} and ClpP_{1C'} subunits, which are part
221 of the ClpP core complex. An alignment of amino acid sequences of P- and R-type
222 subunits with EcClpP revealed a conserved region corresponding roughly to amino
223 acids 27-175 in EcClpP (Extended Data Fig. 5), which served to build the structures of
224 the P- and R-type subunits (Extended Data Fig. 6A). The structure of ClpP_{1N} could not
225 be built after many trials. To assign the subunits into the ClpP core complex, the 14
226 subunits in the complex were designated as D1-D14 (Fig. 4A, left) and the individually
227 built structures were manually fit into the central region of the ClpP-S1 map. Some
228 specific sequences of ClpP and ClpR subunits were used for assigning them to the ClpP-
229 S1 map (Extended Data Fig. 5). Visualizations of map fitting of specific side chains

230 allowed us to identify subunit locations in the Clp-S1 map (Extended Data Fig. 6B).

231 After subunit assignment into the ClpP-S1 map, the individual subunit structures
232 were built manually with continuous map densities. The structures of the N- and C-
233 termini of some subunits could not be solved due to discontinuity in the densities.
234 Information on the resolved sequences is summarized in Supplementary Table 2. The
235 superimposition of the individual subunit structure onto its corresponding electron
236 density indicates that the models fit well into the corresponding maps (Extended Data
237 Fig. 7A). Overall, we could assign one ClpR6, three ClpP4, and three ClpP5 subunits
238 to the top ring (designated as P-ring), and one of each ClpR1-4 and three ClpP1_C
239 subunits to the bottom ring (Figs. 4A-C). Values of around 0.7 in the correlation
240 coefficient (CC) chart corroborate our model-to-map fitting for each subunit (Extended
241 Data Fig. 7B). Still, some regions in the ClpP structure could not be solved, as shown
242 by white regions in the core complex map density (Fig. 4B and Extended Data Fig. 7C).
243 Notice that no subunits or sequences could be fit into the additional A2 and A3 maps
244 (Extended Data Fig. 7C), including the long C-terminal sequences (Val196-Trp296) of
245 the adjacent ClpP4 subunit, which remains unsolved. Conversely, no density could be
246 ascribed to the large IS1 sequence characteristic of ClpP1_H^{24, 26}. The T4 subunit could
247 be assigned to the A1 density, which is located adjacent to the interface of the two rings
248 (Figs. 4B-C). The additional A1 map density, located at the interface of the two rings,
249 is formed by a very long helix originating from the C-terminus of the ClpR3 subunit in
250 the bottom ring, as well as by the T4 subunit and several amino acids from the ClpR6
251 subunit in the top ring (Fig. 4D). Some unassigned map density in A1 might be
252 contributed by the C-terminus of ClpR3, and by ClpT3 or ClpP1_N subunits. The A1
253 region connects the two rings like a handle and can be assumed to stabilize the ClpP
254 core complex. Moreover, the C-terminus of ClpR4 protrudes into a region next to the
255 ClpP5 subunits in the top ring, and this might contribute to stabilize the core complex,
256 as well (Fig. 4E).

257 The electrostatic potential is not equally distributed around the central pores of the
258 two rings through which substrates enter the catalytic chamber, suggesting a differential

259 affinity to substrates (Fig. 4F). The AAA+ chaperones interact with EcClpP by
260 protruding their flexible IGF loops into the hydrophobic clefts of ClpP which are
261 formed at the interface of two subunits¹⁶. Similarly, seven hydrophobic clefts were
262 found to form at the surface of the top ring of ClpP by amino acids from two adjacent
263 subunits. These hydrophobic clefts are arranged in a circular manner with seven-fold
264 symmetry and were observed only on the top ring (Fig. 4G). However, we cannot
265 completely rule out that similar hydrophobic clefts exist on the bottom ring, because
266 the models of the three ClpP1c subunits in the bottom ring are not complete. A
267 schematic model of the overall structure of the ClpP core complex with assigned
268 subunits is shown in Fig. 4A.

269 **The cap on the top of the ClpP core is the co-chaperonin**

270 We have shown that the co-chaperonin complex interacts with the ClpP core *in vivo* and
271 *in vitro* (Fig. 1). Previous work from us and others showed that the authentic co-
272 chaperonin complex *in vivo* consists of two Cpn20 subunits, one Cpn23 subunit, and
273 one Cpn11 subunit^{37,38}. Compared to Clp-S1, the ClpP-S2c map showed a dome-like
274 density located on the top of the ClpP core complexes (Fig. 3, labeled A4 in Extended
275 Data Fig. 8A) which we attributed to the co-chaperonin. Since Cpn20 and Cpn23 each
276 have two GroES-like domains, we split them into Cpn20-N, Cpn20-C, Cpn23-N, and
277 Cpn23-C regions after cleavage of transit peptide³⁹ and performed a sequence
278 alignment with Cpn11 and GroES (Fig. 5A). Sequences contributing to the roof of the
279 dome-shaped co-chaperonin complex were present in GroES, Cpn23-N, Cpn23-C and
280 Cpn20-C. This sequence was much shorter in Cpn11 and missing in Cpn20-N. Based
281 on these specific roof characteristics we could assign Cpn11, Cpn23, and two Cpn20
282 subunits (Cpn20-1 and Cpn20-2) to the co-chaperonin map (Fig. 5B). The CC chart
283 indicates a good model-to-map fitting for both Clp and co-chaperonin subunits
284 (Extended Data Fig. 8B) and the superimposition of the individual co-chaperonin
285 subunit structure onto its corresponding electron density indicates that the models fit
286 well into the corresponding maps (Extended Data Fig. 8C). We further refined the
287 model and fit five of the seven mobile loops into the densities that extended from the

288 bottom of the co-chaperonin dome (belonging to the Cpn20 and Cpn11 subunits, Fig.
289 5B, left) (Fig. 5B, left). No densities were observed for the loops of the Cpn23 subunit,
290 although the sequences are conserved (Fig. 5A).

291 Inspection of the map revealed that the co-chaperonin cap is tilted by about 4°
292 relative to the ClpP symmetry axis (Figs. 5C and 5F). This tilt increases the distance
293 between ClpP and Cpn23 versus ClpP and Cpn20, suggesting an intimate interaction at
294 the Cpn20 side. A similar interaction, but with a tilt angle of 11° , has been observed for
295 the asymmetrical EcClpP with its AAA+ chaperone ClpX¹⁶. While that interaction is
296 accompanied by a symmetry mismatch, this is not the case for the interaction between
297 co-chaperonin and ClpP, which share a 7-fold symmetry. In *E.coli*, the IGF loops of
298 ClpX insert into hydrophobic clefts located at the surface of EcClpP¹⁵. Close inspection
299 of the density map of the ClpP core complex in ClpP-S2c clearly revealed five extra
300 densities in the hydrophobic clefts of the top ClpP ring (red and orange in Fig. 5D).
301 Three of these densities were confidently identified as mobile loops stemming from
302 Cpn11, contacting the ClpP5/R6 interface, and Cpn20-1C and Cpn20-2C, both
303 contacting a ClpP5/P4 interface (red in Fig. 5E). The mobile loops from Cpn20-1N and
304 Cpn20-2N might contact ClpP clefts as well, but at a slightly outward location, which
305 could account for two more extra densities observed in clefts of the top ClpP ring
306 (orange in Fig. 5D; Extended Data Fig. 8D). An assignment of the densities in the two
307 cleft regions was complicated by the close-by unassigned A2/A3 densities. For the two
308 remaining hydrophobic clefts, in register with Cpn23, no extra map densities were
309 observed (broken red lines in Fig. 5D), indicating that there is no direct interaction
310 between Cpn23 and the ClpP core. In summary, the co-chaperonin appears to interact
311 with the ClpP core by inserting at least three mobile loops into hydrophobic clefts on
312 the surface of the top ring of ClpP (Fig. 5F), similar to the interaction of EcClpP with
313 its AAA+ chaperone ClpX. Notice that the same loops of the co-chaperonin mediate its
314 interaction with the Cpn60 chaperonin⁶. The interactions of the co-chaperonin with
315 ClpP, the co-chaperonin with Cpn60, and ClpP with AAA+ chaperones are thus
316 mutually exclusive.

317

318 **DISCUSSION**

319 **The chloroplast co-chaperonin has multiple functions**

320 Many lines of evidence indicate that Cpn20 functions as a co-chaperonin for Cpn60 to
321 assist protein folding in chloroplasts^{38, 40-42}. However, Cpn20 appears to exhibit
322 functions independent of the Cpn60 chaperonin: in *Arabidopsis*, Cpn20 has been shown
323 to play roles in abscisic acid (ABA) signaling and the activation of iron superoxide
324 dismutase (FeSOD)⁷⁻⁹. Here we show that a fraction of the *Chlamydomonas* chloroplast
325 Cpn11/20/23 co-chaperonin complex co-purified with the ClpP core complex, and this
326 interaction was robust enough to withstand several chromatography purification steps
327 (Fig. 1). This interaction was shown before in *Arabidopsis*²⁰. The complexes between
328 co-chaperonin and ClpP were confirmed by cryo-EM and *in vitro* reconstitution, and
329 the co-chaperonin was shown to slow down the proteolytic activity of chloroplast ClpP
330 (Figs. 2B and 5). The co-chaperonin caps the top-ring of the ClpP complex like a dome,
331 similar to its interaction with Cpn60. Thus, the co-chaperonin might inhibit ClpP
332 activity by blocking the entrance of unfolded proteins into the central cavity of the ClpP
333 core complex via the pore of the top ring. The co-chaperonin complex interacts with
334 ClpP and the Cpn60 chaperonin via the same mobile loops extending from the bottom
335 of the dome (Fig. 5E). Therefore, the Cpn20/23/11 co-chaperonin complex appears to
336 play a dual role in chloroplast protein quality control. Since its interaction with the
337 Cpn60 chaperonin contributes to protein folding, the binding of the co-chaperonin to
338 ClpP would simultaneously limit the efficiency of protein folding and inactivate a
339 substantial part of the chloroplast's degradation capacity. Releasing the co-chaperonin
340 from ClpP would activate both activities at the same time, which might be important
341 upon rapid challenges of proteostasis by environmental changes. In future studies, it
342 will be interesting to quantify the relative stoichiometry of the chaperonin, co-
343 chaperonin, and ClpP complexes in the *Chlamydomonas* chloroplast and their
344 interactions under various growth or stress conditions. It is of note that the very
345 abundant RbcL subunit of Rubisco is a substrate for both chaperonin and ClpP, and that

346 translational attenuation of *clpP1* is lethal in the presence of *rbcL* mutations preventing
347 its folding²⁷. While we initially postulated that the lack of sufficient ClpP would cause
348 poisoning of the chaperonin by an unfoldable substrate, we must now also consider an
349 additional hypothesis, that the mutant RbcL, by increasing the residence time of the co-
350 chaperonin on and ever-busy Cpn60, would deregulate ClpP proteolysis.

351 The fact that the three Cpn60 subunits were also detected by mass spectrometry in
352 our preparation (Supplementary Dataset 1) raises another possibility, namely that ClpP
353 and Cpn60 interact to regulate chloroplast protein homeostasis. Non-foldable substrate
354 proteins, released from the chaperonin, need to be recognized by Clp to be degraded to
355 prevent poisoning of the chaperonin. The co-chaperonin, by maintaining its interaction
356 with the non-foldable substrate, could help direct it to the protease. In line with this
357 idea, a cooperation between Trigger Factor, an ATP-independent ribosome-associated
358 chaperone, and ClpXP was recently reported to promote the degradation of some
359 substrates⁴³.

360 **Molecular architecture of chloroplast ClpP**

361 The first structure of the *E. coli* ClpP core complex, obtained by X-ray crystallography,
362 revealed 14 identical subunits in two stacked rings¹³. The ClpP structures solved
363 afterwards all consisted of a single or two types of ClpP subunits. In contrast, the
364 chloroplast ClpP core complex solved here combines 10 different subunits, either of
365 the P-type (harboring a functional catalytic site) or R-type (lacking catalytic residues)
366 or T-type (homologous to the N-domain of AAA+ chaperones), distributed with uneven
367 stoichiometry between two rings of different subunit composition. The subunit
368 composition of the two rings of plastid ClpP from *Arabidopsis* was reported previously
369 by the van Wijk and Clarke groups^{20, 22, 30, 44, 45}. The designated P-ring contains ClpP3,
370 4, 5 and 6 in a 1:2:3:1 ratio, while the designated R-ring consists of ClpP1 and ClpR1,
371 2, 3 and 4 in a 3:1:1:1:1 ratio. It was suggested that the two rings might exhibit different
372 proteolytic capacities because only three catalytic subunits are present in the R-ring
373 compared to seven in the P-ring. In the *Chlamydomonas* ClpP core complex reported
374 here, the top ring consists of three ClpP4, three ClpP5 and one ClpR6 subunits. Because

375 algal ClpP4 forms a sister clade to land plant ClpP3/ClpP4, and ClpR6 is clearly derived
376 from ClpP6 by loss of a catalytic residue ³², the top ring of *Chlamydomonas* is clearly
377 homologous to the *Arabidopsis* P-ring. Similarly, the bottom ring, with three ClpP1c
378 and one each of ClpR1, 2, 3, and 4 subunits (Fig. 4C) can be called the R-ring. Because
379 the diameter of the protein entrance pore in the P-ring is much larger than that in the R-
380 ring (Fig. 3A), we propose that the former is the functional entry site of substrates. Only
381 the P-ring was observed to interact with the co-chaperonin (Fig. 5), in line with a central
382 regulatory role for this interaction. Our structure supports the notion that substrates are
383 degraded inside a unique central cavity by a variety of active sites, possibly showing
384 distinct chemical specificities as in the proteasome ⁴⁶. Each subunit within each ring
385 may also be functionally unique in terms of substrate selection, delivery and unfolding,
386 and of interaction with the co-chaperonin. A good example is the specific interactions
387 of the ClpP4/5 clefts with Cpn20 subunits (Fig. 5E, Extended Data Fig. 8D).

388 Interesting differences can be noted between algal and land plant ClpP. With only
389 6 active subunits in its P-ring, the *Chlamydomonas* enzyme is slightly less asymmetrical
390 than that of land plants in terms of catalysis (both have 3 catalytic sites in the R-ring).
391 Importantly, the inactive ClpR6 subunit of the P-ring contributes, together with two
392 other alga-specific features, namely the C-terminal extension of the R-ring ClpR3 and
393 ClpT4, to the A1 side density that connects the two rings. In land plants, T-type Clp
394 subunits were suggested to stabilize the ClpP core by interacting with both rings ^{20, 21},
395 and it will be interesting to see if they form a structure similar to the A1 side-mass.
396 Note however that the *Chlamydomonas* ClpT3/4 are vastly different from the ClpT1/2
397 of land plants ³², to the extent that they may not even be orthologous. Given that the N-
398 terminal domain of AAA+ chaperones, to which ClpTs are homologous, play a role in
399 substrate selection, the possibility arises that their dual structural/functional role was
400 acquired several times in the evolution of the plastid Clp system. Furthermore, the C-
401 terminus of ClpR4 protrudes to the P-ring, which also might stabilize the ClpP core
402 complex. The massive contacts between the two rings may explain the high stability of
403 the ClpP core, exemplified by our observation that the two-rings could not be separated

404 by high salt treatment, in contrast to its homologs in *E. coli* and *Arabidopsis* plastids⁴⁵,
405 ⁴⁷.

406 **Proteolytic activity of the ClpP is regulated by the co-chaperonin**

407 Structures of several different AAA+ chaperones with their ClpP cores have emerged
408 recently, providing detailed information on the six-seven symmetry mismatch between
409 the two complexes, and how it impacts their interaction dynamics^{15, 16, 18}. ClpX is tilted
410 by 11° without major conformational changes upon binding to ClpP. As a result, the
411 symmetry axes of the protease and the AAA+ chaperone are not aligned so that the
412 translocation pathway for unfolded peptides is not straight but twisted. We also found
413 a tilt of 4° between the axes of the co-chaperonin and the ClpP core (Fig. 5C). However,
414 the fact that a single type of particle was obtained, with specific interactions between
415 co-chaperonin and Clp subunits, together with the absence of a symmetry mismatch,
416 suggest that the two complexes do not rotate and that the co-chaperonin instead operates
417 like a cap stably sealing the entrance to the ClpP core.

418 Because none of the three known chloroplast AAA+ chaperones (ClpC1, ClpD1,
419 ClpB3) were co-purified with ClpP in this and previous studies, they might not
420 cooperate with ClpP in the chloroplast as they do in bacteria and mitochondria. Since
421 the purified ClpP core complex was able to degrade the model substrate casein *in vitro*
422 with no need for addition of ADEP, it appears possible that chloroplast ClpP can
423 degrade proteins *in vivo* without the assistance of any type of chaperone. Regulation of
424 protease activity may be carried out by the co-chaperonin or by small-molecule
425 activators acting like ADEP.

426

427 **METHODS**

428 **Construction of expression plasmids**

429 The *ClpP* gene sequence from *E. coli* was amplified by PCR on genomic DNA and
430 cloned into the pHUE vector with restriction enzymes to generate *EcClp*-pHUE. The
431 produced *EcClp* protein contains ubiquitin and a 6×His tag at its N-terminus that
432 facilitated subsequent protein affinity purification. The forward and reverse primers

433 were 5'-GCGGATCCATGTCATACAGCAGCGGCGAACG-3' and 5'-
434 CCCAAGCTTTCAATTACGATGGGTCAGAATCGAATCGACCAG-3' containing
435 BamHI and HindIII restriction sites, respectively. The construction of the co-
436 chaperonin expression plasmids, GroES-pET11a, CrCPN20-pQlinkT and CrCPN11/
437 20/23-pQlinkT, were described earlier³⁷.

438 ***Chlamydomonas reinhardtii* strains and growth conditions**

439 *Chlamydomonas* strain ClpP1-strep (#8), used for the purification of the Clp complex
440 via Strep-tag affinity purification, has been describe previously²⁶ and is freely available
441 upon request. The strain was kept on solid TAP medium containing 100 µg/ml of
442 spectinomycin at 25°C under continuous illumination (40 µmol photo m⁻² s⁻¹) or a
443 photoperiod rhythm (12 h light/12 h dark).

444 **Purification of the ClpP complex from *C. reinhardtii***

445 Purification of the ClpP complex from *Chlamydomonas reinhardtii* was conducted as
446 described previously with some modifications⁴⁸. The ClpP1-strep strain was inoculated
447 into 24 L TAP liquid medium for 4 days under continuous light (40 µmol photo m⁻² s⁻¹).
448 When cell numbers reached around 6×10⁶ cells/mL, the cells were collected by
449 centrifugation at 3600 g for 6 min and resuspended in buffer A (20 mM Tris-HCl pH
450 8.0, 150 mM NaCl). The volume of the cell slurry was adjusted to 200 ml by the
451 addition of 1 mM EDTA, 1 mM PMSF, and two tablets of EDTA-free protease inhibitor
452 cocktail (Roche). Unless otherwise stated, subsequent steps were performed at 4°C. The
453 cells were sonicated and centrifuged at 36,000 g for 30 min to remove debris. 1 mg/L
454 avidin and 6 mM MgCl₂ were added to the supernatant to improve the Strep tag binding
455 efficiency with Strep-Tactin beads (Novagen). Next, the supernatant was further
456 clarified by ultracentrifugation at 150,000 g for 1 h (Beckman Coulter rotor, 70Ti).
457 Ammonium sulfate was slowly added to the supernatant at 25% saturation. After gently
458 stirring for 30 min, the produced aggregates were removed by centrifugation at 36000
459 g for 20 min. Buffer E (20 mM Tris-HCl pH 8.0, 1 mM DTT, 10% glycerol) was added
460 to the supernatant to dilute the ammonium sulfate concentration to 0.5 M and the
461 solution was applied to pre-equilibrated hydrophobic column after passing through a

462 0.22 μ m filter (Hitrap phenyl, GE Healthcare). Proteins were eluted by an ammonium
463 sulfate gradient from 0.5 M to 0 M with ten column volumes. Clp-containing fractions
464 were collected after visualization with SDS-PAGE and transferred to a Strep-Tactin
465 gravity column which was pre-equilibrated with buffer B (20 mM Tris-HCl pH 8.0, 100
466 mM NaCl, 1 mM DTT). The protein was eluted with buffer C (buffer
467 B+10% glycerol+2.5 mM Desthiobiotin (Novagen)) in four column volumes. The ClpP
468 complex was concentrated and subjected to a Superdex-200 column pre-equilibrated
469 with buffer D (20 mM Tris-HCl pH 8.0, 80 mM NaCl, 10% glycerol) and the desired
470 fractions were collected. Purified ClpP complexes were concentrated to ~2 mg/ml by
471 Amicon Ultra-15 Centrifugal Filter Units (Merck Millipore, Beijing China) with 100
472 kDa cut-off, supplemented with 10% glycerol and frozen at -80°C.

473 **Purification of recombinantly expressed proteins**

474 **EcClpP** The *EcClp*-pHUE plasmid was transformed into the *E. coli* BL21 (DE3) strain,
475 and then transferred to 4 L lysogeny broth (LB) medium containing 100 μ g/L ampicillin.
476 When *E. coli* was grown to an OD 600 of ~ 0.6, 1 mM isopropyl β -D-1-
477 thiogalactopyranoside (IPTG) was added to induce protein expression. After 4 hours,
478 cells were collected by centrifugation at 4000 g. Unless otherwise stated, all purification
479 steps were performed at 4°C. Cells were resuspended in lysis buffer (20 mM Tris-HCl,
480 pH 8.0, 300 mM NaCl, 1 mM DTT, 10 mM imidazole, 1 mM phenylmethylsulfonyl
481 fluoride (PMSF)) and lysed by sonication. Debris was removed by centrifugation at
482 36000 g for 40 min and the supernatant was passed through a 0.22 μ m filter followed
483 by transferring into the Ni-NTA gravity column. Protein was eluted with elution buffer
484 (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1 mM DTT, 250 mM imidazole) after two
485 washes with buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1 mM DTT, 25 mM
486 imidazole). Proteins in the eluted fractions were collected and digested with
487 deubiquitinating enzyme USP2cc (1:50 molar ratio) to remove the ubiquitin-6 \times His tag.
488 The digested solution was applied to Ni-NTA column again and the flow through was
489 collected which contained EcClpP protein without tag. The collected protein was
490 concentrated using Amicon Ultra-15 Centrifugal Filter Units (Merck Millipore, Beijing

491 China) with 100 kDa cut off and frozen at -80 °C.

492 **GroES, Cpn20 and Cpn11/20/23.** All co-chaperonins were purified using the same
493 method. Individual expression plasmid GroES-pET11a, Cpn20-pQlinkT or
494 Cpn11/20/23-pQlinkT was transformed into *E. coli* strain BL21 (DE3). Transformants
495 were picked and transferred to 6 L of LB medium containing 100 µg/L ampicillin.
496 When cells were grown to an OD 600 of ~ 0.6, 1 mM IPTG was added to induce protein
497 expression at 37°C for 4 hours. Cells were harvested by centrifugation at 3600 g for 6
498 min. Unless otherwise stated, all subsequent steps were performed at 4°C. Cell pellets
499 were resuspended in lysis buffer (30 mM Tris-HCl, pH 8.0, 60 mM NaCl, 1 mM DTT,
500 1 mM EDTA and 1 mM PMSF) sonicated. Lysates were centrifuged at 36,000 g for 30
501 min to remove cell debris. The supernatant was passed through a 0.22 µm filter followed
502 by transfer to a pre-equilibrated source 30Q column (GE Healthcare) with 1 ml/min
503 flow rate. Co-chaperonins were eluted by a linear salt gradient from 30 mM to 1 M
504 NaCl with 10-fold column volume. Fractions containing protein according to UV
505 absorption were separated by 15% SDS-PAGE and visualized by Coomassie staining.
506 Eluted proteins were concentrated, then injected into a pre-equilibrated Superdex75
507 column (GE Healthcare) with a flow rate of 0.8 ml/min. According to UV absorption,
508 the protein fraction was further analyzed with 15% SDS-PAGE and Coomassie staining.
509 The pure co-chaperonin proteins were concentrated with Amicon Ultra-15 Centrifugal
510 filters with 30 kDa cutoff, then flash-frozen in liquid nitrogen and stored at -80°C.

511 **Identification of CrClpP subunits by mass spectrometry**

512 The protein subunits of the CrClpP complex were separated on an SDS-PAGE gel and
513 visualized by Coomassie staining. Individual bands were cut out and digested by trypsin.
514 Liquid chromatography-mass spectrometry (LC-MS) was done on a Thermo Scientific
515 Q Exactive (QE) mass spectrometer at the Beijing Huada Protein R&D Center Co., Ltd.
516 (Beijing, P.R.China). The Q Exactive mass spectrometry data were searched against the
517 Phytozome v12.1(*Chlamydomonas reinhardtii*) database and NCBI-*Chlamydomonas*
518 (taxid: 3052) database using 15 ppm peptide mass tolerance and 20 m/z fragment mass
519 tolerance.

520 **Immunoprecipitation assays**

521 Protein A-sepharose beads coupled with CrCpn20 or strep-tag antibodies were pre-
522 equilibrated in lysis buffer containing 20 mM Hepes-KOH (pH7.5), 150 mM NaCl, 10
523 mM MgCl₂, 20 mM KCl and 2 mM EDTA, then incubated with total protein or stroma
524 protein which was prepared with the same lysis buffer under gentle stirring at 4°C.
525 Protein A beads were washed three times with lysis buffer containing 0.1% Tween 20.
526 Bound protein complexes were eluted with 2% SDS for 1 h at 4°C. The eluted proteins
527 were separated by 12% SDS-PAGE and analyzed by immunoblotting with serum in
528 1:3000 dilution.

529 **Analytical gel filtration**

530 The ClpP interaction with co-chaperonins was analyzed by analytical gel filtration as
531 described previously with some modifications³⁸. 1 μM ClpP and 2 μM co-chaperonin
532 were incubated for 30 min at 4°C in 20 mM MOPS-KOH, pH 7.5, 80 mM NaCl, 10
533 mM MgCl₂, 10 mM KCl, 1 mM DTT, 10% glycerol. Then the protein complexes were
534 loaded onto a Superdex 200 PC 3.2/10 column (GE Healthcare) at a flow rate of 0.05
535 ml/min. 50 μL fractions were collected and analyzed by immunoblotting with Strep-tag
536 and Cpn20 antibodies.

537 **Asymmetric flow field-flow fractionation with multi-angle light scattering** 538 **(AFFFF-MALS)**

539 50 μg protein of purified CrClpP were loaded into a AFFFF-MALS device with a flow
540 rate of 0.8 ml/min and a cross-flow rate of 2 ml/min using a 350 mm spacer and 10 kDa
541 RC membrane (Wyatt Technology, Santa Barbara, CA, USA). The monitor methods
542 employed a multiple-angle light scattering detector (DAWN HELEOS II, 658 nm;
543 Wyatt Technology), a UV detector (1100 series, 280 nm; Agilent Technologies In.,
544 Santa Clara, CA, USA) and a differential refractive index detector (Optilab rEX, 658
545 nm; Wyatt Technology)⁴⁹. The CrClpP molecular weight was calculated in the presence
546 of Dn/dc values equal to 0.185 ml/g.

547 **β-casein degradation assays**

548 Degradation of the substrate protein β-casein was visualized by Coomassie Blue R

549 staining after separating proteins via SDS-PAGE (12% acrylamide). The reaction
550 mixture contained 20 mM Tris-HCl pH 8.0, 120 mM NaCl, 10 mM KCl, 10 mM MgCl₂,
551 1 mM DTT, 10% glycerol, 0.4 μM EcClpP or CrClpP, 0.4 μM β-casein protein with or
552 without 0.4 μM co-chaperonin (GroES, Cpn20 and Cpn11/20/23). 4 to 18 μM protease
553 activator ADEP (#sc-397312, Santa Cruz) dissolved in DMSO was added to the
554 reaction as indicated. The reaction was performed at 30°C, and the aliquots were taken
555 at the indicated time points. The reaction was stopped by heating to 98°C for 10 min.
556 Each reaction was performed at least three times and densitometric quantification of β-
557 casein from the reactions was made with Photoshop CS6.

558 **Cryo-EM sample preparation and data collection**

559 Holey carbon grids (Quantifoil R2/1, 200 mesh) were plasma cleaned using a Solarus
560 plasma cleaner (Gatan), and an aliquote of 2 μl CrClpP sample was placed onto the
561 glow-discharged grid. Then the grid was flash-frozen in liquid ethane by a Vitrobot
562 Mark IV (Thermo Fisher Scientific). Movies were taken on a Titan Krios transmission
563 electron microscope (Thermo Fisher Scientific) equipped with a Cs corrector and
564 operated at an accelerating voltage of 300 kV with a nominal magnification of 18,000x
565 (Supplementary Table 3). Movies were collected by using a K2 Summit direct electron
566 detector (Gatan) in super-resolution mode (yielding a pixel size of 1.318 Å after 2 times
567 binning). Each movie was dose-fractioned into 38 frames and the exposure time was
568 7.6 s with 0.2 s for each frame, producing a total dose of ~38 e⁻/Å². The defocus value
569 of the data set varied from -0.8 to -2.5 μm. We employed the SerialEM automated data
570 collection software package to collect the images⁵⁰.

571 **Image processing and 3D reconstruction**

572 A total of 8,064 movies were applied for CrClpP structure determination. Unless
573 otherwise specified, single-particle analysis was mainly executed in RELION 3.1^{51, 52}.
574 All images were aligned and summed using MotionCorr2⁵³ and CTF parameters were
575 determined using CTFFIND4^{51, 52}. We obtained 1,351,977 particles by automatic
576 particle picking followed by manual checking, and 578,978 particles remained for
577 further processing after reference-free 2D classification. Through one round of 3D

578 classification, a ClpP-S1 dataset of 306,743 particles and a CrClpP-S2 dataset of
579 134,904 particles were obtained. Then multiple rounds of reference free 2D and 3D
580 classification were applied to clean up the particles for each dataset. Classes with better
581 structural features were combined and yielded ClpP-S1 dataset consisting 131,245
582 particles. After further Bayesian polishing and CTF refinement, a map at 3.3 Å
583 resolution was obtained. Two classes with better structural features were combined and
584 yielded ClpP-S2 dataset consisting 49,759 particles. After Bayesian polishing and CTF
585 refinement, a ClpP-S2 map at 3.6 Å resolution was obtained. To improve the local
586 resolution of the Cpn cap in the ClpP2-S2 map, the particles were subtracted by a soft
587 mask focusing on the cap region and re-centered. We then applied 3D classification and
588 obtained a cleaned-up dataset of 13,040 particles with better structural features
589 especially more complete density of the Cpn mobile loops, which was further refined
590 to 4.8 Å resolution of the Cpn11/20/23 map. ClpP-S2 map was then combined with the
591 Cpn11/20/23 map together by using the *vopmaximum* function in Chimera ⁵⁴,
592 generating the composite ClpP-S2 map (termed ClpP-S2c). The overall resolution was
593 determined based on the gold-standard criterion using an FSC of 0.143. The local
594 resolution estimation was determined by Local resolution function in RELION 3.1.

595 **Model building of Clp subunits with the co-chaperonin**

596 The structures of conserved regions of Clp subunits were predicted using the Rosetta
597 server ^{55,56}. The predicted structures of conserved domains were docked rigidly into the
598 density map in UCSF Chimera. The coordinates were further refined by the Real Space
599 Refine module of the Phenix suite ⁵⁷. On this basis of refinement, the Rosetta
600 enumerative sampling method was applied to build the remaining residues of each Clp
601 subunit *de novo* ^{56,58,59}. The resulting model was adjusted manually in Coot ⁶⁰. About
602 the co-chaperonin Cpn11/20/23 model, each Cpn11, Cpn20, Cpn23 homology model
603 was built with the tFold server (Tencent AI Lab) or Rosetta server . These models were
604 docked into the cryo-EM map of ClpP-S2c using the fit in map command in UCSF
605 Chimera ⁵⁴. The resulting model was subjected to Rosetta and Phenix refinement ^{57,56}.
606 The geometries and atomic model refinement statistics were evaluated by Molprobity

607 in Phenix ⁶¹.

608 Cryo-EM data acquisition, 3D process information and model refinement statistics
609 are summarized in Supplementary Table 3. Figures were generated with either UCSF
610 Chimera and ChimeraX ^{54, 62}.

611 **Data availability**

612 Electron density maps have been deposited in the Electron Microscopy Data Bank
613 under accession codes EMD-31171 for CrClpP-S1, EMD-31175 for ClpP-S2, EMD-
614 31173 for ClpP-S2c and EMD-31174 for Cpn11/20/23. Related atom coordinates file
615 also has been submitted to the Protein Data Bank, with accession codes 7EKO for
616 CrClpP-S1, and 7EKQ for CrClpP-S2c.

617 **Acknowledgments**

618 We are grateful to the staff of the NCPSS EM facility, Mass Spectrometry facility, and
619 Database and Computing facility for instrument support and technical assistance. This
620 work was funded by the Strategic Priority Research Program of Chinese Academy of
621 Sciences (Grant No. XDA24020103-2, XDB37040103), the National Key Research
622 and Development Program of China (2016YFD0100405, 2017YFA0503503) and the
623 Ministry of Agriculture of China (2016ZX08009-003-005), the ‘Initiative d'Excellence’
624 program from the French State (Grant ‘DYNAMO’, ANR-11-LABX-0011-01), and the
625 DFG (TRR 175, project C02). We thank Prof. Jean David Rochaix and Dr. Silvia
626 Ramundo for their fruitful discussion.

627 **Author Contribution**

628 C. L. and Y.C. supervised the project. N. W. executed all biochemical experiments. Y.
629 W. and X. Z. collected the cryo-EM data. Y. W. did data processing with initial map
630 from X. Z. Y.W. and N.W. did model building and structural analysis. Q. Z. started the
631 project and optimized the protein purification. C. P. performed the MS analysis. W. Z.
632 and Y. L. helped to purify protein. O. V. and M. S. were involved in the project design,
633 data analysis and interpretation. C. L., N. W., O. V. and M. S. wrote the manuscript with
634 modification from Y. W. and Y. C.

635 **Competing Interests**

636 The authors declare that they have no competing interests.

637

638 **FIGURE LEGENDS**

639 **Figure 1. Co-chaperonins interact with CrClpP**

640 (A) Purified *Chlamydomonas* ClpP complexes were separated on 12-18% SDS-
641 polyacrylamide gels and visualized by Coomassie staining. Proteins in visible bands
642 were identified by mass spectrometry and the bands identified as Clp or co-chaperonin
643 subunits are marked. The gel is a representative of four independent experiments and
644 mass spectrometry was performed twice with similar results.

645 (B) Purified ClpP and Cpn60 complexes were analyzed on a 5% non-denaturing (ND)
646 polyacrylamide gel and visualized with Coomassie staining. The arrowhead indicates
647 the position of the oligomeric 820-kDa Cpn60 complex and the asterisk the position of
648 the 60-kDa Cpn60 monomer. The gel is a representative of three independent
649 experiments.

650 (C) Immunoblots analysis. Purified ClpP complexes were separated on a 15% SDS-
651 polyacrylamide gel, transferred to nitrocellulose and immunodecorated with anti-strep
652 and anti-Cpn20 antibodies. The Strep-antibody recognized the three gene products of
653 the plastid-encoded *clpP* gene, ClpP_{1H}, ClpP_{1C} and ClpP_{1C'}. The experiment was
654 repeated three times independently with similar results.

655 (D) Immunoprecipitation of Cpn20. Total cell lysates from the ClpP-Strep (#8) strain
656 (input) were incubated with protein A-Sepharose beads coupled to antibodies of either
657 preimmune serum (pre) or anti-Cpn20 serum. Precipitated proteins were analyzed by
658 immunoblotting using antisera against ClpP4, the Strep-tag (detecting ClpP_{1C}), ClpR6,
659 ClpT4 and Cpn60. The experiment was repeated three times independently with similar
660 results.

661 (E) Co-migration of ClpP core complexes and co-chaperonins in size exclusion
662 chromatography. 1 μ M ClpP protein complex was injected into a Superdex 200 PC
663 3.2/10 column after incubation with or without 2 μ M co-chaperonin for 30 min at 4°C.
664 Proteins were eluted with 20 mM MOPS-KOH, pH 7.5, 80 mM NaCl, 10 mM MgCl₂,
665 10 mM KCl, 1 mM DTT, 10% glycerol. Then the corresponding fractions were
666 analyzed by SDS-PAGE and immunoblotting using antisera against Cpn20 or GroES.
667 The positions of the ClpP and co-chaperonin complexes are indicated at the top of the
668 panel. The experiment was repeated three times independently with similar results and
669 a representative result was shown.

670

671 **Figure 2. Co-chaperonins inhibit ClpP proteolytic activity**

672 (A) and (C) Degradation of β -casein was monitored in reactions containing β -casein
673 (16 μ M), CrClpP (0.4 μ M), EcClpP (0.4 μ M), Cpn20 (0.4 μ M), Cpn11/20/23 (0.4 μ M)
674 and ADEP dissolved in DMSO (4 or 18 μ M) as indicated. Reactions were performed at
675 30°C, aliquots were taken at the indicated time points, and analyzed via SDS-PAGE
676 (15% gels) and Coomassie staining. Arrowheads indicate the position of β -casein, stars

677 indicate the position of the EcClpP protein and arrows indicate the position of Cpn20.
678 (B) and (D) Densitometric quantification of β -casein from the reactions shown in (A)
679 and (C), respectively. Shown are mean values from three independent replicates, error
680 bars represent SD. Quantifications were made with Photoshop CS6.

681

682 **Figure 3. Cryo-EM structure of the *Chlamydomonas* ClpP complex**

683 (A) Overview of ClpP-S1 particles. The dimensions are given next to the particles. The
684 cut-away views of the central pore in the two rings of the ClpP core complex are shown.

685 (B) Overview of ClpP-S2_C particles. The dimensions are given next to the particles.

686 (C) Superimposition of EcClpP (PDB ID:1TYF) with density maps of CrClpP-S1. The
687 subunits located in the top ring of EcClpP are shown in light or dark green and the
688 subunits of the bottom ring are shown in yellow or orange. The dotted line indicates the
689 interface between the two rings. Additional densities of CrClpP not overlapping with
690 EcClp are shown in gray. The most prominent ones are labeled A1, A2 and A3.

691

692 **Figure 4. Properties of the *Chlamydomonas* ClpP core complex**

693 (A) Cartoon presentation of the CrClp core complex. The subunits designated as D1-
694 D14 are shown in the left grey cartoon and the corresponding assigned subunits are
695 colored at the right. Subunit ClpP1_C occupies positions D8, D10, D13 positions.
696 Subunit ClpP4 occupies positions D3, D5 and D7. Subunit ClpP5 occupies positions
697 D2, D4 and D6. Subunits ClpR1, R2, R3, R4 and R6 occupy single positions at D9,
698 D12, D14, D11 and D1 positions, respectively.

699 (B) Density map of the ClpP core complex. Densities with assigned Clp subunits are
700 shown in colors, densities that could not be assigned are shown in white.

701 (C) Subunit composition and arrangement in the top and bottom rings of the ClpP core
702 complex.

703 (D) Assignment of Clp subunits to the additional density map A1. This region is
704 composed of full-length ClpT4 (cyan), the long C-terminus of ClpR3 (pink), and
705 several amino acids from the C-terminus of ClpR6 (purple). The zoom-in view shows
706 the fitting of these three subunits to the density.

707 (E) Position of the ClpR4 subunit (turquoise) in the bottom ring of the ClpP core
708 complex. The long C-terminus of ClpR4 protrudes to a region next to the top ring. The
709 fitting of ClpR4 is shown in the zoom-in view.

710 (F) Electrostatic potential of the ClpP core complex. Acidic amino acids are shown in
711 red, basic amino acids in blue.

712 (G) Hydrophobicity of amino acids at the ClpP core surface. Hydrophobic amino acids
713 are shown in red, hydrophilic amino acids are shown in light blue. The seven
714 hydrophobic binding clefts observed in the top ring of Clp-S1 are encircled by black
715 broken lines.

716

717 **Figure 5. Cryo-EM structure of the *Chlamydomonas* ClpP-Cpn11/20/23 complex**

718 (A) Sequence alignment of Cpn11, Cpn20-N, Cpn20-C, Cpn23-N, Cpn23-C and GroES.
719 N and C indicate N- and C-terminal GroES-like domains of Cpn20 and Cpn23. The
720 sequence regions of mobile loop and roof are indicated with green boxes.

721 (B) Superimposition of the Cpn11/20/23 model and the cryo-EM density map. The
722 mobile loops extending from the bottom of the co-chaperonin are shown (left). The
723 enlarged picture shows the roof regions of Cpn23 and Cpn20C.
724 (C) Interaction between ClpP and co-chaperonin. The central axes of co-chaperonin
725 and CrClpP are given as black lines. The tilted angle between the axes of co-chaperonin
726 and CrClpP complex has a value of 4°.
727 (D) Top view of the extra densities located in the hydrophobic clefts at the top ring of
728 ClpP. The densities colored in red were confidently identified as mobile loops from
729 Cpn11 and Cpn20C. The densities colored in orange might be the mobile loops from
730 Cpn20N. The regions encircled with a broken red line are the two remaining clefts in
731 which no additional densities could be identified.
732 (E) Insertion of the mobile loops of Cpn11 and Cpn20C into the hydrophobic clefts at
733 the surface of the top ring of ClpP.
734 (F) Cartoon presentation of the interaction between the top ring of ClpP and
735 Cpn11/20/23. Red arrowheads indicate the hydrophobic clefts. High-confidence mobile
736 loops inserted into the clefts are indicated by solid lines, low-confidence mobile loops
737 inserted into the clefts are indicated by broken lines.
738

739 REFERENCES

- 740 1. Cyr, D.M., Hohfeld, J. & Patterson, C. Protein quality control: U-box-containing E3
741 ubiquitin ligases join the fold. *Trends in biochemical sciences* **27**, 368-375 (2002).
- 742 2. Bukau, B., Weissman, J. & Horwich, A. Molecular chaperones and protein quality
743 control. *Cell* **125**, 443-451 (2006).
- 744 3. Janska, H., Kwasniak, M. & Szczepanowska, J. Protein quality control in organelles—
745 AAA/FtsH story. *Biochimica Et Biophysica Acta (BBA)-Molecular Cell Research* **1833**,
746 381-387 (2013).
- 747 4. Baker, B.M. & Haynes, C.M. Mitochondrial protein quality control during biogenesis and
748 aging. *Trends in biochemical sciences* **36**, 254-261 (2011).
- 749 5. Sontag, E.M., Samant, R.S. & Frydman, J. Mechanisms and Functions of Spatial
750 Protein Quality Control. *Annual review of biochemistry* **86**, 97-122 (2017).

- 751 6. Hayer-Hartl, M., Bracher, A. & Hartl, F.U. The GroEL-GroES Chaperonin Machine: A
752 Nano-Cage for Protein Folding. *Trends in biochemical sciences* **41**, 62-76 (2016).
- 753 7. Kuo, W. et al. CHAPERONIN 20 mediates iron superoxide dismutase (Fe SOD) activity
754 independent of its co-chaperonin role in Arabidopsis chloroplasts. *New Phytologist* **197**,
755 99-110 (2013).
- 756 8. Zhang, X. et al. Arabidopsis co-chaperonin CPN20 antagonizes Mg-chelatase H
757 subunit to derepress ABA-responsive WRKY40 transcription repressor. *Science China*
758 *Life Sciences* **57**, 11-21 (2014).
- 759 9. Zhang, X.-F. et al. Cochaperonin CPN20 negatively regulates abscisic acid signaling
760 in Arabidopsis. *Plant molecular biology* **83**, 205-218 (2013).
- 761 10. Rawlings, N.D. et al. The MEROPS database of proteolytic enzymes, their substrates
762 and inhibitors in 2017 and a comparison with peptidases in the PANTHER database.
763 *Nucleic Acids Research* **46**, D624-D632 (2017).
- 764 11. Nishimura, K., Kato, Y. & Sakamoto, W. Chloroplast Proteases: Updates on Proteolysis
765 within and across Suborganellar Compartments. *Plant Physiol* **171**, 2280-2293 (2016).
- 766 12. Gottesman, S. proteases and their targets in escherichia coli. *Annual Review of*
767 *Genetics* **30**, 465-506 (1996).
- 768 13. Wang, J., Hartling, J.A. & Flanagan, J.M. The structure of ClpP at 2.3 Å resolution
769 suggests a model for ATP-dependent proteolysis. *Cell* **91**, 447-456 (1997).
- 770 14. Sauer, R.T. & Baker, T.A. AAA+ proteases: ATP-fueled machines of protein destruction.
771 *Annual review of biochemistry* **80**, 587-612 (2011).
- 772 15. Ripstein, Z.A., Vahidi, S., Houry, W.A., Rubinstein, J.L. & Kay, L.E. A processive rotary

773 mechanism couples substrate unfolding and proteolysis in the ClpXP degradation
774 machinery. *Elife* **9**, e52158 (2020).

775 16. Lopez, K.E. et al. Conformational plasticity of the ClpAP AAA+ protease couples
776 protein unfolding and proteolysis. *Nature Structural & Molecular Biology*, 1-11 (2020).

777 17. Fei, X. et al. Structures of the ATP-fueled ClpXP proteolytic machine bound to protein
778 substrate. *Elife* **9** (2020).

779 18. Gatsogiannis, C., Balogh, D., Merino, F., Sieber, S.A. & Raunser, S. Cryo-EM structure
780 of the ClpXP protein degradation machinery. *Nature structural & molecular biology* **26**,
781 946-954 (2019).

782 19. de Sagarra, M.R. et al. Mitochondrial localization and oligomeric structure of HClpP,
783 the human homologue of E. coli ClpP. *Journal of molecular biology* **292**, 819-825
784 (1999).

785 20. Kim, J. et al. Structures, Functions, and Interactions of ClpT1 and ClpT2 in the Clp
786 Protease System of Arabidopsis Chloroplasts. *Plant Cell* **27**, 1477-1496 (2015).

787 21. Sjögren, L.L. & Clarke, A.K. Assembly of the chloroplast ATP-dependent Clp protease
788 in Arabidopsis is regulated by the ClpT accessory proteins. *The Plant Cell* **23**, 322-332
789 (2011).

790 22. Olinares, P.D., Kim, J., Davis, J.I. & van Wijk, K.J. Subunit stoichiometry, evolution,
791 and functional implications of an asymmetric plant plastid ClpP/R protease complex in
792 Arabidopsis. *Plant Cell* **23**, 2348-2361 (2011).

793 23. Stanne, T.M., Pojidaeva, E., Andersson, F.I. & Clarke, A.K. Distinctive types of ATP-
794 dependent Clp proteases in cyanobacteria. *The Journal of biological chemistry* **282**,

- 795 14394-14402 (2007).
- 796 24. Majeran, W., Friso, G., van Wijk, K.J. & Vallon, O. The chloroplast ClpP complex in
797 *Chlamydomonas reinhardtii* contains an unusual high molecular mass subunit with a
798 large apical domain. *The FEBS journal* **272**, 5558-5571 (2005).
- 799 25. Schroda, M. & Vallon, O. in *The Chlamydomonas Sourcebook (Second Edition)*. (eds.
800 E.H. Harris, D.B. Stern & G.B. Witman) 671-729 (Academic Press, London; 2009).
- 801 26. Derrien, B., Majeran, W., Wollman, F.-A. & Vallon, O. Multistep processing of an
802 insertion sequence in an essential subunit of the chloroplast ClpP complex. *Journal of*
803 *Biological Chemistry* **284**, 15408-15415 (2009).
- 804 27. Majeran, W., Wostrikoff, K., Wollman, F.A. & Vallon, O. Role of ClpP in the Biogenesis
805 and Degradation of RuBisCO and ATP Synthase in *Chlamydomonas reinhardtii*. *Plants*
806 *(Basel)* **8** (2019).
- 807 28. Ramundo, S. et al. Conditional Depletion of the *Chlamydomonas* Chloroplast ClpP
808 Protease Activates Nuclear Genes Involved in Autophagy and Plastid Protein Quality
809 Control. *Plant Cell* **26**, 2201-2222 (2014).
- 810 29. Sjogren, L.L., Stanne, T.M., Zheng, B., Sutinen, S. & Clarke, A.K. Structural and
811 functional insights into the chloroplast ATP-dependent Clp protease in *Arabidopsis*.
812 *Plant Cell* **18**, 2635-2649 (2006).
- 813 30. Kim, J. et al. Subunits of the plastid ClpPR protease complex have differential
814 contributions to embryogenesis, plastid biogenesis, and plant development in
815 *Arabidopsis*. *Plant Cell* **21**, 1669-1692 (2009).
- 816 31. Vahidi, S. et al. An allosteric switch regulates *Mycobacterium tuberculosis* ClpP1P2

- 817 protease function as established by cryo-EM and methyl-TROSY NMR. *Proceedings*
818 *of the National Academy of Sciences* **117**, 5895-5906 (2020).
- 819 32. Derrien, B. et al. The purification of the *Chlamydomonas reinhardtii* chloroplast ClpP
820 complex: additional subunits and structural features. *Plant molecular biology* **80**, 189-
821 202 (2012).
- 822 33. LaBreck, C.J., May, S., Viola, M.G., Conti, J. & Camberg, J.L. The Protein Chaperone
823 ClpX Targets Native and Non-native Aggregated Substrates for Remodeling,
824 Disassembly, and Degradation with ClpP. *Front Mol Biosci* **4**, 26 (2017).
- 825 34. Gersch, M. et al. AAA+ chaperones and acyldepsipeptides activate the ClpP protease
826 via conformational control. *Nat Commun* **6**, 6320 (2015).
- 827 35. Sass, P. et al. Antibiotic acyldepsipeptides activate ClpP peptidase to degrade the cell
828 division protein FtsZ. *Proceedings of the National Academy of Sciences* **108**, 17474-
829 17479 (2011).
- 830 36. Kirstein, J. et al. The antibiotic ADEP reprogrammes ClpP, switching it from a regulated
831 to an uncontrolled protease. *EMBO molecular medicine* **1**, 37-49 (2009).
- 832 37. Zhao, Q. et al. Hetero-oligomeric CPN60 resembles highly symmetric group-I
833 chaperonin structure revealed by Cryo-EM. *The Plant journal : for cell and molecular*
834 *biology* **98**, 798-812 (2019).
- 835 38. Tsai, Y.-C.C., Mueller-Cajar, O., Saschenbrecker, S., Hartl, F.U. & Hayer-Hartl, M.
836 Chaperonin cofactors, Cpn10 and Cpn20, of green algae and plants function as hetero-
837 oligomeric ring complexes. *Journal of Biological Chemistry* **287**, 20471-20481 (2012).
- 838 39. Schroda, M. The *Chlamydomonas* genome reveals its secrets: chaperone genes and

- 839 the potential roles of their gene products in the chloroplast. *Photosynth Res* **82**, 221-
840 240 (2004).
- 841 40. Bracher, A., Whitney, S.M., Hartl, F.U. & Hayer-Hartl, M. Biogenesis and Metabolic
842 Maintenance of Rubisco. *Annu Rev Plant Biol* **68**, 29-60 (2017).
- 843 41. Weiss, C., Bonshtien, A., Farchi-Pisanty, O., Vitlin, A. & Azem, A. Cpn20: siamese
844 twins of the chaperonin world. *Plant molecular biology* **69**, 227 (2009).
- 845 42. Koumoto, Y., Shimada, T., Kondo, M., Hara-Nishimura, I. & Nishimura, M. Chloroplasts
846 have a novel Cpn10 in addition to Cpn20 as co-chaperonins in *Arabidopsis thaliana*.
847 *Journal of Biological Chemistry* **276**, 29688-29694 (2001).
- 848 43. Rizzolo, K. et al. Functional cooperativity between the trigger factor chaperone and the
849 ClpXP proteolytic complex. *Nature communications* **12**, 1-18 (2021).
- 850 44. Kim, J. et al. Modified Clp protease complex in the ClpP3 null mutant and
851 consequences for chloroplast development and function in *Arabidopsis*. *Plant Physiol*
852 **162**, 157-179 (2013).
- 853 45. Olinares, P.D., Kim, J. & van Wijk, K.J. The Clp protease system; a central component
854 of the chloroplast protease network. *Biochim Biophys Acta* **1807**, 999-1011 (2011).
- 855 46. Gallastegui, N. & Groll, M. The 26S proteasome: assembly and function of a destructive
856 machine. *Trends in biochemical sciences* **35**, 634-642 (2010).
- 857 47. Maurizi, M.R., Thompson, M.W., Singh, S.K. & Kim, S.-H. in *Methods in enzymology*,
858 Vol. 244 314-331 (Elsevier, 1994).
- 859 48. Derrien, B. & Vallon, O. One-step Affinity Purification of the Chloroplast ClpP Complex
860 from the Green Alga *Chlamydomonas reinhardtii* Using the Strep-tagII Epitope Tag.

- 861 *Bio-protocol* **3**, e315 (2013).
- 862 49. Bai, C. et al. Protomer Roles in Chloroplast Chaperonin Assembly and Function.
863 *Molecular plant* **8**, 1478-1492 (2015).
- 864 50. Mastronarde, D.N. Automated electron microscope tomography using robust prediction
865 of specimen movements. *Journal of Structural Biology* **152**, 36-51 (2005).
- 866 51. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from
867 electron micrographs. *Journal of structural biology* **192**, 216-221 (2015).
- 868 52. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure
869 determination. *Journal of structural biology* **180**, 519-530 (2012).
- 870 53. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for
871 improved cryo-electron microscopy. *Nature methods* **14**, 331-332 (2017).
- 872 54. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research
873 and analysis. *Journal of computational chemistry* **25**, 1605-1612 (2004).
- 874 55. Raman, S. et al. Structure prediction for CASP8 with all-atom refinement using Rosetta.
875 *Proteins* **77 Suppl 9**, 89-99 (2009).
- 876 56. Song, Y. et al. High-resolution comparative modeling with RosettaCM. *Structure* **21**,
877 1735-1742 (2013).
- 878 57. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for
879 macromolecular structure solution. *Acta crystallographica. Section D, Biological*
880 *crystallography* **66**, 213-221 (2010).
- 881 58. Frenz, B., Walls, A.C., Egelman, E.H., Veisler, D. & DiMaio, F. RosettaES: a sampling
882 strategy enabling automated interpretation of difficult cryo-EM maps. *Nature methods*

883 **14**, 797-800 (2017).

884 59. Wang, R.Y.-R. et al. De novo protein structure determination from near-atomic-
885 resolution cryo-EM maps. *Nature methods* **12**, 335-338 (2015).

886 60. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot.
887 *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).

888 61. Williams, C.J. et al. MolProbity: More and better reference data for improved all-atom
889 structure validation. *Protein science : a publication of the Protein Society* **27**, 293-315
890 (2018).

891 62. DeLano, W.L. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On*
892 *Protein Crystallography* **40**, 82-92 (2002).

893

894