



Research article

## Computational analysis of homo-dimerization of chloroplast-localized chaperonin from the alga *Chlamydomonas reinhardtii*

Rungdawan Wongsamart<sup>a</sup>, Duangnapa Kiriwan<sup>b</sup>, Chonticha Suwattanasophon<sup>c</sup>, Kittisak Yokthongwattana<sup>a,d</sup>, Kiattawee Choowongkamon<sup>b,e,\*</sup>

<sup>a</sup> Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

<sup>b</sup> Department of Biochemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

<sup>c</sup> Department of Pharmaceutical Technology and Biopharmaceutics, University of Vienna, Vienna 1010 Austria

<sup>d</sup> Center for Excellence in Protein and Enzyme technology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

<sup>e</sup> Center for Advanced Studies in Nanotechnology for Chemical, Food and Agricultural Industries, KU Institute for Advanced Studies, Kasetsart University, Bangkok 10900, Thailand

### Article Info

#### Article history:

Received 26 April 2021

Revised 8 September 2021

Accepted 11 September 2021

Available online 31 October 2021

#### Keywords:

Binding free energy,

*Chlamydomonas reinhardtii*,

Cpn60,

Homodimer,

MD simulation

### Abstract

Molecular chaperones are major groups of proteins responsible for proteostasis regulation. GroEL is a protein belonging to the chaperonin family of molecular chaperones. Unlike GroEL in prokaryotes, chloroplast chaperonin 60 (Cpn60) has distinct subunits. The chaperonin for the alga *Chlamydomonas reinhardtii* has one  $\alpha$  and two  $\beta$  subunits ( $\beta 1$  and  $\beta 2$ ). Even though a crystal structure of homo-oligomer (Cpn60 $\beta 1$ ) has been reported, it is still unclear how these subunits assemble. This study modeled homo-subunit dimers of *C. reinhardtii*. Dimers  $\alpha\alpha$ ,  $\beta 1\beta 1$  and  $\beta 2\beta 2$  were analyzed based on molecular dynamics simulation. Many hydrogen bonds presented in the same position among the three dimers, even though the amino acids were different. Charged amino acids were an important factor influencing the binding interface orientation. The binding free energy of the  $\beta 1\beta 1$  dimer was the lowest, followed by  $\beta 2\beta 2$  and  $\alpha\alpha$ , respectively. The results suggested that the  $\beta 1\beta 1$  assembly was more favorable than  $\beta 2\beta 2$  and  $\alpha\alpha$ .

### Introduction

Maintenance of proteostasis is crucial for cellular health in all organisms; molecular chaperones are major groups of

proteins responsible for protein stasis regulation (Gething and Sambrook (1992). Chaperonins, which belong to a subfamily of molecular chaperones, have a primary role in assisting newly translated, imported or denatured proteins to correctly fold/refold in an ATP-dependent manner (Gething and Sambrook, 1992; Horwich et al., 2007). Group I chaperonins are found

\* Corresponding author.

E-mail address: [kiattawee.c@ku.th](mailto:kiattawee.c@ku.th)

in prokaryotes, mitochondria and plastids, while group II chaperonins can be found in archaea and eukaryotic cytosol (Hayer-Hartl et al., 1995; Boston et al., 1996; Dickson et al., 2000). GroEL is probably the best characterized chaperonin protein. Braig et al. (1994) reviewed the structure of GroEL that its functional complex consists of two rings stacked back to back into a double barrel structure, each of which has seven identical subunits (~57 kDa each). Structurally, individual subunits contain three domains: equatorial, intermediate and apical. The equatorial domain is the interaction zone between the subunits within or between the rings as well as providing a binding site for ATP. The apical domain contains binding sites for the protein substrate and the co-chaperonin (GroES, also known as a lid protein). The intermediate domain is a linker between the apical and equatorial domains. Functional activity of GroEL is achieved by cooperation between the core tetradecameric structures with detachable lids (Saibil et al., 2013). The overall double-barrel structure binds to the target protein. Over time, one of the two barrels is in a conformation so that the inner surface becomes hydrophobic to facilitate the substrate binding. Concomitantly, the co-chaperone GroES binds and closes the barrel. Subsequent ATP hydrolysis promotes a conformational change leading to expansion of the volume inside the barrel as well as a transition of the inner surface from hydrophobic to hydrophilic, allowing the denatured substrate to refold (Bukau and Horwich, 1998). Another round of ATP binding at the other barrel releases the now-correctly-folded substrate as well as the GroES and the cycle repeats itself (Bukau and Horwich, 1998).

From bacteria to eukaryotic mitochondria, there exists only one gene copy for GroEL or chaperonin 60 and one copy for GroES or chaperonin 10 (Schroda, 2004a). However, notably in plant chloroplasts, the GroEL homologs exist in more than one form as A and B and sometimes referred to as  $\alpha$  and  $\beta$  (Schroda, 2004a). Each form can also exist in more than one copy. The *Arabidopsis thaliana* Cpn60 has two  $\alpha$  and four  $\beta$  subunits, *Pisum sativum* has one  $\alpha$  and one  $\beta$  subunit and *Chlamydomonas reinhardtii* has one  $\alpha$  and two  $\beta$  (Hemmingsen and Ellis, 1986; Schroda, 2004b). At the protein level, *Chlamydomonas*  $\alpha$  and  $\beta$  subunits share 50% identity while the two  $\beta$  isoforms are 80% identical to each other (Schroda, 2004a). At present, it is still not completely clear how these subunits are arranged in the active tetradecameric structure. The *A. thaliana* Cpn60 was thought to contain equal amounts of  $\alpha$  and  $\beta$  subunits (Martel et al., 1990; Nishio et al., 1999), and in plant chloroplasts, there exist two types of oligomers: homo and hetero (Vitlin et al., 2011). Studies of

*C. reinhardtii* revealed that the  $\beta$ 1 and  $\beta$ 2 subunits could form homo-tetradecameric complexes *in vitro* but the population of the homocomplex was readily reduced when the  $\alpha$  subunit was present (Bai et al., 2015). On the other hand, the  $\alpha$  subunit alone could not form the oligomer unless there was cross-assembling with the  $\beta$ 1 and  $\beta$ 2 subunits. It was thought that the equatorial domain of  $\alpha$  subunit could be the cause of  $\alpha$  homo-oligomeric disassembly (Zhang et al., 2016). In nature, the ratio of  $\alpha$ : $\beta$ 1: $\beta$ 2 subunits was suggested to be 5:6:3, respectively (Bai et al., 2015).

Aside from the aforementioned reports, no other studies support or disprove the existing hypothesis. There are many possibilities of how  $\alpha$ ,  $\beta$ 1 and  $\beta$ 2 subunits combine in nature. The current focused on why  $\beta$ 1 subunits were able to form a homo-tetradecameric structure *in vitro* and even have a crystal structure (Zhang et al., 2016), while there was no evidence of a homo-oligomer of the  $\alpha$  subunit. However, as running all the combination would take years, the current study looked at how the same type of two subunits form together and molecular dynamics simulation was used to elucidate the dynamic aspects of Cpn60 subunit homo-dimerization including the  $\alpha\alpha$ ,  $\beta$ 1 $\beta$ 1,  $\beta$ 2 $\beta$ 2 dimers. Study of hetero-dimers which requires six systems ( $\alpha\beta$ 1,  $\beta$ 1 $\alpha$ ,  $\alpha\beta$ 2,  $\beta$ 2 $\alpha$ ,  $\beta$ 1 $\beta$ 2 and  $\beta$ 2 $\beta$ 1) because of the difference in left-right interfaces will be subsequently investigated. Interactions at the interfaces of the dimers were analyzed including hydrogen bonding, electrostatic attraction. Furthermore, binding free energy was calculated using the MM-PBSA method. The results should provide more information on how the subunits interact and assemble.

## Materials and Methods

### Model setup

The amino acid sequence and 3D structure of GroEL of *Escherichia coli* strain K-12 was obtained from the Protein Data Bank (structure code 4WSC), 548 amino acids long. The *Chlamydomonas reinhardtii* chaperonin 60 $\alpha$  amino acid sequence was retrieved from the NCBI database (www.ncbi.nlm.nih.gov), accession number XP\_001703692.1, 580 amino acids long. For chaperonin 60 $\beta$ 1, the amino acid sequence was retrieved from Protein Data Bank (structure code 5CDI), 581 amino acids long. For chaperonin 60 $\beta$ 2, the amino acid sequence was retrieved from the NCBI database, accession number XP\_001692504.1, 577 amino acids long. The signal peptides were predicted and removed from amino acid sequences of Cpn60 $\alpha$  (1–38), Cpn60 $\beta$ 1(1–32) and Cpn60 $\beta$ 2

(1–33) using the signalP and ChloroP programs. (Emanuelsson et al., 1999)

The 3D models of the Cpn60 $\alpha$ , Cpn60 $\beta$ 1 and Cpn60 $\beta$ 2 dimers were first constructed in the SWISS-PDB program (Guex and Peitsch, 1997) using GroEL 4WSC as a template.

There is no information on the C-terminal structure of GroEL 4WSC from amino acid 526 to 548 because it is the flexible tail of the complex; therefore, the end parts of Cpn60 $\alpha$ , Cpn60 $\beta$ 1 and Cpn60 $\beta$ 2 were removed according to their template. The removed C-terminal regions of Cpn60 $\alpha$ , Cpn60 $\beta$ 1 and Cpn60 $\beta$ 2 were the last 22, 21 and 18 amino acids, respectively. Only some of the C-terminal regions were cut; most of residues important for the equatorial binding were still there. The rest of the sequences were newly aligned and assigned numbers as shown in [Supplementary Fig. S1](#)

#### *Molecular dynamics simulation*

The GROMACS 5.1 software (Abraham et al., 2015) with a 54a7 force field (Huang et al., 2011) was used to run the molecular dynamics (MD) simulation of each dimer complex. Each complex was solvated in a cubic box of extended simple point charge water molecules (Bonvin et al., 2000). These systems were neutralized by adding 34 Na<sup>+</sup> counter ions. All systems were subjected to 50,000 steps of energy minimized with the steepest descent algorithm to release conflicting contacts; the equilibration phase (at constant number of particles, constant volume, and constant temperature or NVT) was done at 300 K for 200 ps followed by MD production runs for 5 ns. For the purpose of analysis, MD was used to perform for 100 ns.

#### *Binding free energy calculations*

The free binding energies of  $\alpha\alpha$ ,  $\beta$ 1 $\beta$ 1 and  $\beta$ 2 $\beta$ 2 complexes were calculated using the MM/PBSA method after MD stabilization.

The binding free energies of the complexes between dimers were analyzed during the equilibrium phase based on a snapshot generated at an interval of 1.5 ps in the 80–100 ns range of MD trajectories using the g\_mmpbsa tool with default settings (Kumari et al., 2014).

#### *Hydrogen bond interactions*

Hydrogen bond interactions and %occupancy were observed at the last 20 ns of molecular dynamics simulation.

#### *Ionic bond interactions*

Ionic bonds were observed at 100 ns MD structures. The protein interactions calculator (PIC) from the Molecular Biophysics Unit, Indian Institute of Science, Bangalore (Tina et al., 2007) was used to find the interaction. The distance cut off was 500 pm.

#### *Calculation of binding free energy using molecular mechanics combined with Poisson-Boltzmann and surface area calculations approach*

Binding free energies for each dimer were calculated using the g\_mmpbsa tool in GROMACS. The molecular mechanics combined with Poisson-Boltzmann and surface area calculations (MM-PBSA) method was used as follows:

$$\begin{aligned} G_{\text{binding}} &= E_{\text{gas}} + G_{\text{sol}} - T\Delta S \\ E_{\text{gas}} &= E_{\text{int}} + E_{\text{vdw}} + E_{\text{coulomb}} \\ G_{\text{sol}} &= G_{\text{pb}} + G_{\text{np}} \\ &= G_{\text{pb}} + \text{SASA} \end{aligned}$$

Therefore;  $G_{\text{binding}} = (E_{\text{int}} + E_{\text{vdw}} + E_{\text{coulomb}}) + (G_{\text{pb}} + \text{SASA}) - T\Delta S$

where  $E_{\text{gas}}$  is the average molecular mechanics potential energy in a vacuum or gas phase energy including internal energy ( $E_{\text{int}}$ ), van der Waals forces ( $E_{\text{vdw}}$ ) and Coulomb or electrostatic force ( $E_{\text{coulomb}}$ ).  $G_{\text{sol}}$  denotes the contribution to the solvation free energy consisting of polar solvation ( $G_{\text{pb}}$ ) and nonpolar solvation ( $G_{\text{np}}$ ) which was calculated from the solvent accessible area (SASA). In this study, the  $-T\Delta S$  term was excluded since the system was calculated at constant temperature.

All the binding free energies of dimers were calculated from 20 snapshots extracted every 1,000 ps from the last 20 ns of molecular dynamics trajectories.

## **Results and Discussion**

#### *Structural stabilities and flexibilities of systems*

There are many possibilities for how  $\alpha$ ,  $\beta$ 1 and  $\beta$ 2 subunits combine in nature. The sequence identity and similarity between individual subunits and GroEL (4WSC) are shown in Table 1.

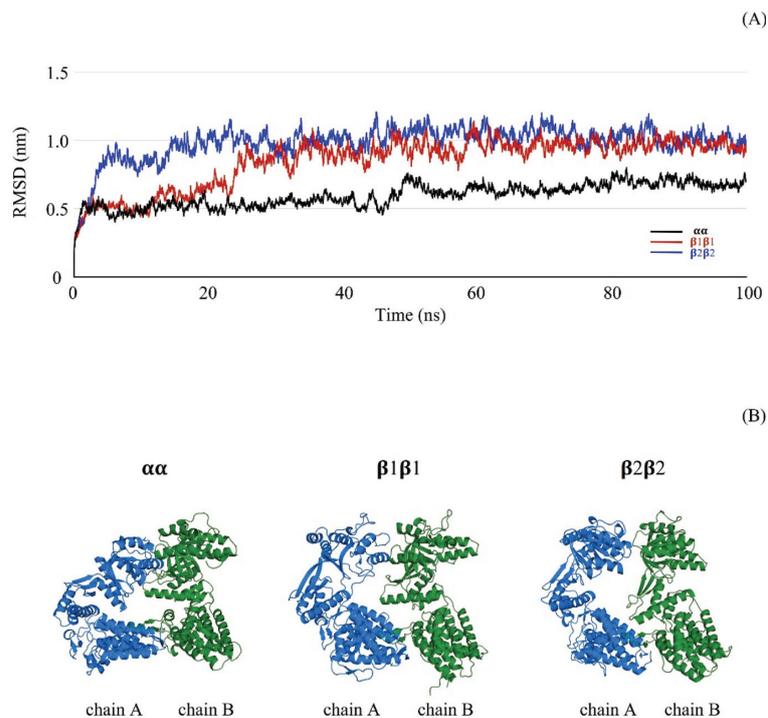
**Table 1** Amino acid sequence identity and similarity of individual Cpn60 subunits and GroEL

	%Identity (%Similarity)			
	GroEL	$\alpha$	$\beta 1$	$\beta 2$
GroEL		46.7% (69.0%)	49.4% (69.2%)	45.7% (69.7%)
$\alpha$	46.7% (69.0%)		50.8% (83.5%)	52.7% (84.2%)
$\beta 1$	49.4% (69.2%)	50.8% (83.5%)		82.8% (95.5%)
$\beta 2$	45.7% (69.7%)	52.7% (84.2%)	82.8% (95.5%)	

The current study investigated the simple possibilities of binding by considering how the same type of subunits formed together. So, three systems were set in this simulation: first, the homodimer of Cpn60 $\alpha$ , called  $\alpha\alpha$ ; second, the homodimer of Cpn60 $\beta 1$ , called  $\beta 1\beta 1$ ; and third, the homodimer of Cpn60 $\beta 2$ , called  $\beta 2\beta 2$ . The trajectories of molecular dynamic simulation were analyzed to assess the structural stabilities of the systems. The structural stabilities of all systems were informed by constant in root mean square deviations (RMSD). All three systems were stable at 80–100 ns (Fig. 1A).  $\beta 1\beta 1$  and  $\beta 2\beta 2$  had similar RMSD values which were higher than for  $\alpha\alpha$ . The higher RMSD only indicated a higher deviation of heavy atoms in the MD structure compared with heavy atoms in the original structure before energy minimization. However, the higher RMSD does not indicate a higher stability;

instead, the stability of the system can be represented as a form of constant RMSD. Therefore, at the time 80–100 ns, all three systems were stable and were in native state as confirmed by a Ramachandran plot in PROCHECK (Laskowski et al., 1993). The structures in this time period were used to analyze root mean square fluctuations (RMSF), hydrogen bond and ionic bond interactions, energy per residue and binding free energy.

The structures of the  $\alpha\alpha$ ,  $\beta 1\beta 1$  and  $\beta 2\beta 2$  dimers at 100 ns of simulation are shown in Fig. 1B. The radius of gyration (RG) which accounts for how much the MD is intact was calculated and averaged from three dimensions. The RG values of  $\alpha\alpha$ ,  $\beta 1\beta 1$  and  $\beta 2\beta 2$  were 3.37, 3.59, and 3.55 nm, respectively, indicating that the  $\beta 1\beta 1$  and  $\beta 2\beta 2$  dimers were extended more than the  $\alpha\alpha$  dimer.



**Fig. 1** Structural stability and flexibility of simulated systems: (A) root-mean-square deviation (RMSD) plots of  $\alpha\alpha$ ,  $\beta 1\beta 1$ , and  $\beta 2\beta 2$  dimers of 100 ns simulation; (B) simulated structures of  $\alpha\alpha$ ,  $\beta 1\beta 1$  and  $\beta 2\beta 2$  dimers at time 100 ns

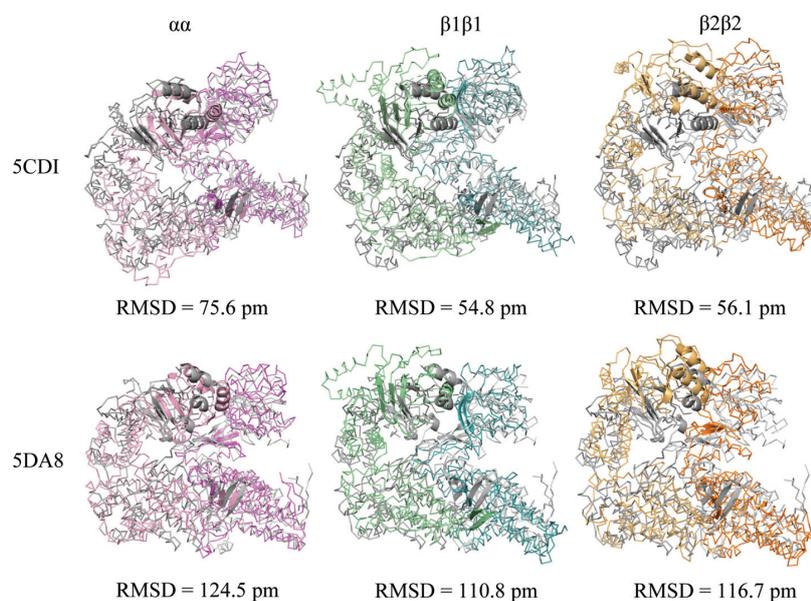
The overall structures of the simulated dimers were a good fit with both crystal structures. 5CDI is a crystal structure of the  $\beta 1$  homo-tetradecamer of *Chlamydomonas reinhardtii* (Fig. 2). Therefore, the crystal was used as the positive control. Overall the RMSD between crystal structures and all homology models was less than 200 pm indicating good quality models. However, some parts of the simulated dimers were different from the 5CDI crystal structure due to greater flexibility of the molecular dynamics, resulting in greater interaction at the interface of the simulated  $\beta 1\beta 1$  (17 H-bond, 11 ionic bond) compared to 5DCI (13 H-bond, 8 ionic bonds). 5DA8 is a crystal structure of chaperonin GroEL, consisting of 14 identical subunits arranging in two rings, of *Chlorobaculum tepidum* TLS (Chang, 2015). It is the closest structure available for chaperonin 60 with amino acid similarity of 51%. The alignment showed that the simulated dimers were a good fit with the dimer of 5DA8, ensuring that the simulated dimers were suitable for use in further analysis and to the study nature of Cpn60 dimerization.

After stable simulation systems of all three dimers had been achieved, the flexibility of their interfaces was examined. Some studies have suggested that  $\beta 1$  subunits could form homooligomers, while  $\alpha$  subunits cause oligomer disassembly (Bai et al., 2015; Zhang et al., 2016). Investigating how residues behave at the interface would explain the phenomenon. To see the flexibility of the dimer interfaces, the RMSF of

alpha carbon atoms of MD-simulated dimers were compared with the starting structures. The average RMSF of residues at the dimer interfaces at 80–100 ns of the simulation were considered. To have comparable information of each residue at the interface, the residue numbers were newly defined according to the sequence alignment of GroEL, Cpn60 $\alpha$ , Cpn60 $\beta 1$  and Cpn60 $\beta 2$  as shown in the supplementary S1. The residues were selected from the interfaces of dimers or because they had the same alignment position with interface residues, as shown in Fig. 3A. Fig. 3B indicates that in A chains, every monomer had a flexible region at residues 199–312. Subunit  $\alpha$  seemed to have a lower RMSF than  $\beta 2$  and  $\beta 1$ , indicating a more constrained dynamic structure. B chains all had similar fluctuations; however, subunit  $\alpha$  was less flexible than the other two subunits.

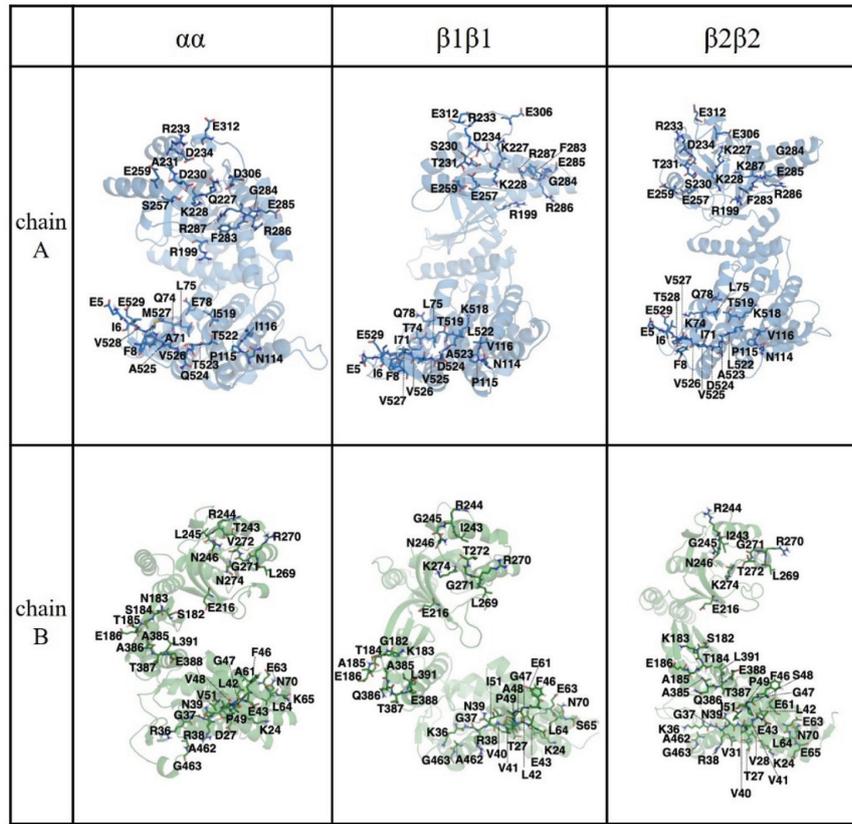
#### Interactions at interface of each dimer

To confirm the dimer strength, hydrogen bonding was analyzed between residues in each chain. The hydrogen bond information between the  $\alpha\alpha$ ,  $\beta 1\beta 1$  and  $\beta 2\beta 2$  dimers was calculated from the last 3 ns of the simulation. Fig. 4A and [Supplementary Fig. S2](#) the H-bond formation at 100 ns. Dimer  $\alpha\alpha$  had a higher number of H-bonds than  $\beta 1\beta 1$  and  $\beta 2\beta 2$  that both appeared to have similar numbers of H-bonds. The %occupancy showed bond stability during this time period, with most H-bonds having more than 50% occupancy.



**Fig. 2** Structural alignment, where dimer part of  $\beta 1$  homo-oligomer crystal structure (5CDI) in gray and dimer part of chaperonin GroEL crystal structure (5DA8) in gray aligns with  $\alpha\alpha$  dimer (pink),  $\beta 1\beta 1$  dimer (green) and  $\beta 2\beta 2$  dimer (orange) and some interaction parts of interfaces are shown in cartoon

(A)



(B)

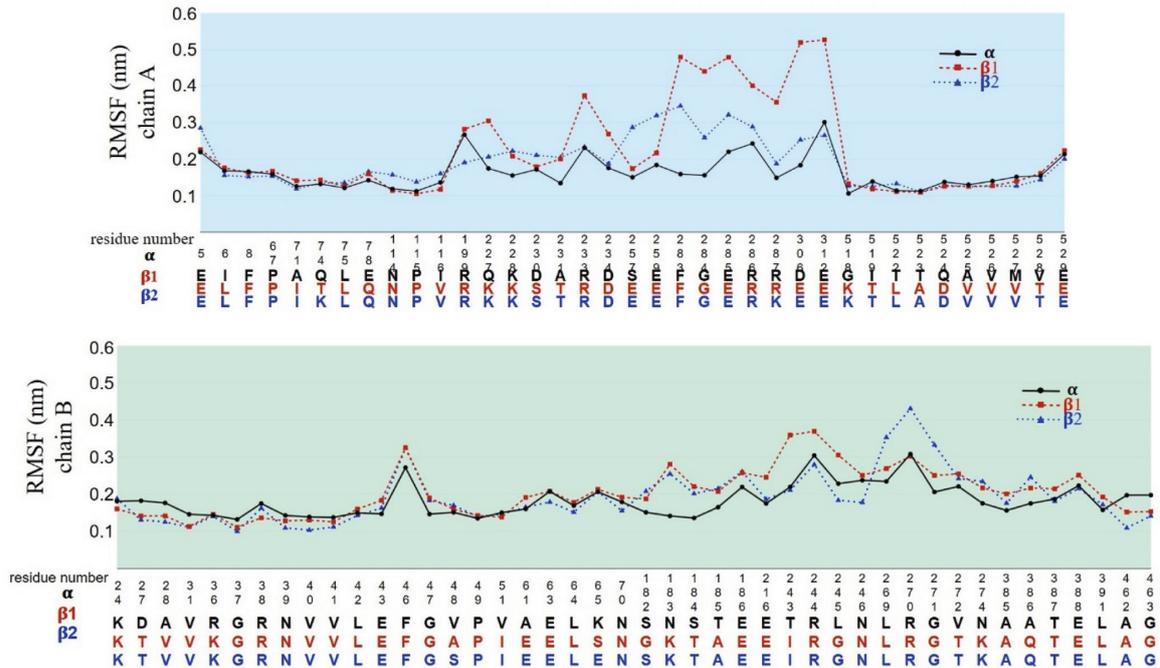
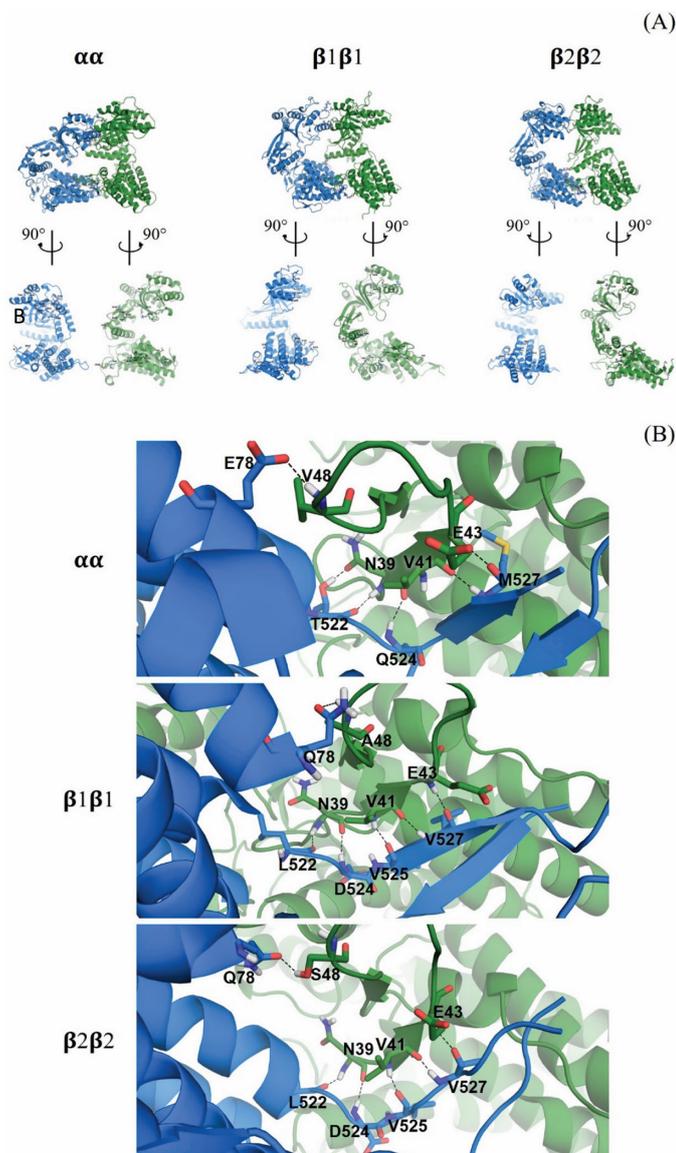


Fig. 3 Structural flexibility of simulated systems: (A):  $\alpha\alpha$ ,  $\beta1\beta1$ , and  $\beta2\beta2$  dimers opened to show residues lying at interface and their root mean square fluctuation (RMSF); (B) RMSF at interfaces of  $\alpha\alpha$ ,  $\beta1\beta1$  and  $\beta2\beta2$  dimers, where black, red and blue lines represent  $\alpha$ ,  $\beta1$  and  $\beta2$  subunits, respectively

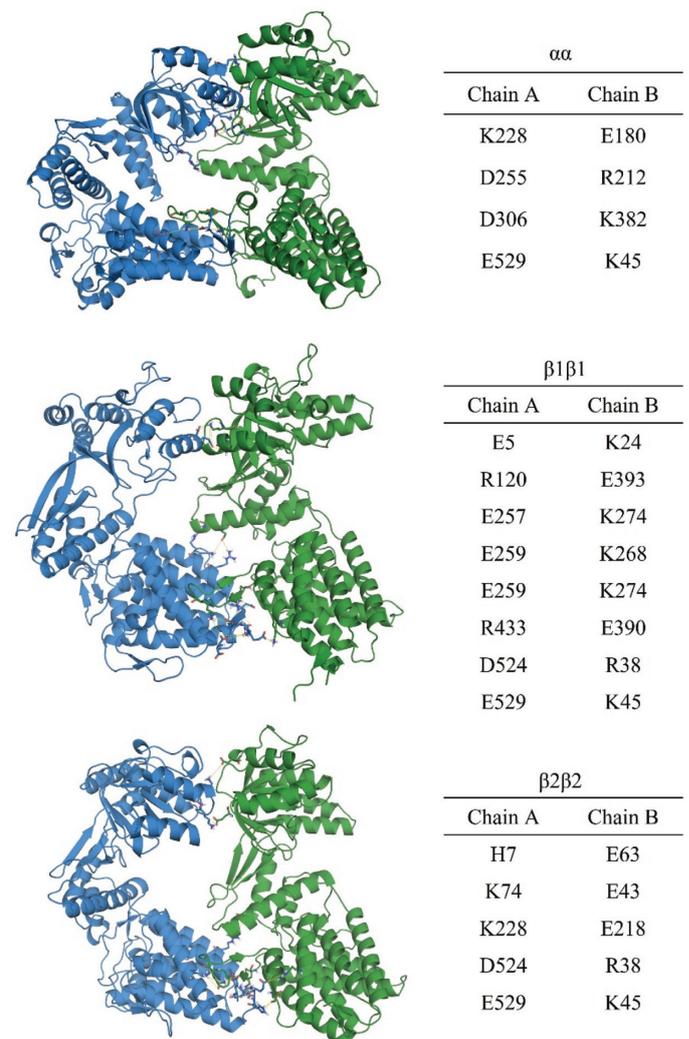


**Fig. 4** Hydrogen bonds at interfaces of each dimer at time 100 ns, for chain A in blue and chain B in green: (A) dimer interfaces opened to show residue forming an H bond; (B) some equatorial residues forming H-bonds at same alignment positions among three dimers

Interestingly, many H-bonds formed at the same positions among dimers. For example in the equatorial domain, the two monomers tried to form a  $\beta$ -sheet by making H-bonds between their backbones, as shown in Fig. 4B. Interfaces of the  $\beta 1\beta 1$  and  $\beta 2\beta 2$  dimers had the same five pairs of H-bonds, namely 1) L522 and N39; 2) D524 and N39; 3) V525 and V41; 4) V527 and V41; and 5) V527 and E43, with one similar pair being Q87 with A48 in  $\beta 1\beta 1$  and with S48 in  $\beta 2\beta 2$ . Despite the similarity between  $\beta 1$  and  $\beta 2$ ,  $\beta 2\beta 2$  could not form a  $\beta$ -sheet because of the distance and angle of the overall interface structure.

The residue pairs from the  $\alpha\alpha$  equatorial domain were quite different. The same N39, V41 and E43 were used to make H-bonds, but instead they bonded with T522, Q524 and M527, respectively. The simulation structures along with the %occupancy showed that these positions were crucial for subunit formation.

The electrostatic interaction was also important. Ionic bonds may stabilize or break various biological complexes. The ionic bonds between each dimer at 100 ns are shown in Fig. 5. Dimer  $\beta 1\beta 1$  had more ionic interactions compared to  $\beta 2\beta 2$  and  $\alpha\alpha$ . Interestingly, some ionic interaction pairs were the same among dimers. E529 and K45 were conserved residues in every type of subunit and they contributed to the stability of the dimers through ionic interaction. D524 and R38 made a pair commonly formed in  $\beta 1\beta 1$  and  $\beta 2\beta 2$ ; however, in  $\alpha\alpha$ , the position 524 was glutamine which repelled R38.

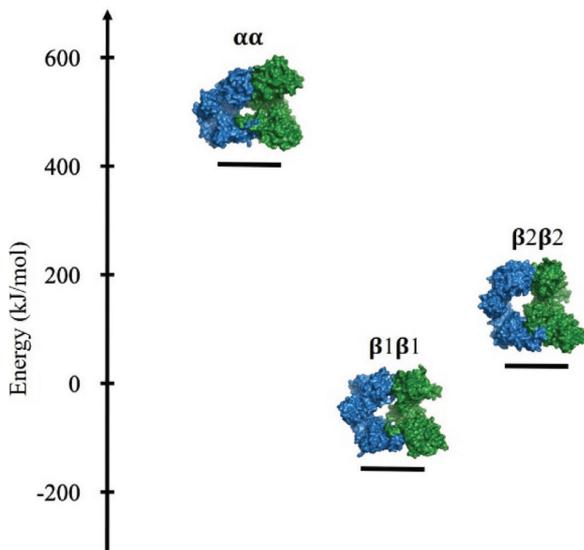


**Fig. 5** Residues forming ionic bonds at simulation time of 100 ns

*Binding energy of dimers calculated using molecular mechanics combined with Poisson-Boltzmann and surface area calculations method*

The binding energies of each complex were calculated to identify the stable structures; the lower the energy, the more stable the complex. Therefore, the energy was calculated of each dimer to see which homodimer was more likely to form in the cells. The binding energy of each dimer was calculated using the MM-PBSA method. The total binding energy is shown in Table 2 and Fig. 6.

Dimer  $\beta 1\beta 1$  had the lowest binding energy (mean  $\pm$  SD,  $-157.51 \pm 2.29$  kJ/mol) showing the ability to form in biological environments. The electrostatic energy between the  $\beta 1$  monomer interface ( $E_{lec}$ ) mainly contributed to this stability of the dimer. Dimers  $\alpha\alpha$  and  $\beta 2\beta 2$  had positive value binding energies of  $403.84 \pm 2.82$  kJ/mol and  $33.11 \pm 1.95$  kJ/mol, respectively.  $\alpha\alpha$  had a high binding energy due to its high polar electrostatic solvation energy, indicating that it was unlikely to form in a water environment.

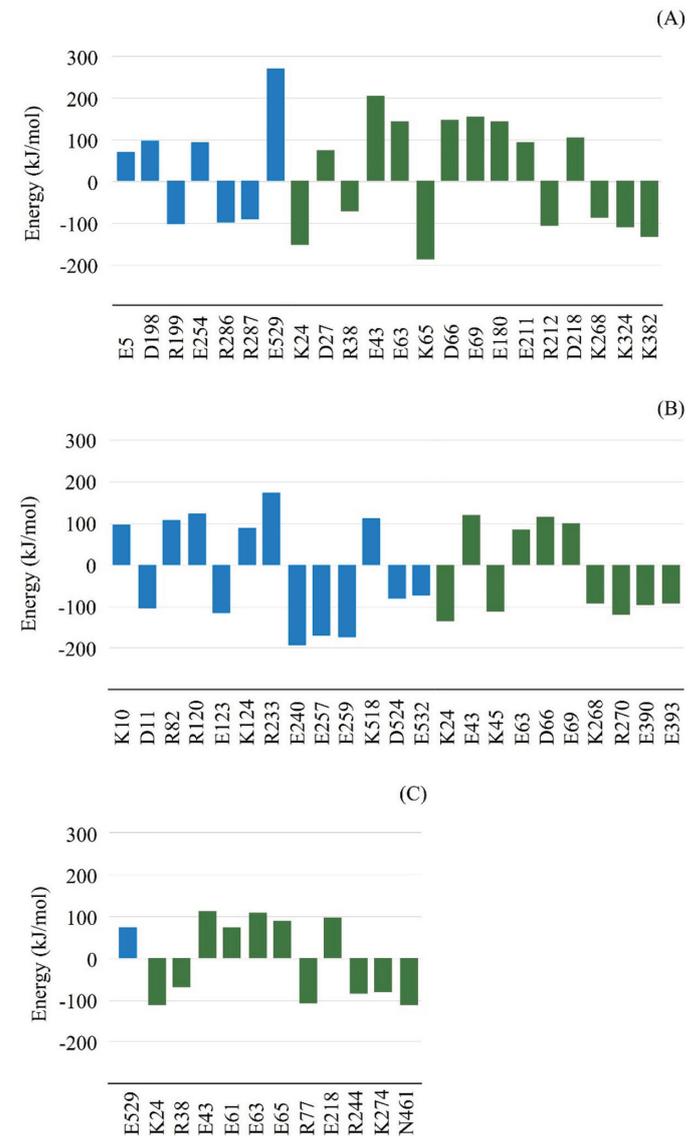


**Fig. 6** Graph showing total binding energy EMM-PBSA of  $\alpha\alpha$ ,  $\beta 1\beta 1$  and  $\beta 2\beta 2$  dimers

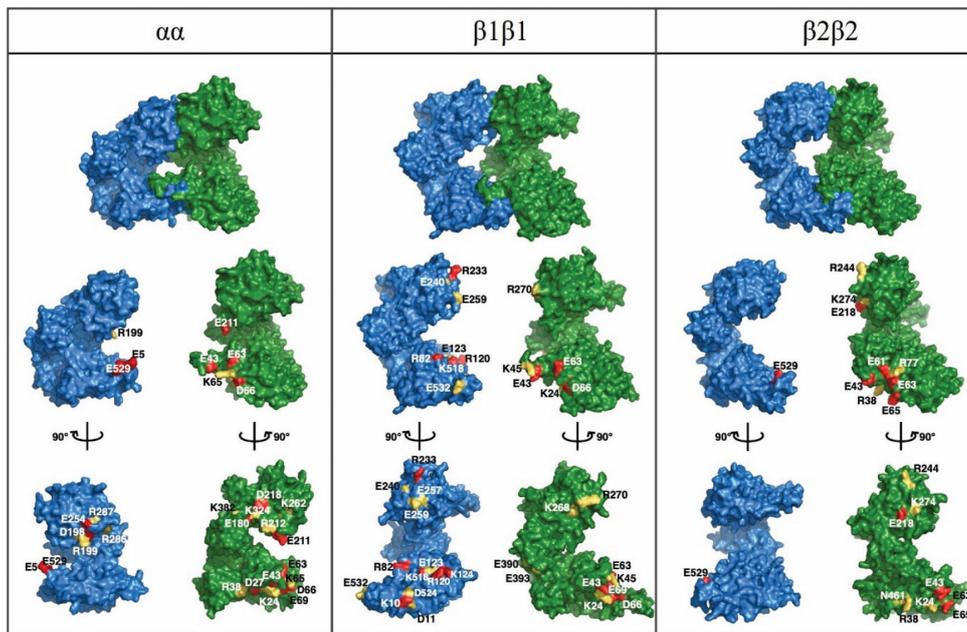
**Table 2** Mean ( $\pm$  SD) total binding energy EMM-PBSA of  $\alpha\alpha$ ,  $\beta 1\beta 1$  and  $\beta 2\beta 2$  dimers using structures at times 80–100 ns of molecular dynamics simulation every 500 ps

	$\alpha\alpha$ (kJ/mol)	$\beta 1\beta 1$ (kJ/mol)	$\beta 2\beta 2$ (kJ/mol)
$E_{vdw}$	$-821.89 \pm 0.90$	$-390.03 \pm 0.62$	$-428.31 \pm 0.68$
$E_{elec}$	$-642.39 \pm 3.18$	$-1037.72 \pm 3.36$	$-518.20 \pm 2.95$
$E_{pb}$	$1980.03 \pm 5.06$	$1333.38 \pm 4.70$	$1045.53 \pm 4.09$
SASA	$-111.88 \pm 0.12$	$-62.96 \pm 0.12$	$-65.81 \pm 0.11$
$E_{MM-PBSA}$	$403.84 \pm 2.82$	$-157.51 \pm 2.29$	$33.11 \pm 1.95$

The energy per residue was investigated to provide more detail of the binding (Figs. 7A–7C). The residues were picked from residues at the interface that were less than 3.5 Å away from another monomer. The residues are shown in Fig. 8 that had a high impact on the binding by having higher energy than 70 kJ/mol or less than -70 kJ/mol.



**Fig. 7** Energy per residues of key residues having energy more than 70 kJ/mol or less than -70 kJ/mol of chain A (blue) and chain B (green): (A)  $\alpha\alpha$  interface; (B)  $\beta 1\beta 1$  interface; (C)  $\beta 2\beta 2$  interface



**Fig. 8** Surfaces of dimers, where first row is binding dimers, second row shows surface of each monomer when separated, third row is monomers from second row rotated 90° along y-axis to left for chain A and to right for chain B, red shows residues having energy greater than 70 kJ/mol and yellow shows residues having energy less than -70 kJ/mol

The residue pairs from the  $\alpha\alpha$  equatorial domain were quite different. The same N39, V41 and E43 were used to make H-bonds, but instead they bonded with T522, Q524 and M527, respectively. The simulation structures along with the %occupancy showed that these positions were crucial for subunit formation.

The energy of key residues from the  $\alpha\alpha$  interface are shown in Fig. 7A. Twelve residues had high unfavorable binding energies. Even though E529 from chain A formed an ionic bond with K45 from chain B, it had a strong repelling effect with E43 and E63, as shown in Fig. 6. Ten residues had low energies; K24 and K65 were the key residues that held the interaction in the equatorial domain of the  $\alpha\alpha$  dimer.

The  $\beta1\beta1$  interface had less high-energy residues and more low-energy residues to stabilize the binding (Fig. 7B). Furthermore, these unfavorable residues had less energy on average than those from  $\alpha\alpha$ . The interface of  $\beta1\beta1$  had the most low-energy residues among these three dimers. Residues E240, E257 and E259 from chain A and K268 and R270 from chain B were the key residues stabilizing the binding in the apical domain (Fig. 8). E390 and E393 from the chain B intermediate domain created an interaction with R120 and K124 from the chain A equatorial domain. It was the only intermediate-equatorial domain interaction among these three

dimers. However, the R120 and K124 energies per residue were positive because of the repelling effect with R397 from chain B. Residues K45 and K24 were important in holding the equatorial domain by interacting with E5, E529 and E532. There were some high-energy residues in the equatorial domain namely E43, E63, D66 and E69 from chain B that were almost the same positive set in all three dimers.

For the  $\beta2\beta2$  interface, most residues had moderate energy, with only 6 residues having energy greater than 70 kJ/mol and 6 residues having energy less than -70 kJ/mol, as shown in Figs. 7C and 8. Most of the stabilization at the interface came from chain B.

In conclusion, the molecular dynamic structures of  $\beta1\beta1$  and  $\beta2\beta2$  were more flexible than for  $\alpha\alpha$ . The binding free energy of  $\beta1\beta1$  was the lowest, mostly due to ionic bonds at the interface, followed by  $\beta2\beta2$  and  $\alpha\alpha$ . Despite all the subunits having similar interface residues,  $\alpha$  subunits had a high binding energy because of the many strong repulsion effects of the charged residues that might explain the oligomer disassembly phenomenon in the presence of  $\alpha$  subunits and no  $\alpha$  homo-oligomers being present *in vitro*. On the other hand,  $\beta$  subunits had more favorable interactions and had low binding energies, which supported the *in vitro* behavior of the  $\beta$  subunits forming homo-tetradecameric structures.

## Conflict of Interests

The authors declare that there are no conflicts of interests.

## Acknowledgements

This work was supported by the Thailand Research Fund through the Royal Golden Jubilee Ph.D. program (Grant No. PHD/0074/2556), the Kasetsart University Research and Development Institute (KURDI), Bangkok, Thailand (FF(KU) 25.64) and the Center for Advanced Studies in Nanotechnology for Chemical, Food and Agricultural Industries, the KU Institute for Advanced Studies, Kasetsart University.

## References

- Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindahl, E. 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2: 19–25. doi.org/10.1016/j.softx.2015.06.001
- Bai, C., Guo, P., Zhao, Q., et al. 2015. Protomer roles in chloroplast chaperonin assembly and function. *Mol. Plant* 8: 1478–1492. doi.org/10.1016/j.molp.2015.06.002
- Bonvin, A., Mark, A., van Gunsteren, W. 2000. The GROMOS96 benchmarks for molecular simulation. *Comput. Phys. Commun.* 128: 550–557. doi.org/10.1016/S0010-4655(99)00540-8
- Boston, R.S., Viitanen, P.V., Vierling, E. 1996. Molecular chaperones and protein folding in plants. *Plant Mol. Biol.* 32: 191–222. doi.org/10.1007/BF00039383
- Bukau, B., Horwich, A.L. 1998. The Hsp70 and Hsp60 chaperone machines. *Cell* 92: 351–366. doi.org/10.1016/S0092-8674(00)80928-9
- Chang, C., Marshall, N., Feldmann, B., Joachimiak, A., Midwest Center for Structural Genomics (MCSG). 2015. Crystal structure of chaperonin GroEL from 5DA8. Protein Data Bank. <https://www.rcsb.org/structure/5DA8>, 20 August 2021.
- Dickson, R., Weiss, C., Howard, R.J., Alldrick, S.P., Ellis, R.J., Lorimer, G., Azem, A., Viitanen, P.V. 2000. Reconstitution of higher plant chloroplast chaperonin 60 tetradecamers active in protein folding. *J. Biol. Chem.* 275: 11829–11835. doi.org/10.1074/jbc.275.16.11829
- Emanuelsson, O., Nielsen, H., Heijne, G.V. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8: 978–984. doi.org/10.1110/ps.8.5.978
- Gething, M.J., Sambrook, J. 1992. Protein folding in the cell. *Nature* 355: 33–45. doi.org/10.1038/355033a0
- Guex, N., Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-Pdb viewer: An environment for comparative protein modeling. *Electrophoresis* 18: 2714–2723. doi.org/10.1002/elps.1150181505
- Hayer-Hartl, M.K., Martin, J., Hartl, F.U. 1995. Asymmetrical interaction of GroEL and GroES in the ATPase cycle of assisted protein folding. *Science* 269: 836–841. doi.org/10.1126/science.7638601
- Hemmingsen, S.M., Ellis, R.J. 1986. Purification and properties of ribulosebiphosphate carboxylase large subunit binding protein. *Plant Physiol.* 80: 269–276. doi.org/10.1104/pp.80.1.269
- Horwich, A.L., Fenton, W.A., Chapman, E., Farr, G.W. 2007. Two families of chaperonin: Physiology and mechanism. *Annu. Rev. Cell Dev. Biol.* 23: 115–145. doi.org/10.1146/annurev.cellbio.23.090506.123555
- Huang, W., Lin, Z., van Gunsteren, W.F. 2011. Validation of the GROMOS 54A7 force field with respect to  $\beta$ -peptide folding. *J. Chem. Theory Comput.* 7: 1237–1243. doi.org/10.1021/ct100747y
- Kumari, R., Kumar, R., Consortium, O., Lynn, A. 2014. g\_mmpbsa-A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* 54: 1951–1962. doi.org/10.1021/ci500020m
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26: 283–291. doi.org/10.1107/S0021889892009944
- Martel, R., Cloney, L.P., Pelcher, L.E., Hemmingsen, S.M. 1990. Unique composition of plastid chaperonin-60:  $\alpha$  and  $\beta$  polypeptide-encoding genes are highly divergent. *Gene* 94: 181–187. doi.org/10.1016/0378-1119(90)90385-5
- Nishio, K., Hirohashi, T., Nakai, M. 1999. Chloroplast chaperonins: Evidence for heterogeneous assembly of  $\alpha$  and  $\beta$  Cpn60 polypeptides into a chaperonin oligomer. *Biochem. Biophys. Res. Commun.* 266: 584–587. doi.org/10.1006/bbrc.1999.1868
- Saibil, H.R., Fenton, W.A., Clare, D.K., Horwich, A.L. 2013. Structure and allostery of the chaperonin GroEL. *J. Mol. Biol.* 425: 1476–1487. doi.org/10.1016/j.jmb.2012.11.028
- Schroda, M. 2004. The *Chlamydomonas* genome reveals its secrets: Chaperone genes and the potential roles of their gene products in the chloroplast. *Photosynth. Res.* 82: 221–240. doi.org/10.1007/s11120-004-2216-y
- Tina, K.G., Bhadra, R., Srinivasan, N. 2007. PIC: Protein interactions calculator. *Nucleic Acids Res.* 35: W473–W476. doi.org/10.1093/nar/gkm423
- Vitlin, A., Weiss, C., Demishtein-Zohary, K., Rasouly, A., Levin, D., Pisanty-Farchi, O., Breiman, A., Azem, A. 2011. Chloroplast beta chaperonins from *A. thaliana* function with endogenous cpn10 homologs *in vitro*. *Plant Mol. Biol.* 77: 105–115. doi.org/10.1007/s11103-011-9797-6
- Zhang, S., Zhou, H., Yu, F., Bai, C., Zhao, Q., He, J., Liu, C. 2016. Structural insight into the cooperation of chloroplast chaperonin subunits. *BMC Biol.* 14: 29. doi.org/10.1186/s12915-016-0251-8