Comparative Modeling of TCP1 Ring Complex (TRiC) From a Psychrophilic Yeast, *Glaciozyma antarctica*

Nur Athirah Yusof[1,2], Farah Diba Abu Bakar[2], Nor Muhammad Mahadi[3], Abdul Munir Abdul Murad[2#]

1 Biotechnology Research Institute, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, MALAYSIA
2 School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, MALAYSIA
3 Malaysia Genome Institute, Jalan Bangi, Bangi 43600, Selangor, MALAYSIA.
Corresponding author. E-Mail: munir@ukm.edu.my; Tel: +603-89215696; Fax: +603-89252698

ABSTRACT The TRiC chaperonin belongs to the group II chaperonin that is ubiquitously expressed in the cytosol of archae and eukaryotes. Well known as the complex machinery of protein folding and biogenesis of many cytoskeletal proteins, including tubulin and actin, this chaperonin is indispensable for cell survival as an essential subset of cytosolic proteins requires TRiC for proper folding. Life in extremely cold environment faces energetic challenges to protein folding where psychrophiles have evolved some important cellular adaptations. This indicates that psychrophilic TRiC has undergone positive selection, structural evolution and mechanistic features that distinguish it from other chaperones. The knowledge of this unique complex is in its infancy, therefore we illustrate a systematic tertiary model of the first eukaryotic psychrophilic chaperonin that open the platform to understand the secrets of its folding chamber. The unique ability displayed by the psychrophilic TRiC offers a great opportunity to study the relationship between protein function and structure in terms of stability, flexibility and dynamic conformation.

KEYWORDS: Chaperonin; *Glaciozyma antarctica*; psychrophile; homology modeling; cylindrical architecture

Full Article - *Industrial biotechnology*
Received 30 August 2017 **Online** 28 November 2017

TRANSACTIONS ON SCIENCE AND TECHNOLOGY

INTRODUCTION

TRiC complex has been known to be the gatekeeper of cellular protein homeostasis especially in assisting the large number of protein folding of multiple structural classes and complex topology (Dekker *et al.*, 2008; Thulasiraman *et al.*, 1999; Yam et al., 2008). Hence, the cells' metabolic systems could have gone haywire if the TRiC complex is structurally prone to be compromised by temperature fluctuation. Our subject, *Glaciozyma antarctica* is a psychrophilic yeast, which was isolated from the surface of the sea ice of Antarctica near the Casey Research Station (Hashim *et al.*, 2013). It is considered as an obligate psychrophile as it has an optimal temperature for growth at 12ºC, a maximal temperature for growth at about 20ºC and a minimal temperature for growth at 0ºC and lower (Boo *et al.*, 2013), as defined by psychrophilic term by Moyer and Morita (2007). Therefore, this paper illustrates the first tertiary model of a psychrophilic yeast chaperonin from *G. antarctica* using comparative modeling followed by model validation and structure assessment. The structure of the chaperonin will open a platform for structural study at the sequence level and tertiary structures where residue substitutions and deletions effected several interactions such as hydrophobicity, hydrogen bonding, ionic, aromatic-aromatic, aromatic-sulphur and cation-pi interactions in the intra- and inter-protein that are related to structure flexibility, stability and dynamic conformation will be accessible for protein folding analysis.

METHODOLOGY

Culturing G. antarctica

The isolated *G. antarctica* culture was obtained from the courtesy of School of Biological Sciences, Universiti Sains Malaysia (Hashim *et al.*, 2013) The cells were cultured in yeast peptone dextrose broth (10% (w/v) yeast extract, 20% (w/v) peptone, and 20% (w/v) dextrose) at 12°C until until its OD_{600} reached approximately 0.6-0.8.

Total RNA extraction of G. antarctica

RNA extraction was done using TRIzol® Reagent (Invitrogen, USA) as described by Bharudin *et al.* (2014). The concentration and purity of total RNA were measured using a Nanodrop spectrophotometer at 260 nm (ThermoScientific, USA).

Cloning and sequence analysis

The cDNA amplification of all eight *G. antarctica* TRiC subunits were done and cloned into pGEMT-T Easy vector (PROMEGA, USA). The sequence of the *G. antarctica* TRiC subunits have been deposited in the GenBank™ database under these accession numbers (KU659477 TRiC_alpha; KU659478 TRiC_beta; TRiC_gamma KU659479; KU659480 TRiC_delta; KU659481 TRiC_epsilon; KU659482 TRiC_zeta; KU659483 TRiC_eta and KU659484 TRiC_theta).

Template recognition and alignment

Using the amino acid sequences of the *G. antarctica* TRiC subunits, the sequences were compared to known structures stored in Protein Data Bank (PDB) (http://www.rcsb.org/pdb/home/home.do). A search with BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against the database for optimal local alignments with the query, gave a list of known protein structures that matches the sequence using PSI-BLAST as the program selection.

Model building for the target based on the 3D structure of the template

After the target-template alignment, the next step in the comparative modeling was the model building. Protein modeling was done using SWISS-MODEL (http://swissmodel.expasy.org/). In this work, a complex structure from PDB with accession number 4V81 was selected as the template.

Model validation

PROCHECK was used to check for proper protein stereochemistry which were the symmetry and geometry checks; chirality, bond lengths, bond angles, and torsion angles (Morris *et al.*, 1992). VERIFY3D was used to analyse the compatibility of an atomic model (3D) with its own amino acid sequence (1D) (Bowie *et al.* 1991). The scores range from -1 (bad score) to +1 (good score). All these validation programs were accessed using Structural Analysis and Verification Server (SAVeS) server at (http://services.mbi.ucla.edu/SAVES/). Atomic Non-Local Environment Assessment (ANOLEA) was used to detect errors in the build protein models by performing energy calculations at the atomic level in protein structures (Melo *et al.*, 1997).

Protein structure analysis

Interactive visualisation and molecular structure analysis were performed using UCSF Chimera (http://www.cgl.ucsf.edu/chimera/).

RESULT AND DISCUSSION

Structure overview

In this study, the structure of *S. cerevisiae* TRiC chaperonin (PDB ID: 4V81) was selected as template as it has a better diffraction quality of 3.4 Å compared to *Bos taurus* TRiC diffraction which was 4 Å (PDB ID: 3IYG). All of the *G. antarctica* TRiC subunits have more than 50% sequence identity to template from PDB. Even if homology modeling is generally much less accurate than experimental methods, it is arguable that this method can be helpful in proposing and testing hypothesis in molecular biology such as hypotheses about the location of ligand binding sites, substrate specificity, function annotation and drug design (Ginalski *et al.*, 2004, Takeda-Shitaka *et al.*, 2004).

The constructed models were evaluated based on the Global Model Quality Estimation (GMQE) and Qualitative Model Energy Analysis (QMEAN4) values. The resulting GMQE score is expressed as a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability. From all the models that were built, the *G. antarctica* subunits have identical constructs with three distinct domains; apical, intermediate and equatorial.

Table 1. A summary of data collection and refinement statistics

Ga TRiC unit	% Sequence identity	GMQE	QMEAN	RAMACHANDRAN PLOT	VERIFY3D	ANOLEA	RMSD
α	71% 4V81.A	0.81	0.688	97% (Favored) 2.4% (Allowed)	90.74%	13.15%	0.776
β	69% 4V81.B	0.79	0.703	94% (Favored) 5.2% (Allowed)	83.56%	21.66%	0.561
γ	63% 4V81.C	0.74	0.714	96.1% (Favored) 2.9% (Allowed)	87.33%	24.17%	1.243
δ	65% 4V81.D	0.76	0.653	93.8% (Favored) 5.4% (Allowed)	85.58%	20.58%	0.588
ε	65% 4V81.E	0.78	0.689	97.5% (Favored) 1.9% (Allowed)	82.04%	17.58%	0.677
ζ	54% 4V81.F	0.69	0.597	92.9% (Favored) 6.2% (Allowed)	81.49%	37.72%	0.666
η	67% 4V81.G	0.77	0.641	97.5% (Favored) 2.5% (Allowed)	85.74%	15.02%	0.791
θ	54% 4V81.H	0.73	0.608	95.4% (Favored) 3.8% (Allowed)	77.86%	30.34%	0.776

The models were evaluated using PROCHECK, Verify3D and ANOLEA. A summary of data collection and refinement statistics were presented in Table 1. PROCHECK analysis showed that all build models have at least 99% in favored and allowed regions. Protein model GaTRiCη has the best evaluation as all residues were in favored and allowed regions followed by GaTRiCα and GaTRiCε with both models have 99.4% residues in the favored and allowed regions. As all models were at 99% and above, these values succeeded the standard evaluation where at least 95% of the residues should be in the favored and allowed regions. Besides that, model verification using Verify3D showed that all constructed models have positive scores between 81.49% (GaTRiCζ) and 90.74% (GaTRiCα). Using ANOLEA, the highest percentage of total amino acids with high energy was GaTRiCζ and GaTRiCθ with 37.73% and 30.34% respectively. Other models have a total percentage of lower than 30% where GaTRiCα has the lowest value of 13.15%. The model evaluation using three different software showed that all constructed models of GaTRiC subunits were acceptable for further analysis except for GaTRiCζ and GaTRiCθ that fulfilled the PROCHECK and Verify3D analysis, however values slightly below the standard acceptance of ANOLEA of 30% and below.

G. antarctica TRiC constructed models

The whole structure of *G. antarctica* TRiC was oriented accordingly based from the crystal structure of *S. cerevisiae* TRiC (PDB: 4V81). The order within each of the *G. antarctica* TRiC ring was consistent with the order defined for *S. cerevisiae* TRiC. The *G. antarctica* TRiC complex was built by superimposed of *G. antarctica* TRiC subunits and *S. cerevisiae* TRiC subunits where no pair exceeds 2 Å which reflected the high homology between both yeasts TRiC structures. GaTRiCα, GaTRiCβ and GaTRiCδ models have the best quality for every evaluation analysis. Hence, the constructed models of GaTRiCα, GaTRiCβ and GaTRiCδ were described in Table 2.

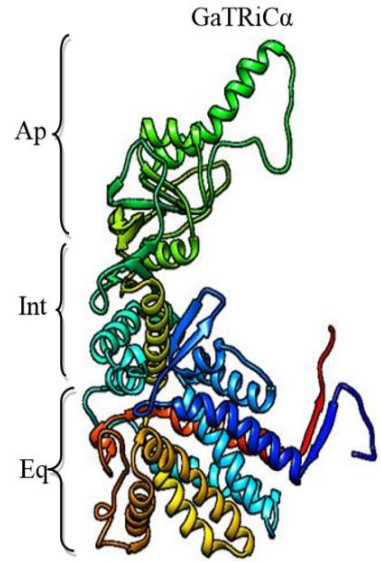
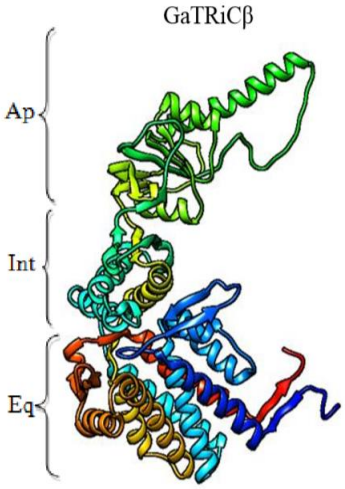
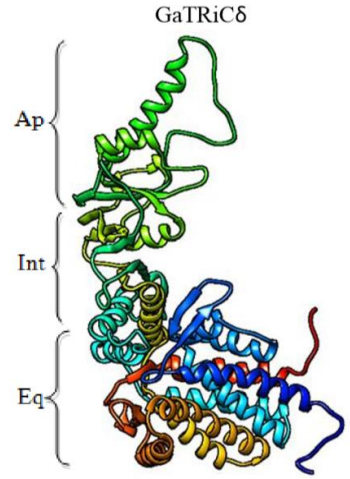
CONCLUSION

Comparative modeling and model assessment of *G. antarctica* TRiC successfully create an equitable model of the whole structure of *G. antarctica* TRiC for structure analysis. It is also compatible to be used as starting models for solving structures of TRiC from X-ray crystallography, NMR and electron microscopy. The *G. antarctica* TRiC complex was built by superimposed of *G. antarctica* TRiC subunits and *S. cerevisiae* TRiC subunits where no pair exceeds 2 Å which reflected the high homology between both yeasts TRiC structures.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Science, Technology and Innovation, Malaysia (MOSTI), for funding this project under grant numbers 10-05-16-MB002 and 07-05-MGIGMB014. The authors would also like to thank Malaysia Genome Institute for the facilities, services and training provided during this project.

Table 2. Constructed models for *G. antarctica* TRiC

Protein structure	Description
	**Equatorial domain** (Eq): β1 (18-22), α1 (23-43), β2 (51-63), α2(65-73), α3 (77-93), α4 (98-117), α5 (121-143), α14 (422-438), α15 (440-465), α16 (469-484), β17 (508-512), β18 (514-518), α17 (519-537), β19 (539-543) **Intermediate domain** (Int): α6 (155-165), α7 (168-171), α8 (172-187), β3 (188-191), β4 (195-197), β5 (204-210), β6 (216-219), β7 (221-225), β15 (382-387), α13 (389-413), β16 (416-418) **Apical domain** (Ap): β8 (236-238), β9 (239-245), α9 (265-289), α10 (292-296), α11 (302-311), β10 (313-318), α12 (321-332), β11 (350-352), β12 (354-357), β13 (358-364), β14 (367-373)
	**Equatorial domain** (Eq): β1 (11-14), α1 (15-35), β2 (44-47), β3 (53-57), α2 (59-67), α3 (71-89), α4 (92-114), α5 (117-138), α14 (406-420), α15 (424-449), α16 (452-465), β14 (470-474), β15 (478-491), α17 (492-510), β16 (513-516) **Intermediate domain** (Int): α6 (146-158), α7 (165-168), α8 (169-184), β4 (190-200), α9 (202-205), β5 (206-209), β6 (213-216), β12 (365-372), α13 (373-397), β13 (400-403) **Apical domain** (Ap): β7 (228-234), α10 (255-279), β8 (283-286), α11 (292-301), β9 (304-307), α12 (311-320), β10 (338-349), β11 (351-357)
	**Equatorial domain** (Eq): α1 (22-42), β1 (50-54), β3 (58-62), α2 (64-72), α3 (76-92), α4 (97-117), α5 (121-140), α9 (419-433), α10 (437-462), α11 (465-478), β15 (483-486), β16 (491-494), β17 (500-503), β18 (504-523) **Intermediate domain** (Int): α6 (151-162), α7 (171-184), β4 (195-204), β5 (216-219), β11 (353-360), β12 (362-368), β13 (378-385), α8 (386-410), β14 (413-415) **Apical domain** (Ap): β6 (232-234), β7 (236-285), α8 (263-285), β8 (288-293), α9 (303-312), β9 (315-320), α10 (321-332), α11 (340-342), α12 (345-347), β10 (349-352)

REFERENCES

[1] Bharudin, I., Zaki N. Z., Bakar, F. D. A., Illias, R. M. & Murad, A. M. A. (2014) Comparison of RNA extraction methods for transcript analysis from the psychrophilic yeast, *Glaciozyma antarctica*. *Malaysian Applied Biology*, **43**, 71-79

[2] Boo, S. Y., Wong, C. M. V. L., Rodrigues, K. F., Najimudin, N., Murad, A. M. A. & Mahadi, N. M. (2013). Thermal stress responses in Antarctic yeast, *Glaciozyma antarctica* PI12, characterized by real-time quantitative PCR. *Polar Biology*, **36**(3), 381-389

[3] Bowie, J. U., Lüthy, R. & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, **253**(5016), 164-170.

[4] Dekker, C., Stirling, P. C., McCormack, E. A., Filmore, H., Paul, A., Brost, R. L., Costanzo, M., Boone, C., Leroux, M. R. & Willison, K. R. (2008). The interaction network of the chaperonin CCT. *EMBO Journal*, **27,** 1827-1839.

[5] Ginalski, K., Zhang, H. & Grishin, N. V. (2004). Raptor protein contains a caspase-like domain. *Trends in Biochemical Sciences*, **29**(10), 522-524.

[6] Goddard, T. D., Huang, C. C. & Ferrin, T. E. (2007). Visualizing density maps with UCSF Chimera. *Journal of Structural Biology*, **157,** 281-287.

[7] Hashim, N. H. F., Bharudin, I., Nguong, D. L. S., Higa, S., Bakar, F. D. A., Nathan, S., Rabu, A., Kawahara, H., Illias, R. M., Najimudin, N., Mahadi, N. M. & Murad, A. M. A. (2013) Characterization of Afp1, an antifreeze protein from the psychrophilic yeast *Glaciozyma antarctica* PI12. *Extremophiles*, **17**, 63-73

[8] Melo, F., Devos, D., Depiereux, E. & Feytmans, E. (1997). ANOLEA: a www server to assess protein structures. *Proceedings on International Conference Intelligent System for Molecular Biology*, **5**, 187-190.

[9] Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins*, **12**(4), 345-364.

[10] Moyer, C. L. & Morita, R. Y. (2007). Psychrophiles and psychrotrophs. *Encyclopedia of life sciences*. New York: Wiley.

[11] Takeda-Shitaka, M., Takaya, D., Chiba, C., Tanaka, H. & Umeyama, H. (2004). Protein structure prediction in structure based drug design. *Current Medicinal Chemistry*, **11**(5), 551-558.

[12] Thulasiraman, V., Yang, C. F. & Frydman, J. (1999). *In vivo* newly translated polypeptides are sequestered in a protected folding environment. *EMBO Journal*, **18,** 85-95.

[13] Yam, A. Y., Xia, Y., Lin, H. T., Burlingame, A., Gerstein, M. & Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nature Structural & Molecular Biology*, **15**, 1255-1262.