Received: 6 November 2020 | Revised: 14 November 2020 | Accepted: 17 November 2020

DOI: 10.1002/pro.4002

PROTEIN STRUCTURE REPORTS

Structural insights into a C2 domain protein specifically found in tardigrades

Yohta Fukuda | **Tsuyoshi Inoue**

Graduate School of Pharmaceutical Science, Osaka University, Osaka, Japan

Correspondence
Tsuyoshi Inoue, Graduate School of Pharmaceutical Science, Osaka University, Yamadaoka 1-6, Suita, Osaka, 565-0871, Japan.
Email: t_inoue@phs.osaka-u.ac.jp

Funding information
Japan Science and Technology Agency, Grant/Award Number: JPMJAX191C; Japan Society for the Promotion of Science, Grant/Award Number: 20K15971

Abstract

Some tardigrades can survive extremely desiccated conditions through transition into a state called anhydrobiosis. Anhydrobiotic tardigrades have proteins unique to them and they are thought to be keys to the understanding of unusual desiccation resistance. In fact, previous transcriptome data show that several tardigrade-specific proteins are significantly upregulated under desiccated conditions. However, their physiological roles and chemical properties have been ambiguous because they show low or no similarity of amino acid sequences to proteins found in other organisms. Here, we report a crystal structure of one of such proteins. This protein shows a β-sandwich structure composed of 8 β-strands, three Ca^{2+}-binding sites, and hydrophobic residues on Ca^{2+}-binding (CBD) loops, which resemble characteristics of C2 domain proteins. We therefore conveniently describe this protein as tardigrade C2 domain protein (TC2P). Because the C2 domain functions as a Ca^{2+}-mediated membrane docking module, which is related to signal transduction or membrane trafficking, TC2Ps may play a role in Ca^{2+}-triggered phenomenon under desiccated situations. Our finding provides not only structural insights into a newly discovered desiccation-related protein family but also insights into the evolution and diversity of C2 domain proteins.

KEYWORDS

C2 domain, calcium binding, crystal structure, tardigrade

1 | INTRODUCTION

Tardigrades are microscopic organisms ubiquitously found on Earth.[1] Some terrestrial tardigrades can survive extremely desiccated conditions through transition into a state called anhydrobiosis with undetectable metabolism.[2,3] In preparation for anhydrobiosis, the bodies of tardigrades shrink to form a so-called "tun." The tun shows tolerances to high (151°C)[4] or low (−273°C)[5] temperature, exposure to high energy radiations,[6–8] vacuum,[9,10] high pressure,[11,12] and toxic chemicals.[13,14] Moreover, anhydrobiotic tardigrades are famous for the survival record in space.[15] To understand the molecular basis of the extraordinary abilities of anhydrobiotic tardigrades, genomics and transcriptome analyses have recently uncovered tardigrade-specific proteins.[16,17] These proteins show very low or no amino acid sequence similarities to those of well-studied proteins from other

organisms; therefore, their functions are ambiguous. A protein family, which we here designated tardigrade C2 domain protein (TC2P), seems to be conserved in tardigrades and is not found in phyla other than Tardigrada. The earlier transcriptome studies[16,17] also show that in two different anhydrobiotic tardigrades, *Ramazzottius varieornatus* and *Hypsibius exemplaris* (the same strain as *Hypsibius dujardini* described in[17], which has been recently renamed *Hypsibius exemplaris*.[18] We use the new terminology in this report), some TC2Ps are upregulated when they are exposed to dried conditions. Therefore, TC2Ps are promising candidates for anhydrobiosis-related proteins. However, due to its no amino acid sequence similarities to known proteins, functional analyses of the proteins at molecular and atomic levels are difficult. Here, we performed the structural analysis on a TC2P from *R. varieornatus* strain YOKOZUNA-1. Our structure provides insights into chemical properties and a possible physiological role of TC2Ps, which will help future biochemical studies.

2 | RESULTS AND DISCUSSION

The crystal structure of the TC2P was determined at a resolution of 2.50 Å by a multi-crystal native single-wavelength anomalous dispersion (native SAD) method[19] using calcium atoms (Table 1) because of its low sequence similarity to known proteins. A high-resolution dataset was also collected from another crystal and its structure was solved at 1.70 Å resolution. The final R_{work} and R_{free} values of the high-resolution structure are 0.175 and 0.212, respectively. The overall structure of TC2P shows a β-sandwich core composed of 4 + 4 β-strands (Figure 1a). The Dali server[20] reports that the obtained TC2P model is especially similar to the structures of the calcium-phospholipid binding domain from cytosolic phospholipase A2 (cPLA2 C2 domain; PDB ID: 1RLW)[21] and the C2A domain of otoferlin (PDB ID: 3L9B)[22] with Z-scores higher than 10. Although the amino acid sequence of the TC2P is distinct from those of the C2 domain proteins (Figure S1), superimposition of the TC2P structure on the cPLA2 C2 domain and the otoferlin C2A domain reveals that the TC2P has the same fold and topology as these proteins (Figure 1b). Furthermore, all other top-hit structures at the Dali server belong to C2 domain proteins. These results show that the TC2P is a previously unknown C2 domain protein. The C2 domain functions as a Ca^{2+}-mediated membrane-docking module,[23,24] which is related to signal transduction or membrane trafficking. Some C2 domain proteins such as cPLA2 C2 domain have hydrophobic residues on the Ca^{2+}-binding (CBD) loops (CBD loop1 and 2) positioned above the Ca^{2+}-binding sites, which facilitate direct interaction with membrane.[23] The TC2P structure also

TABLE 1 Data collection and refinement statistics

Data collection at SPring-8 BL44XU		
	High-resolution data	**Native SAD data**
No. of crystals	1	2
Wavelength (Å)	0.9000	1.9000
Total images	1,800	6,600
Space group	$P4_12_12$	
Unit cell *a*, *b*, *c* (Å)	62.92, 62.92, 100.5	62.69, 62.69, 100.9
Resolution range (Å)	40.7–1.70 (1.73–1.70)[a]	44.3-2.50 (2.60–2.50)
Total no. of reflections	299,105 (16,254)	323,225 (37,127)
No. of unique reflections	22,431 (1,128)	7,373 (796)
Completeness (%)	98.4 (97.1)	99.3 (98.2)
Redundancy	13.3 (14.4)	43.8 (46.6)
$\langle I/\sigma(I)\rangle$	16.2 (2.5)	24.0 (13.5)
R_{meas} (all I+ & I−)	0.091 (1.290)	0.256 (0.932)
R_{meas} (within I+/I−)	0.095 (1.334)	0.256 (0.938)
$CC_{1/2}$	0.999 (0.870)	0.998 (0.990)
Refinement		
Resolution range (Å)	39.3–1.70 (1.76–1.70)	
Completeness (%)	97.9 (96.5)	
No. of reflections, working set	22,402 (2149)	
No. of reflections, test set	1,101 (92)	
R_{work}/R_{free}	0.175/0.212 (0.249/0.286)	
No. of non-H atoms		
Protein	1,449	
Ca ion/Trehalose	3/23	
Water	202	
R.m.s. deviation bonds (Å), angles (°)	0.006, 0.945	
Average *B* factors (Å^2)	29.4	
Protein	27.9	
Ca ion/Trehalose	19.9/54.6	
Water	37.4	
Ramachandran favored/allowed/disallowed (%)	96.43/3.57/0	
PDB code ID	7DF2	

[a]Statistics for the highest-resolution shell are shown in parentheses.

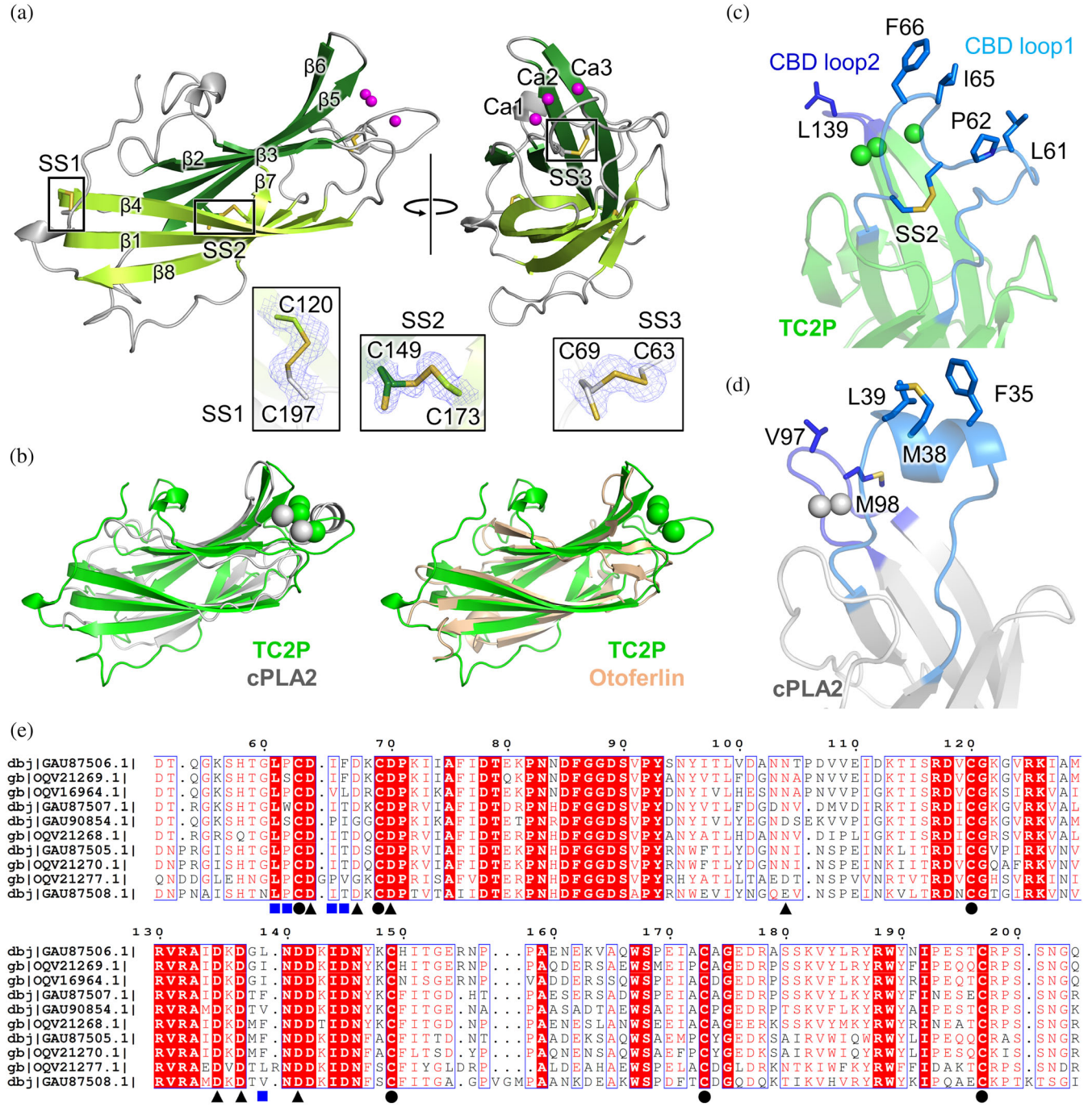

FIGURE 1 Crystal structure of TC2P. (a) Overall structure of TC2P. Ca ions are shown by magenta spheres. Disulfide bonds are shown in insets with $2mF_o$-DF_c maps at 1 σ. (b) Structural comparison of TC2P with C2 domain proteins: cPLA2 (PDB code ID: 1RLW) and otoferlin (PDB code ID: 3L9B). Root-mean-square-deviations are 2.8 and 2.6 Å for 1RLW (123 C^{α} atoms) and 3L9B (126 C^{α} atoms), respectively. Ca ions are illustrated by spheres. (c) Hydrophobic residues on the Ca^{2+}-binding (CBD) loops of TC2P. (d) Hydrophobic residues on the CBD loops of the C2 domain in cPLA2. (e) Sequence alignment of TC2Ps. Black circles: cysteine residues forming the S—S bonds. Black triangles: ligand residues to Ca^{2+} ions. Blue squares: hydrophobic residues on the CDB loops. GenBank accession IDs starting from GAU and OQV mean sequences from *R. varieornatus* and *H. exemplaris*, respectively. GAU87506.1 is the TC2P studied here

displays hydrophobic residues on its CBD loops and they are conserved among TC2Ps (Figure 1c–e), suggesting membrane docking of TC2Ps.

Because of the presence of Ca^{2+} ions in our crystallization condition, we could observe a Ca^{2+}-bound state of the TC2P (Figure 2a, Table S1). These Ca^{2+} ions are verified by anomalous signals that are weak but stronger than those from sulfur atoms (Figure S2). There are three Ca^{2+}-binding sites in the TC2P structure, which resembles aspartate clusters for multiple Ca^{2+}-binding in

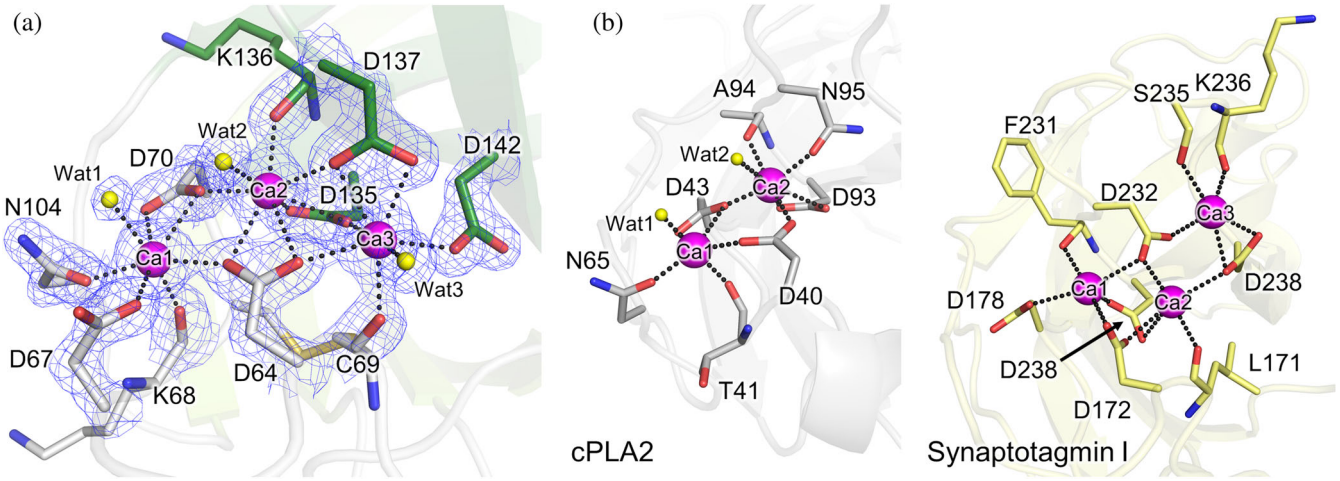

FIGURE 2 Ca^{2+}-binding sites of TC2P. (a) Aspartate cluster for Ca^{2+}-binding. $2mF_o$-DF_c map is contoured at 2.0 σ. Ca^{2+} ions and ligand water molecules are represented by magenta and yellow spheres. Coordination bonds are shown by dotted black lines. Colors of residues are in accordance with Figure 1a. (b) Well-characterized Ca^{2+}-binding sites in the C2 domains from cPLA2 (PDB code ID: 1RLW) and synaptotagmin I (PDB code ID: 1BYN). Other examples of Ca^{2+}-binding sites in C2 domain proteins are shown in Figure S3

known C2 domain proteins (Figure 2b, Figure S3).[21,25] The Ca^{2+} ions in the TC2P structure are coordinated by 7 (at Ca1 and Ca3) or 8 (at Ca2) ligands with coordination distances of 2.32–2.68 Å, which are typical coordination spheres of Ca^{2+}-binding sites in proteins.[26] The Ca^{2+}-ligand aspartate residues are conserved among TC2Ps (Figure 1e), indicating that the Ca^{2+}-binding ability is an important characteristic of TC2Ps.

A striking feature of TC2Ps is the presence of three disulfide bonds, which are not observed in other known C2 domain proteins (Figure S1). Two disulfide (S—S) bonds (SS1 and SS3) are formed near each end of the β-sandwich structure (Figure 1). Additionally, one S—S bond (SS2) is formed between β6 and β7. Because the cysteine residues forming these S—S bonds are conserved among TC2Ps (Figure 1e), it is likely that the bonds have biological functions. Especially, SS3 is located on CBD loop1 and close to Ca^{2+}-binding sites. In our structure, Cys69 shows two different conformations (Figure 1a, inset). One forms the S—S bond and the other is in a free state. Although the observed alternative conformation of Cys69 could be induced by X-ray radiation damages during data collection, our structure implies that this S—S bond is redox sensitive. Because CBD loop1 in the TC2P structure is a random coil longer than those in other C2 domain proteins (Figure 1c,d), reduction of the S—S bond will give the TC2P CDB loop high flexibility. Moreover, Cys69 directly binds to Ca3 through its carbonyl O atom (Figure 2a). These observations suggest that the redox states of SS3 can affect the Ca^{2+}-binding mode or affinity to Ca^{2+} ions. Cys149 at SS2 also shows two conformations (Figure 1a, inset), indicating that SS2 is an additional modulator of the TC2P structure. Because tardigrades are exposed to severe oxidative stresses under desiccated conditions, a redox sensitive switch such as a disulfide bond may be useful to rapidly respond to environmental changes. Biochemical analyses of TC2Ps, such as verification of the hypothesis that TC2Ps can bind membrane lipids and that the S—S bond functions as a redox switch module, are currently under way.

While many tardigrade-specific proteins with low or no sequence similarities to well-characterized proteins are found, structural biology has revealed that some of their three-dimensional structures are unexpectedly quite similar to those of known proteins. For example, secretory abundant heat soluble (SAHS) proteins are only found in some anhydrobiotic tardigrades,[27] but their structures are almost the same as those of fatty acid binding proteins (FABPs) and SAHS proteins can bind to fatty acids and other hydrophobic compounds as FABPs do.[28,29] We here showed that the TC2P is another example for "much different primary structure but the same tertiary structure." More structural and functional analyses on tardigrade proteins may answer to why tardigrade proteins have evolved in such a unique way.

3 | MATERIALS AND METHODS

3.1 | Sequence and structural alignment

Sequence alignment was performed by Clustal Omega.[30] The alignment figure was generated by ESpript.[31] The Dali server was used to find similar fold proteins.

3.2 | Protein expression and purification

The GenBank accession ID of the gene for a *R. varieornatus* TC2P, which was used in this study, is GAU87506.1. A synthesized and codon optimized DNA of the TC2P without the signal peptide region ($TC2P_{21–213}$) was purchased from GenScript and cloned into a pET28a vector. A 6×His tag followed by a TEV protease site (ENLYFQG) was attached at the N-terminus of TC2P for purification. Its complete sequence is shown in Appendix S1 of Supporting Information. The protein was expressed in *Escherichia coli* Shuffle T7 (New England Bio Labs, Ipswich, MA, USA). At culture optical density of ∼0.6, 0.5 mM isopropyl β-D-1 thiogalactopyranoside was added to induce expression. After 18 hr at 18°C, the bacterial pellet was collected and then sonicated in a buffer containing 20 mM Tris–HCl pH 8, 200 mM NaCl, and a cOmplete Protease Inhibitor Cocktail tablet (Roche, Basel, Basel-Stadt, Switzerland). The resulting solution was centrifuged and supernatant was purified using a HiTrap TALON column (GE healthcare, Chicago, IL). The sample was incubated with TEV protease and imidazole was removed through dialysis against 20 mM Tris–HCl pH 8 overnight at 4°C. The sample was then loaded on a HisTrap column (GE healthcare) equilibrated by 20 mM Tris–HCl pH 8 and 40 mM imidazole. The flowthrough fraction was further purified using a Hiload 16/60 Superdex 75 gel filtration column (GE healthcare) in 20 mM Tris–HCl buffer pH 8.

3.3 | Crystallization

Crystallization was performed by the sitting drop and hanging drop vapor-diffusion method. After more than 3 months, crystals appeared under the condition of 30 mg/mL TC2P, 0.1 M calcium chloride dihydrate, 0.1 M Tris pH 6.5, 13% (wt/vol) polyethylene glycol monomethyl ether 2000 at 20°C. Before the crystals were frozen by liquid nitrogen, they were soaked in the crystallization solutions supplemented by 25% vol/vol glycerol (SAD data) or 15% xylitol and 15% trehalose (high-resolution data). The soaking time should be very short otherwise crystals are dissolved. A trehalose molecule was observed in the crystal structure (Figure S3).

3.4 | X-ray data collection, processing, structure solution, and refinement

X-ray diffraction experiment was performed on the BL44XU beamline of SPring-8, Hyogo, Japan. Diffraction images were collected at 100 K using an EIGER X 16 M detector (Dectris, Philadelphia, PA). For native SAD and high-resolution data collection, X-ray wavelengths were set to 1.9 and 0.9 Å, respectively. The datasets were processed using XDS.[32] The processed data were scaled and merged by Aimless.[33] Phase determination and initial model building was performed by CRANK2.[34] Manual model building was performed using Coot.[35] The program phenix.refine[36] was used for structural refinement. The stereochemical quality of the final model was checked by Molprobity.[37] Data collection and refinement statistics are summarized in Table 1. The coordinate and structure factor files are deposited at the Protein Data Bank (PDB code ID: 7DF2). Raw data is available at Integrated Resource for Reproducibility in Macromolecular Crystallography (https://proteindiffraction.org/).

ACKNOWLEDGMENTS

X-ray diffraction experiments were performed at the Osaka University beamline BL44XU of SPring-8 under the Collaborative Research Program of Institute for Protein Research, Osaka University (Proposal No. 2019A6943, 2019B6943, 2020A6543). The authors would like to acknowledge the beamline staffs for their kind support during data collection. This research was supported by Japan Science and Technology Agency ACT-X "Life and Chemistry" Grant No. JPMJAX191C and JSPS KAKENHI Grant-in-Aid for Young Scientists Grant No. 20K15971 (Y. F.).

AUTHOR CONTRIBUTIONS

Yohta Fukuda: Conceptualization; data curation; formal analysis; funding acquisition; investigation; writing-original draft; writing-review and editing. **Tsuyoshi Inoue:** Writing-review and editing.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

ORCID

Yohta Fukuda https://orcid.org/0000-0002-7386-8201

REFERENCES

1. Nelson DR. Current status of the Tardigrada: Evolution and ecology. Integr Comp Biol. 2002;42:652–659.
2. Keilin D. The problem of anabiosis or latent life: History and current concept. Proc R Soc Lond B. 1959;150:149–191.
3. Møbjerg N, Halberg KA, Jørgensen A, et al. Survival in extreme environments – on the current knowledge of adaptations in tardigrades. Acta Physiol. 2011;202:409–420.
4. Rahm PG. Biologische und physiologische Beiträge zur Kenntnis de Moosfauna Zeitschrift allgemeine. Physiologie. 1921;20:1–35.
5. Becquerel P. La suspension de la vie au dessus de 1/20 K absolu par demagnetization adiabatique de l'alun de fer dans le vide les plus eléve. C R Hebd Seances Acad Sci. 1950;231:261–263.

6. May RM, Maria M, Guimard MJ. Action différentielle des rayons x et ultraviolets sur le tardigrade Macrobiotus areolatus, a l'état actif et desséché. Bull Biol Fr Belg. 1964; 98:349–367.
7. Horikawa DD, Sakashita T, Katagiri C, et al. Radiation tolerance in the tardigrade *Milnesium tardigradum*. Int J Radiat Biol. 2006;82:843–848.
8. Horikawa DD, Cumbers J, Sakakibara I, et al. Analysis of DNA repair and protection in the Tardigrade Ramazzottius varieornatus and *Hypsibius dujardini* after exposure to UVC radiation. PLoS One. 2013;8:e64793.
9. Horikawa DD, Yamaguchi A, Sakashita T, et al. Tolerance of anhydrobiotic eggs of the Tardigrade *Ramazzottius varieornatus* to extreme environments. Astrobiology. 2012;12:283–289.
10. Utsugi K, Noda H. Vacuum tolerance of tardigrades. Proc NIPR Symp Polar Biol. 1995;8:202.
11. Ono F, Saigusa M, Uozumi T, et al. Effect of high hydrostatic pressure on to life of the tiny animal tardigrade. J Phys Chem Solid. 2008;69(9):2297–2300.
12. Seki K, Toyoshima M. Preserving tardigrades under pressure. Nature. 1998;395:853–854.
13. Ramløv H, Westh P. Cryptobiosis in the Eutardigrade Adorybiotus (Richtersius) coronifer: Tolerance to alcohols, temperature and de novo protein synthesis. Zool Anz. 2001; 240:517–523.
14. Horikawa DD, Kunieda T, Abe W, et al. Establishment of a rearing system of the extremotolerant tardigrade *Ramazzottius varieornatus*: A new model animal for astrobiology. Astrobiology. 2008;8:549–556.
15. Jönsson KI, Rabbow E, Schill RO, Harms-Ringdahl M, Rettberg P. Tardigrades survive exposure to space in low earth orbit. Curr Biol. 2008;18:R729–R731.
16. Hashimoto T, Horikawa DD, Saito Y, et al. Extremotolerant tardigrade genome and improved radiotolerance of human cultured cells by tardigrade-unique protein. Nat Commun. 2016;7: 12808.
17. Yoshida Y, Koutsovoulos G, Laetsch DR, et al. Comparative genomics of the tardigrades Hypsibius dujardini and *Ramazzottius varieornatus*. PLoS Biol. 2017;15:e2002266.
18. GĄsiorek P, Stec D, Morek W, Michalczyk Ł. An integrative redescription of *Hypsibius dujardini* (Doyère, 1840), the nominal taxon for Hypsibioidea (Tardigrada: Eutardigrada). Zootaxa. 2018;4415:45–75.
19. Liu Q, Dahmane T, Zhang Z, et al. Structures from anomalous diffraction of native biological macromolecules. Science. 2012; 336:1033–1037.
20. Holm L. DALI and the persistence of protein shape. Protein Sci. 2020;29:128–140.
21. Perisic O, Fong S, Lynch DE, Bycroft M, Williams RL. Crystal structure of a calcium-phospholipid binding domain from cytosolic phospholipase A2. J Biol Chem. 1998;273:1596–1604.
22. Helfmann S, Neumann P, Tittmann K, Moser T, Ficner R, Reisinger E. The crystal structure of the C2A domain of otoferlin reveals an unconventional top loop region. J Mol Biol. 2011;406:479–490.
23. Nalefski EA, Falke JJ. The C2 domain calcium-binding motif: Structural and functional diversity. Protein Sci. 1996;5:2375–2390.
24. Cho W, Stahelin RV. Membrane binding and subcellular targeting of C2 domains. Biochim Biophys Acta Mol Cell Biol Lipid. 2006;1761:838–849.
25. Shao X, Fernandez I, Südhof TC, Rizo J. Solution structures of the Ca^{2+}-free and Ca^{2+}-bound C2A domain of synaptotagmin I: Does Ca^{2+} induce a conformational change? Biochemistry. 1998;3:16106–16115.
26. Zheng H, Cooper DR, Porebski PJ, Shabalin IG, Handing KB, Minor W. CheckMyMetal: A macromolecular metal-binding validation tool. Acta Cryst D. 2017;73:223–233.
27. Yamaguchi A, Tanaka S, Yamaguchi S, et al. Two novel heat-soluble protein families abundantly expressed in an anhydrobiotic tardigrade. PLoS ONE. 2012;7:e44209.
28. Fukuda Y, Miura Y, Mizohata E, Inoue T. Structural insights into a secretory abundant heat-soluble protein from an anhydrobiotic tardigrade, *Ramazzottius varieornatus*. FEBS Lett. 2016;9:2458–2469.
29. Fukuda Y, Inoue T. Crystal structure of secretory abundant heat soluble protein 4 from one of the toughest "water bears" micro-animals *Ramazzottius varieornatus*. Protein Sci. 2018;27: 993–999.
30. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.
31. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014; 42:W320–W324.
32. Kabsch W. XDS. XDS Acta Cryst D. 2010;66:125–132.
33. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Cryst D. 2013;69:1204–1214.
34. Skubak P, Pannu NS. Automatic protein structure solution from weak X-ray data. Nat Commun. 2013;4:2777.
35. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of coot. Acta Cryst D. 2010;66:486–501.
36. Terwilliger TC, Zwart PH. PHENIX: A comprehensive python-based system for macromolecular structure solution. Acta Cryst D. 2010;66:213–221.
37. Chen VB, Arendall WB 3rd, Headd JJ, et al. MolProbity: All-atom structure validation for macromolecular crystallography. Acta Cryst D. 2010;66:12–21.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Fukuda Y, Inoue T. Structural insights into a C2 domain protein specifically found in tardigrades. *Protein Science*. 2021;30:513–518. https://doi.org/10.1002/pro.4002