STRUCTURAL BIOLOGY COMMUNICATIONS

Volume 78 (2022)

Supporting information for article:

Structure of the hypothetical protein TTHA1873 from *Thermus thermophilus*

I. Yuvaraj, Santosh Kumar Chaudhary, J. Jeyakanthan and K. Sekar

S1. Protein expression and purification

The TTHA1873 was expressed and purified [Supplementary Figs. 1A & 1B] and purity was checked using SDS PAGE [Supplementary Fig. 1C]. The LC/ESI-MS of the purified protein showed the 1628.79 Da mass differences between the observed and expected mass. The observed mass of the protein matched with the expected mass equivalent to the N-terminal cleavage of the first fifteen residues (MGNYLEDCATVDVQA), which might be a signal sequence. However, signal prediction servers could not predict any signal sequence from the full-length protein sequence. The molecular weight of the protein after cleavage is 16.85 kDa, which was confirmed using mass spectrometry [Supplementary Fig. 1D].

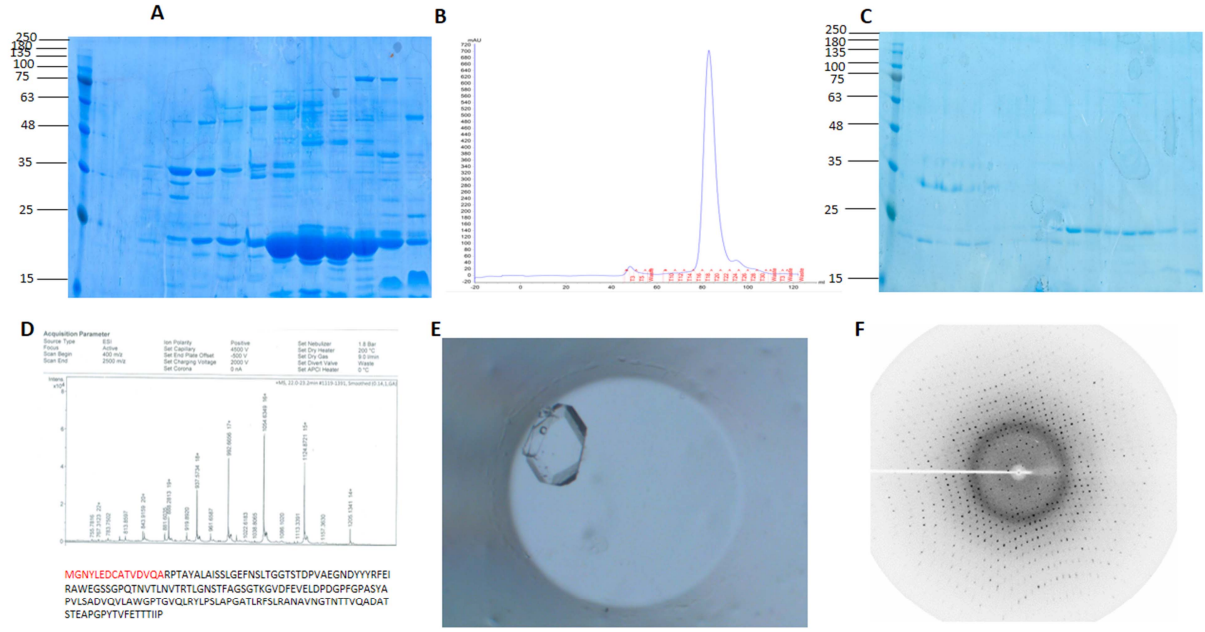

Figure S1 Purification of hypothetical protein TTHA1873. (A) 12% SDS-PAGE showing the expressed protein (B) Gel filtration chromatography for protein purification step shows the single peak (C) 12% SDS-PAGE showing the purified protein after gel filtration chromatography (D) LC-ESI/MS of the purified TTHA1873 protein confirms the N-terminal cleavage of 15 residues and molecular weight of the protein is 16.85 kDa and 15 residues of the TTHA1873 protein sequence, marked in red were missing in the N-terminal. (E) Crystals obtained for the TTHA1873, crystallization condition is 0.1 M Tris pH 8.5, 3.0 M NaCl and protein concentration is 45 mg/ml of protein (F) Diffraction pattern of TTHA1873 crystal

S2. Protein Similarity Search

Protein Similarity Search (PSI-Search) to find distantly related protein sequences revealed the protein is similar to the N-terminal domain (1-183) of capsid polyprotein of Human astrovirus. The N-terminal region of astroviruses capsid proteins contain a region of basic amino acids (Arginine and Lysine) that are predicted to interact with the genomic RNA inside the virus (Arias & Dubois, 2017), (Geigenmüller *et al.*, 2002). The presence of viral capsid (N-terminal domain) like protein in bacterium *T. Thermophilus* is interesting [Supplementary Fig. 2]. DALI server was used for structural similarity search [Supplementary Fig. 3]

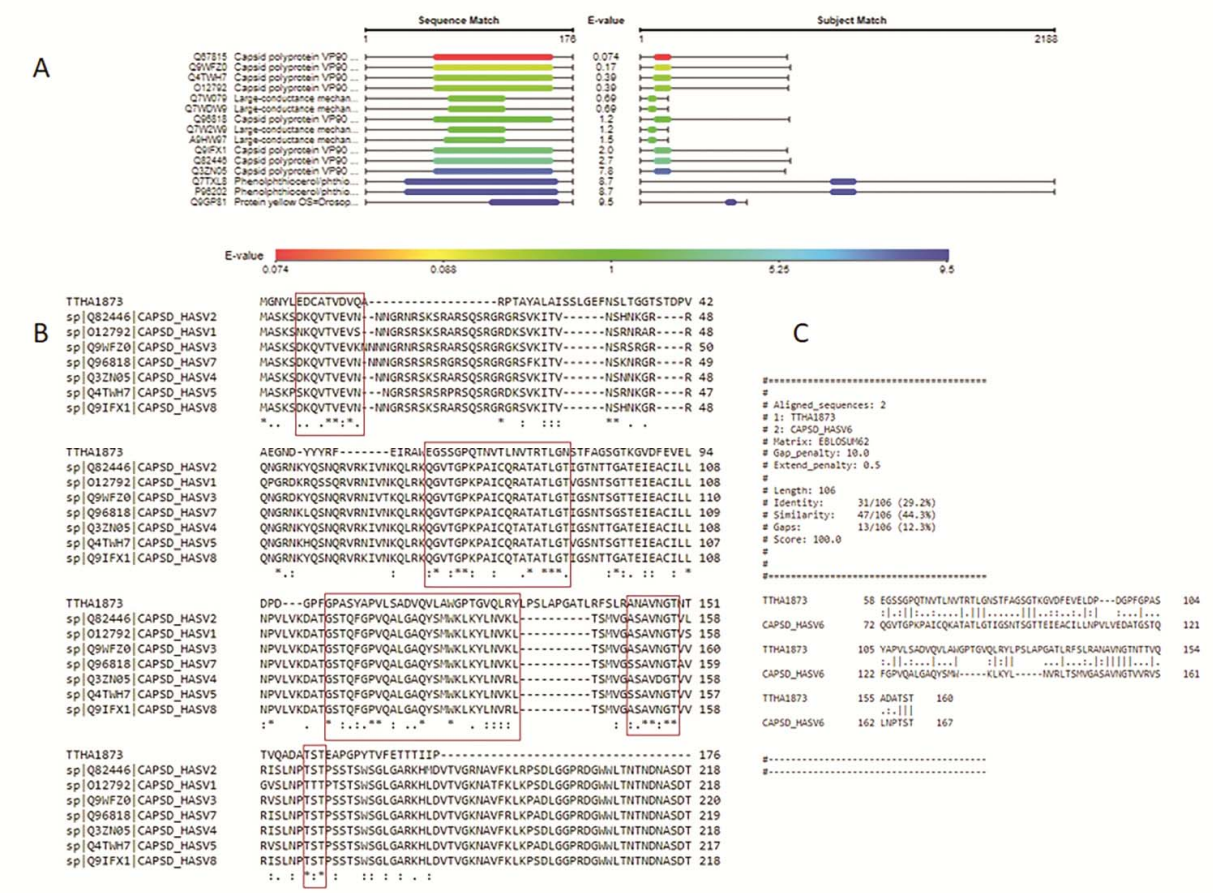

Figure S2 Sequence alignment of TTHA1873 with N-terminal domain of capsid polyprotein VP90 of Human astrovirus. (A) Protein Similarity Search (PSI-Search) (Madeira *et al.*, 2022) to find distantly related protein sequences showing the matches with capsid polyprotein VP90 of Human astrovirus with the E-value of 0.074. (B) Multiple Sequence alignment of TTHA1873 with N-terminal domain (1-183) of capsid polyprotein of Human astroviruses. The red boxed residues highlights the domains with similar protein sequences. (C) Sequence alignment between TTHA1873 (58-160) and Q67815 (72-167), N-terminal domain of capsid polyprotein of Human astrovirus 6.

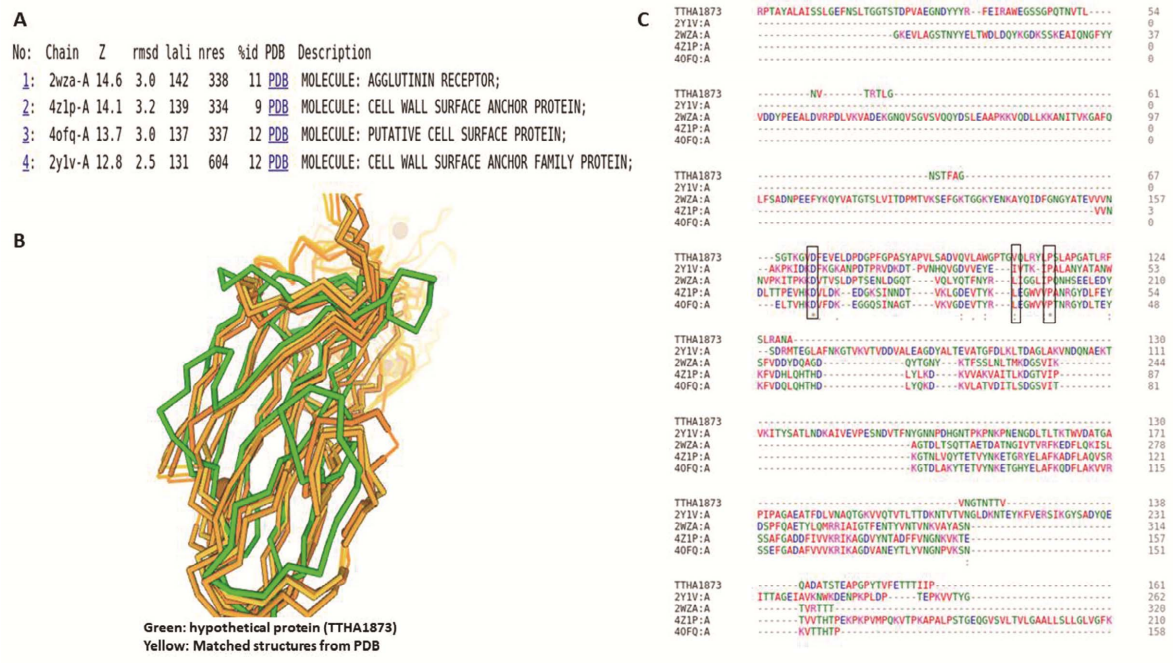

Figure S3 Structural alignment of TTHA1873 using DALI server (A) Similar structures identified by DALI (B) Similar structures from PDB were superimposed using DALI server, TTHA1873 structure is shown in Green and matched structures from PDB are in yellow. (C) Multiple sequence alignment of the proteins with top hits from DALI server.

S3. Structural disparity analysis upon mercury II potassium iodide binding

The structural alignment of the native structure with the mercury II potassium iodide bound structure was performed to analyze the structural changes upon mercury binding. The structures have a root mean square deviation of 0.126. The major divergence occurred in the region of residues (107-116) [Fig. 4A]. The mercury II potassium iodide bound in the loop between β4 and β5, this resulted in loop stabilization and, β5 and β6 strand extension compared to the native structure. Ser110, Ala111, and Asp112 form a 3_{10} helix in the mercury II bound structure, whereas in the native structure, it forms a loop [Supplementary Fig. S4].

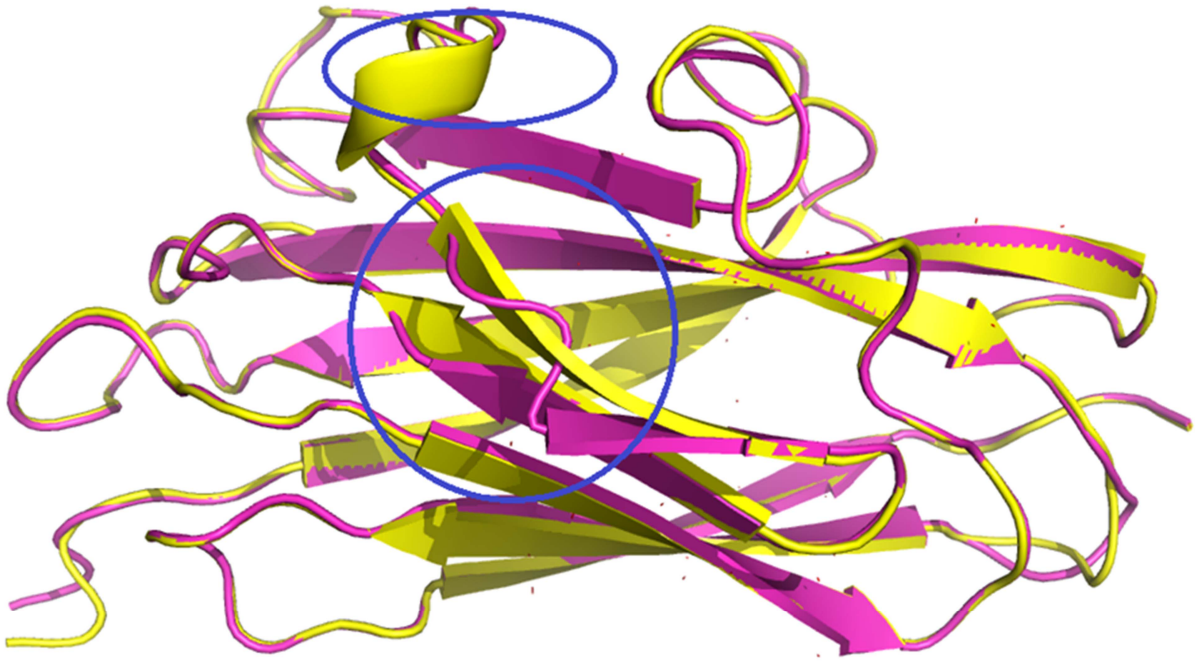

Figure S4 : Comparison of mercury bound and native TTHA1873 structures. (A) The native monomer structure (Pink) with the mercury II bound structure (Yellow), Blue circled are regions showing the deviations occurred upon mercury II potassium Iodide binding.

Table S1 Calcium Ion interactions in the structure of TTHA1873

	Bond Length(Å)	Residue
CALCIUM-I	**2.4**	**Ile24(O)**
	2.4	**Ser25(OG)**
	2.4	**Glu29 (OE1)**
	2.5	**Glu29 (OE2)**
	2.3	**Asp40(OD2)**
	2.3	**Glu170 (OE2)**
	2.3	**Thr171(O)**
CALCIUM-II	**2.5**	**Asp95(OD1)**
	2.4	**Asp95(OD2)**
	2.3	**Asp97(OD1)**
	2.4	**Gly101(O)**
	2.4	**Ala103(O)**

O, oxygen; OD1, oxygen delta 1; OD2, oxygen delta 2; OG, oxygen gamma; OE1, oxygen epsilon 1; OE2, oxygen epsilon 2.

Table S2 Ion pairs present in the hypothetical protein TTHA1873 structure.

	Bond Length(Å)	Residue
Asp47(OD1)	**3.23**	**Arg142(NH1)**
Asp47(OD2)	**2.81**	**Arg142(NH1)**
Asp89(OD1)	**2.78**	**Arg142(NH2)**
Glu53(OE1)	**2.5**	**Arg55(NH1)**
Glu93(OE1)	**3.4**	**Arg138(NH2)**
Glu93(OE2)	**3.22**	**Arg138(NH2)**
Glu91(OE2)	**3.61**	**Lys86(NZ)**

NZ, nitrogen zeta; NH1, nitrogen eta 1; NH2, nitrogen eta 2; OD1, oxygen delta 1; OD2, oxygen delta 2; OE1, oxygen epsilon 1; OE2, oxygen epsilon 2.

References

Arias, C. F. & Dubois, R. M. (2017). *Viruses* **9**.
Geigenmüller, U., Ginzton, N. H. & Matsui, S. M. (2002). *The Journal of general virology* **83**, 1691-1695.
Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A. & Lopez, R., (2022). Nucleic Acids Research **50,** W276–W279.