



RESEARCH ARTICLE

# Non-native fold of the putative VPS39 zinc finger domain

[version 1; peer review: 2 approved]

Benjamin G. Butt , Edward J. Scourfield, Stephen C. Graham 

Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, UK

**v1** First published: 01 Jul 2020, 5:154  
<https://doi.org/10.12688/wellcomeopenres.16078.1>  
 Latest published: 01 Jul 2020, 5:154  
<https://doi.org/10.12688/wellcomeopenres.16078.1>

## Abstract

**Background:** The multi-subunit homotypic fusion and vacuole protein sorting (HOPS) membrane-tethering complex is involved in regulating the fusion of late endosomes and autophagosomes with lysosomes in eukaryotes. The C-terminal regions of several HOPS components have been shown to be required for correct complex assembly, including the C-terminal really interesting new gene (RING) zinc finger domains of HOPS components VPS18 and VPS41. We sought to structurally characterise the putative C-terminal zinc finger domain of VPS39, which we hypothesised may be important for binding of VPS39 to cellular partners or to other HOPS components.

**Methods:** We recombinantly expressed, purified and solved the crystal structure of the proposed zinc-binding region of VPS39.

**Results:** In the structure, this region forms an anti-parallel  $\beta$ -hairpin that is incorporated into a homotetrameric eight-stranded  $\beta$ -barrel. However, the fold is stabilised by coordination of zinc ions by residues from the purification tag and an intramolecular disulphide bond between two predicted zinc ligands.

**Conclusions:** We solved the structure of the VPS39 C-terminal domain adopting a non-native fold. Our work highlights the risk of non-native folds when purifying small zinc-containing domains with hexahistidine tags. However, the non-native structure we observe may have implications for rational protein design.


## Keywords

CORVET, membrane trafficking, zinc finger domain, class C core

## Open Peer Review

Reviewer Status  

	Invited Reviewers	
	1	2
<b>version 1</b> 01 Jul 2020	 report	 report

1 **Daniel Kümmel** , University of Münster, Münster, Germany

2 **Lauren Jackson** , Vanderbilt University, Nashville, USA

**Boyang Xie** , Vanderbilt University, Nashville, USA

Any reports and responses or comments on the article can be found at the end of the article.

**Corresponding author:** Stephen C. Graham ([scg34@cam.ac.uk](mailto:scg34@cam.ac.uk))

**Author roles:** **Butt BG:** Data Curation, Investigation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Scourfield EJ:** Data Curation, Investigation, Resources, Writing – Review & Editing; **Graham SC:** Conceptualization, Data Curation, Funding Acquisition, Investigation, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** This work was supported by a Sir Henry Dale Fellowship (098406), jointly funded by the Wellcome Trust and the Royal Society (to SCG). BGB is a Wellcome Trust PhD student. Remote synchrotron access was supported in part by the EU FP7 infrastructure grant BIOSTRUCT-X (Contract No. 283570). A Titan V graphics card used for this research was donated by the NVIDIA Corporation.  
*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2020 Butt BG *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Butt BG, Scourfield EJ and Graham SC. **Non-native fold of the putative VPS39 zinc finger domain [version 1; peer review: 2 approved]** Wellcome Open Research 2020, 5:154 <https://doi.org/10.12688/wellcomeopenres.16078.1>

**First published:** 01 Jul 2020, 5:154 <https://doi.org/10.12688/wellcomeopenres.16078.1>

## Introduction

Eukaryotic cells use an interconnected system of membrane-bound compartments to partition intracellular space, allowing a multitude of biological reactions to proceed simultaneously in distinct chemical environments. The primary carriers of macromolecules between these compartments are vesicles, which bud from donor membranes in a cargo-dependent manner before fusing with an acceptor membrane at the destination compartment. Membrane fusion in the endomembrane system is critically dependent on SNARE (soluble *N*-ethylmaleimide sensitive factor attachment protein receptor) proteins, the co-folding of which on opposing membranes provides the energy for membrane bilayer mixing and thus vesicle fusion<sup>1</sup>. SNARE activity is tightly regulated by both Sec1/Munc18 family proteins, which bind directly to SNAREs, and by multi-protein ‘tethering’ complexes that bring vesicles into close apposition to allow the physical contact of SNARE proteins on opposing membranes<sup>2</sup>. The conserved multi-subunit tethering complexes CORVET (class C core vacuole/endosome tethering) and HOPS (homotypic fusion and vacuole protein sorting) combine both of these activities by incorporating the Sec1/Munc18 family protein VPS33A<sup>3–5</sup>. CORVET mediates homotypic fusion of early endosomes<sup>6</sup>, while HOPS mediates heterotypic fusion of late endosomes with lysosomes<sup>4,5</sup> and autophagosomes with lysosomes<sup>7–9</sup>.

The human CORVET and HOPS complexes share four conserved core subunits (VPS11, VPS16, VPS18, VPS33A), known collectively as the class C core<sup>3,10</sup>. Two additional, unique subunits direct each complex to its respective membrane target; VPS8 and TRAP1 direct CORVET to Rab5-positive membranes<sup>6,11</sup>, while VPS41 and VPS39 direct HOPS to Rab7-positive membranes<sup>12,13</sup>. Previous studies using truncation mapping have highlighted the importance of the C-terminal regions of HOPS components in assembly of the HOPS complex<sup>3,14–16</sup>. Recruitment of VPS41 to the class C core is facilitated by the C-terminal RING (really interesting new gene) domains of VPS18 and VPS41, which interact directly<sup>15</sup>. RING domains are a type of zinc finger, with an eight-residue motif containing six or seven cysteine residues and one or two histidine residues that coordinate two zinc ions<sup>17–19</sup>. RING domains may be involved in protein-protein, protein-lipid or protein-nucleic acid interactions, and have a wide variety of cellular functions<sup>17–19</sup>.

The C terminus of VPS39 contains a putative zinc finger domain<sup>15</sup> (Figure 1A), the closest homologue of which is the zinc finger domain of *Saccharomyces cerevisiae* protein Pcf11 (Figure 1B)<sup>20</sup>. This putative VPS39 zinc finger domain is much shorter than those of VPS18 and VPS41, and is predicted to bind only one zinc ion via four ligands<sup>15</sup>. Given that VPS41 is recruited to the class C core by an interaction between two zinc finger domains<sup>14</sup>, and that the C-terminal region of VPS39 is required for its interaction with VPS11<sup>16</sup>, we hypothesised that the putative VPS39 C-terminal zinc finger domain may be required for its incorporation into the HOPS complex or for binding other cellular partners.

There is currently no high-resolution structural information available for any region of human VPS39, nor its yeast

homologue vps39 (a.k.a. vam6). An atomic-resolution structure of the putative VPS39 zinc finger domain may further our understanding of HOPS complex assembly and function. We solved the structure of crystals formed by the VPS39 zinc finger domain to 2.9 Å resolution, but observed that the protein had adopted a non-native fold mediated by interactions between zinc ions and the purification tag.

## Methods

### Protein expression and purification

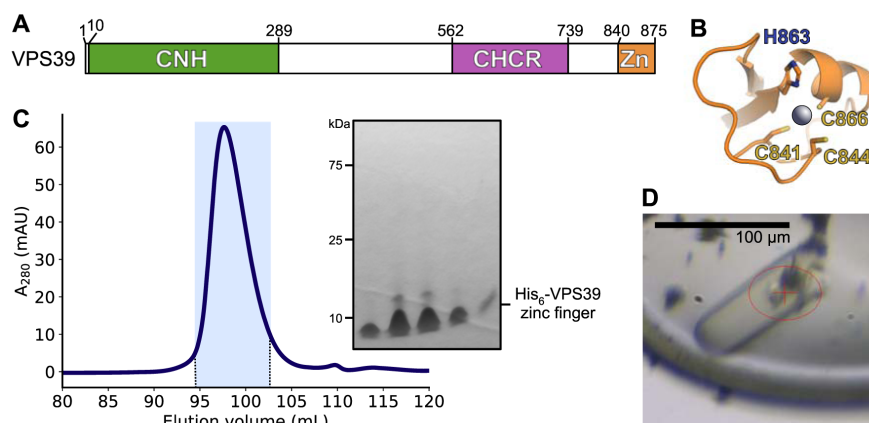
Residues 840–875 of human VPS39 isoform 2 (UniProt ID Q96JC1-2), corresponding to the putative C-terminal zinc finger domain, were cloned into pOPH, (derived from pOPT<sup>21</sup>), with an N-terminal MetHis<sub>6</sub> purification tag and expressed in *Escherichia coli* strain BL21(DE3) pLysS. Bacteria were cultured in 2×TY medium, recombinant proteins being expressed overnight at 22°C following addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside. Cultures were harvested by centrifugation at 5000×g for 15 min and cell pellets were stored at -80°C.

Bacterial cell pellets were resuspended in lysis buffer (20 mM tris pH 7.5, 500 mM NaCl, 20 mM imidazole pH 7.5, 0.5 mM MgCl<sub>2</sub>, 1.4 mM β-mercaptoethanol, 0.05% Tween-20) supplemented with 400 U bovine pancreas DNase I (Merck) and 200 μL EDTA-free protease inhibitor cocktail (Merck) at 4°C. Cells were lysed using a TS series cell disruptor (Constant Systems) at 24 kPSI and the lysate was cleared by centrifugation at 40,000×g for 30 min at 4°C. The cleared lysate was incubated with Ni<sup>2+</sup>-nitrilotriacetic acid agarose resin (Qiagen) equilibrated in wash buffer (20 mM tris pH 7.5, 500 mM NaCl, 20 mM imidazole pH 7.5) for 1 h at 4°C before being applied to a column and washed with >10 column volumes of wash buffer. Bound protein was eluted using elution buffer (20 mM tris pH 7.5, 500 mM NaCl, 250 mM imidazole pH 7.5), concentrated, and further purified by size-exclusion chromatography (SEC) using an S75 16/600 column (GE Healthcare) equilibrated in SEC buffer (20 mM tris pH 7.5, 200 mM NaCl, 1 mM dithiothreitol (DTT)). Purified VPS39 was concentrated, snap-frozen in liquid nitrogen as small aliquots (<100 μL), and stored at -80°C. Protein concentrations were estimated from absorbance at 280 nm using a calculated extinction coefficient<sup>22</sup> for VPS39(840–875), assuming all cysteines were reduced.

### X-ray crystallography

VPS39(840–875) was crystallised in sitting drops by mixing 200 nL of 19.4 mg/mL protein in SEC buffer with 200 nL of reservoir solution (100 mM HEPES pH 7.5, 200 mM ammonium acetate, 45% (v/v) 2-methyl-2,4-pentanediol (MPD)) and equilibrating against 80 μL of reservoir at 20°C for 30 months. The VPS39 crystal was cryo-cooled by plunging into liquid nitrogen, no cryoprotectant being added as the high concentration of MPD in the reservoir solution was predicted to provide sufficient cryoprotection. Diffraction data were recorded at 100 K on a Pilatus 6M-F detector (Dectris) at Diamond Light Source beamline I04. Data were collected in three sweeps, as shown in Table 1.

Images were processed using DIALS version 1.14.13<sup>23</sup> then CCP4 suite version 7.0.078<sup>24</sup> programs POINTLESS version 1.11.21<sup>25</sup> and AIMLESS version 0.7.4<sup>26</sup> as implemented by the



**Figure 1. Purification and crystallisation of the VPS39 zinc finger domain.** (A) Predicted domain organisation of human VPS39, showing the position of the short C-terminal zinc finger domain that is the focus of this study (CNH, citron homology domain; CHCR, clathrin heavy chain repeat; Zn, zinc finger domain). Predicted domain boundary residue numbers are indicated above the schematic. (B) Homology model of VPS39 C-terminal domain based on *S. cerevisiae* Pcf11 zinc finger domain. Putative zinc ligands are shown as sticks and the approximate position of a bound zinc ion is indicated by a grey sphere. (C) SEC elution profile of purified VPS39 C-terminal domain (UV absorbance: solid blue line). Fractions that were analysed by SDS-PAGE (inset) are highlighted in light blue. (Inset). SDS-PAGE analysis of SEC elution fractions 35–39 (94.1–102.1 mL). Approximate positions of molecular weight markers are shown. (D) Crystal of purified VPS39 C-terminal domain mounted on beamline I04 at the Diamond Light Source (scale bar = 100 µm; position of X-ray beam: red crosshair).

**Table 1. Data collection strategy.** Data were recorded from a single crystal in the order Peak 1, Peak 2 and then High-energy remote.

Dataset	Peak 1	Peak 2	High-energy remote
Wavelength (Å)	1.2810	1.2810	0.9795
Exposure (s)	0.5	0.2	0.2
X-ray transmission (%)	3.0	29.9	52.4
Oscillation per frame (°)	0.2	0.2	0.2
Total number of frames	900	900	1800

xia2 version 0.5.902 data processing pipeline<sup>27</sup>. Data collection statistics are shown in Table 2. Two-wavelength multiple anomalous dispersion analysis was performed using the CCP4 suite version 7.1.001<sup>24</sup> CRANK2 version 2.0.229 automated experimental phasing pipeline<sup>28</sup>, with substructure determination performed with SHELXD version 2019/1<sup>29</sup>, density modification performed with Parrot version 0.8<sup>30</sup>, and iterative model building and refinement performed with Buccaneer version 1.1<sup>31,32</sup> and Refmac5 version 5.8.0258<sup>33</sup>. Cycles of iterative manual building with COOT version 0.8.9<sup>34</sup> and TLS plus positional refinement using Refmac5 version 5.8.0258<sup>33</sup> with local non-crystallographic symmetry (NCS) restraints were initially performed using the high-energy remote wavelength dataset (Table 2). Building was assisted by the use of real-time molecular dynamics-assisted model building and map fitting with ISOLDE version 1.0b3<sup>35</sup>. To ameliorate radiation damage evident in the structure, later stages of refinement were performed using the first 300 frames of the second peak wavelength dataset (Peak 2; Table 1), processed using xia2 as above with the same set of reflections kept ‘free’ for cross-validation<sup>36</sup>. Final

cycles of refinement were performed using autoBUSTER version 2.10.3<sup>37</sup> with local NCS restraints and bond length/angle restraints for zinc ligands to ensure chemically-plausible zinc coordination<sup>38</sup>. The quality of the model was monitored throughout refinement using MolProbity version 4.5.1<sup>39</sup> and the validation tools in COOT version 0.8.9<sup>34</sup>. Refinement statistics are shown in Table 2. Molecular images were produced in PyMOL 2.4.0a0 Open-Source<sup>40</sup> and figures were composed in Inkscape version 1.0<sup>41</sup>. VPS39 C-terminal domain residues predicted to bind zinc were identified via generation of a homology model using I-TASSER version 5.1<sup>42</sup> with the structure of *S. cerevisiae* Pcf11 (PDB ID: 2NAX)<sup>20</sup> as the template.

## Results

The C-terminal region of human VPS39 contains a putative zinc finger domain (residues 840–875, Figure 1A) with four predicted zinc-binding residues (Cys841, Cys844, His863, Cys866). These residues are predicted to coordinate a single zinc ion based on homology to the zinc finger domain of *S. cerevisiae* protein Pcf11 (Figure 1B). The coordinates for this theoretical model are available (see *Underlying data*)<sup>43</sup>.

The VPS39 C-terminal domain was expressed with an N-terminal His<sub>6</sub> tag in *E. coli* and purified using nickel affinity capture followed by SEC. The protein eluted from SEC as a single, symmetrical peak near the end of the elution profile (Figure 1C), consistent with expectations for a small folded protein domain. Analysis of the eluted fractions by SDS-PAGE showed a single predominant band that migrated as would be expected for the VPS39 zinc finger domain (5.1 kDa; Figure 1C), with a much less intense band at higher apparent molecular mass that was presumed to be a small amount of SDS-resistant VPS39 dimer.

**Table 2. Data collection and refinement statistics.** The 'Peak' column describes the merged diffraction data from sweeps 'Peak 1' and 'Peak 2' (Table 1) used for structure solution. The 'Peak(1–300)' column describes the subset of 'Peak 2' diffraction data used for structure refinement. Values in parentheses describe the high-resolution shell.

Dataset	Peak	High energy remote	Peak(1–300)
<b>Data collection</b>			
Wavelength (Å)	1.28096	0.97949	1.28096
Space group	$P4_222$	$P4_222$	$P4_222$
Cell dimensions			
$a, b, c$ (Å)	104.17, 104.17, 39.43	104.17, 104.17, 39.43	104.18, 104.18, 39.42
$\alpha, \beta, \gamma$ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	73.65–3.07 (3.12–3.07)	28.89–2.98 (3.03–2.98)	46.59–2.90 (2.95–2.90)
Total reflections	97,700 (2367)	116,580 (4961)	21,099 (1058)
Unique reflections	4399 (197)	4793 (247)	5148 (247)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	99.3 (100.0)
Anomalous completeness (%)	100.0 (100.0)	99.9 (100.0)	96.8 (99.0)
Multiplicity	22.2 (12.0)	24.3 (25.7)	4.1 (4.3)
Anomalous multiplicity	12.6 (6.5)	13.7 (14.1)	2.3 (2.3)
$R_{merge}$	0.121 (0.844)	0.124 (1.154)	0.075 (1.178)
$R_{pim}$	0.027 (0.252)	0.026 (0.231)	0.041 (0.632)
$CC_{1/2}$	0.998 (0.939)	0.999 (0.944)	0.966 (0.602)
$CC_{anom}$	0.699 (0.011)	0.403 (0.020)	0.580 (-0.144)
Mean $I/\sigma(I)$	16.4 (2.4)	16.0 (3.1)	10.1 (0.9)
<b>Refinement</b>			
Resolution (Å)			46.59–2.90 (2.98–2.90)
Reflections			
Working set			4853 (350)
Test set			286 (18)
$R_{work}$			0.2376 (0.2535)
$R_{free}$			0.2686 (0.3028)
No. of atoms			
Protein			922
Solvent			1
Zinc ions			3
Root mean square deviation			
Bond lengths (Å)			0.008
Bond angles (°)			1.07
MolProbity score			2.05
Ramachandran favoured (%)			92.73
Ramachandran outliers (%)			0.00
Poor rotamers (%)			4.90
Mean B value (Å <sup>2</sup> )			122.04



The protein was concentrated and sparse matrix crystallisation screening was performed, but no crystals were obtained in the following six weeks. Approximately 30 months later, the crystallisation trays were re-inspected prior to disposal and a single crystal was observed (Figure 1D). This crystal was harvested and diffraction data were recorded at two wavelengths (Table 2), allowing the structure of the VPS39 zinc finger domain to be solved using anomalous dispersion signal from the incorporated zinc ions. The model was initially refined against the high-energy data, but later stages of the refinement proved challenging because map features were indistinct and loop density was poor. We were concerned that intense X-ray exposure during data collection at the peak wavelength, where the zinc ions would have a large X-ray absorption cross-section<sup>44</sup>, may have caused radiation damage. The final stages of refinement were thus performed using data recorded in the first 300 frames of the second sweep at the peak wavelength (Table 1 and Table 2), which represented the best compromise between total X-ray exposure/damage and data redundancy/resolution. The structure was refined to 2.90 Å resolution with residuals  $R = 0.238$ ,  $R_{\text{free}} = 0.269$  and good stereochemistry, with an overall MolProbity score<sup>39</sup> of 2.05 (Table 2). The structure is available under PDB ID: 6ZE9; raw diffraction images, crystallographic datasets and X-ray fluorescence scans are available (see *Underlying data*)<sup>45</sup>.

The asymmetric unit contains three copies of the VPS39 C-terminal domain: two full-length copies (residues 840–875; purple and teal in Figure 2A) and a third copy spanning residues 840–869 (blue in Figure 2A). The remaining C-terminal residues of the third copy are absent from the electron density and presumably disordered. Each copy of the VPS39 C-terminal domain forms an antiparallel  $\beta$ -hairpin, with residues 849–860 forming a loop linking the two  $\beta$ -strands (Figure 2A). Strikingly, the VPS39 C-terminal domains are all organised around crystallographic symmetry axes such that they form eight-stranded  $\beta$ -barrels (Figure 2B). There are two distinct homotetramers formed: the first comprises two NCS-related chains that interact with two additional chains that are related by crystallographic two-fold rotational symmetry (Figure 2C), while the second homotetramer is formed by a single VPS39 C-terminal domain interacting with three additional chains that are related by two orthogonal two-fold crystallographic symmetry axes (Figure 2D).

The asymmetric unit contains three zinc ions, consistent with the four predicted zinc ligands in each VPS39 copy based on homology to Pcf11 (Figure 1B). All zinc ions have tetrahedral geometry. However, only one of the predicted zinc ligands (Cys844) is involved in zinc ion coordination (Figure 2E). Of the remaining predicted zinc ligands, Cys841 and Cys866 had become oxidised to form an intramolecular disulphide bond in each VPS39 molecule (Figure 2F) and the final predicted ligand (His863) is not in close proximity to the zinc ions. Instead, the remaining zinc ligands are provided by two histidine side chains from the MetHis<sub>6</sub> purification tag (His-3 and His-1) and the terminal carboxylate group of the polypeptide chain (Thr875) or a water molecule (Figure 2G). As

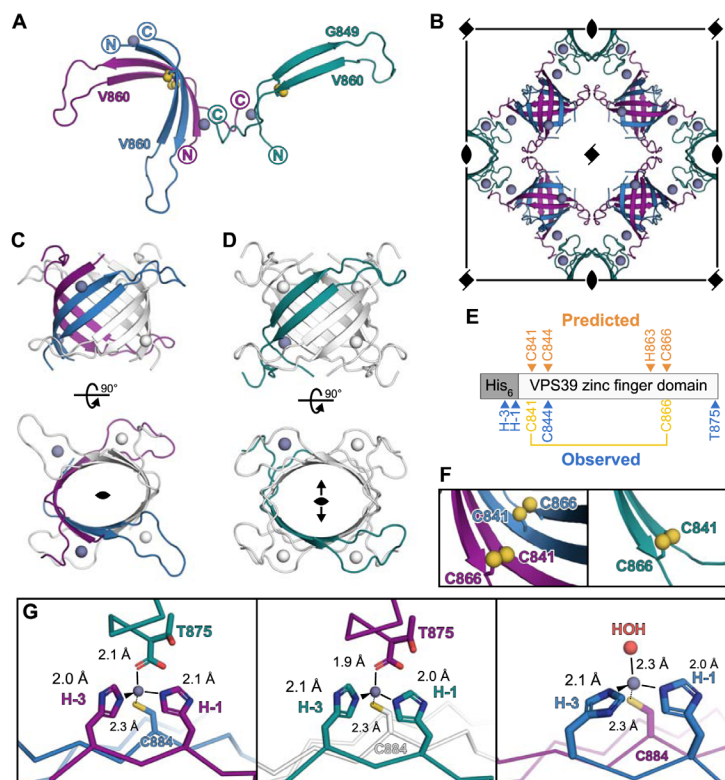
two of the ligands for each zinc ion derive from the affinity purification tag and the fold of the VPS39 C-terminal domain that we observe differs significantly from that of the closest sequence homologue (compare Figure 1B and Figure 2A), we conclude that the observed fold is non-native.

## Discussion

We present the crystal structure of the human VPS39 zinc finger domain in a non-native fold. In the structure, three copies of the VPS39 C-terminal domain in the asymmetric unit (Figure 2A) combine with symmetry-related chains to form two similar, homotetrameric, eight-stranded  $\beta$ -barrels (Figure 2C, D). In each copy of VPS39, two of the residues predicted to bind zinc ions (Cys844 and Cys866; Figure 2E) instead form intramolecular disulphide bonds (Figure 2F), with the remaining zinc ligands provided by side chains from the N-terminal His<sub>6</sub> purification tag and the carboxylate group of the polypeptide chain or a water molecule (Figure 2G).

Structural characterisation of VPS39 was undertaken to complement a yeast two-hybrid screen of HOPS component zinc finger domains, including the putative VPS39 zinc finger domain, with the aim of identifying cellular binding proteins<sup>15</sup>. However, as pull-down experiments failed to validate any of the potential interactions that were tested, structural characterisation of the VPS39 C-terminal domain was not actively pursued. After 30 months, as the crystallisation trials were being discarded, a single VPS39 C-terminal domain crystal was identified and used for successful structure determination. It seems very likely that the non-native fold that we observed arose from re-folding of the purified VPS39 C-terminal domain during the extended crystallisation experiment. The elution of freshly purified VPS39 C-terminal domain from SEC (Figure 1C) was consistent with this small protein being monomeric, whereas the  $\beta$ -barrels of VPS39 in the crystal structure would be likely to elute much earlier, although we concede that formation of a  $\beta$ -barrel fold from the outset remains possible.

Refolding of the VPS39 C-terminal domain to form the observed  $\beta$ -barrels is likely to have been promoted via the concerted actions of zinc binding by the purification tag, disulphide bond formation and formation of  $\beta$ -sheets with unsatisfied backbone hydrogen bonds. The histidine side chains from the MetHis<sub>6</sub> purification tag could have competed with Cys841 and Cys866 for coordination of the zinc ions, thereby liberating the side chains of these two cysteine residues. While the VPS39 C-terminal domain was purified under reducing conditions (the SEC buffer being supplemented with 1 mM DTT), it is likely that the contents of the crystallisation drops became oxidised during their extended incubation. The liberated cysteine side chains may thus have formed the observed intramolecular disulphide bond, prohibiting them from competing with the MetHis<sub>6</sub> tag side chains for re-binding to the zinc ion. Either or both molecular rearrangements could have promoted re-folding of the protein backbone to adopt the extended  $\beta$ -hairpin fold observed in this structure. The refolded VPS39  $\beta$ -sheets would have unsatisfied backbone hydrogen bonds, which could have promoted similar refolding of additional VPS39 molecules (akin



**Figure 2. Non-native structure of VPS39 C-terminal domain.** (A) Three copies of the VPS39 C-terminal domain in the asymmetric unit, showing the antiparallel  $\beta$ -hairpin fold of each molecule. VPS39 is shown in ribbon representation with N and C termini shown. Cysteine sulphur atoms that form disulphide bonds are shown as yellow spheres and zinc ions are shown as grey spheres. Residues at the start and end of the loop that joins the two  $\beta$ -strands are indicated. (B) Unit cell of the VPS39 crystal lattice viewed along the  $c$  axis, showing eight stranded  $\beta$ -barrels formed by symmetry-related VPS39 chains. Selected symmetry axes (four-fold screw and two-fold rotation) are indicated using standard symmetry symbols. Incorporated zinc ions are shown as spheres. (C, D) Eight-stranded  $\beta$ -barrels formed by symmetry-related chains via a single crystallographic two-fold rotational symmetry axis (C) or two orthogonal two-fold rotational axes (D) are shown in ribbon representation. The unique VPS39 molecules from the asymmetric unit (A) are coloured blue/purple (C) or teal (D), with symmetry-related chains shown in light grey. Zinc ions coordinated by visible residues are shown as spheres. Two orthogonal views are shown, with symmetry axes shown in the bottom view where arrows represent a two-fold rotational axis in the plane of the image. (E) Schematic of the His<sub>6</sub>-VPS39 zinc finger construct used in this study. (Top) Residues predicted to coordinate zinc ions by homology to *S. cerevisiae* Pcf11 (Figure 1B) are denoted with arrows. (Bottom) Residues that coordinate zinc ions in the crystal structure, including two residues from the purification tag, are denoted with arrows. Residues involved in the intramolecular disulphide bond in each chain are joined. (F) Enlarged views of intramolecular Cys841-Cys866 disulphide bond in each molecule of the VPS39 C-terminal domain. (G) Enlarged views of the three zinc ions in the asymmetric unit. VPS39 backbone atoms are shown as lines, with side chains involved in coordinating zinc ions shown as sticks. Bond lengths between zinc ions and relevant side chain atoms or water molecules are indicated.

to nucleation of amyloid fibrils). Such stimulated refolding could promote further exchange of zinc ligands and disulphide bond formation, acting as a ratchet to increase the pool of refolded VPS39 for crystallisation. The covalent interaction between  $\beta$ -barrels, mediated by the carboxy terminus of the polypeptide binding to the zinc ions, would have promoted stability of the crystal once nucleated.

While the structure presented here does not provide biological insight into the organisation or function of the putative VPS39 C-terminal zinc finger domain, there are still useful lessons to be learned. Firstly, nickel-affinity chromatography should be used with caution when purifying zinc-binding proteins as the similar chemical properties of zinc and nickel can lead to competition between purification tag residues and

native zinc ligands for zinc ions. If this purification strategy is used, constructs should be engineered to include a protease cleavage site that can be used to remove the purification tag before downstream applications, particularly those involving long incubations such as crystallisation. We have previously reported structures where purification tag residues give rise to folding artefacts<sup>46</sup> and where metal ions help mediate non-natural ‘swapped’  $\beta$ -strand topologies of crystallised molecules<sup>47</sup>. While His<sub>6</sub> tags are generally benign for crystallisation and may indeed be beneficial in some cases<sup>48</sup>, caution should be exercised when using them to purify small zinc-containing domains.

The non-native  $\beta$ -barrel fold of the VPS39 C-terminal domain we observe here highlights the power of metal ion coordination

to strongly promote the stable (re)folding of proteins<sup>49</sup>, especially given the simple sequence requirements for efficient zinc binding (cysteine and histidine side chains or carboxylate groups). As a result, it is not uncommon for such features to arise spontaneously<sup>50,51</sup>, as has been previously noted in studies on directed protein evolution. Small zinc finger domains are often highly thermostable and tolerant to sequence changes outside of the zinc ligands<sup>52</sup>, which has led to their use as scaffolds for modular protein design<sup>53–55</sup>. Novel, non-native, metal ion-coordinating folds such as the VPS39 fold reported in this work are potentially less likely to interact with off-target cellular components when used as biologics<sup>56</sup>. The non-native fold of the VPS39 C-terminal domain presented here therefore expands the number of protein scaffolds available for rational therapeutic design.

## Data availability

### Underlying data

Protein Data Bank: Non-native fold of the putative VPS39 zinc finger domain. Accession number 6ZE9; <https://identifiers.org/rcsb/pdb:6ZE9>.

Apollo: Crystallographic diffraction data for structure of the VPS39 C-terminal domain. <https://doi.org/10.17863/CAM.5386745>.

This project contains raw diffraction images, crystallographic datasets and X-ray fluorescence scans.

Apollo: Theoretical model of the VPS39 zinc finger domain. <https://doi.org/10.17863/CAM.5450343>.

This project contains atomic coordinates for the theoretical model of the VPS39 zinc finger domain shown in [Figure 1B](#).

Data hosted with Apollo are available under the terms of the [Creative Commons Attribution 4.0 International license \(CC-BY 4.0\)](#).

## Acknowledgements

We thank Janet Deane for helpful discussions and Diamond Light Source for access to beamline I04 under proposal MX11235.

## References

- Yoon TY, Munson M: **SNARE complex assembly and disassembly**. *Curr Biol*. 2018; **28**(8): R397–R401. [PubMed Abstract](#) | [Publisher Full Text](#)
- Baker RW, Hughson FM: **Chaperoning SNARE assembly and disassembly**. *Nat Rev Mol Cell Biol*. 2016; **17**(8): 465–79. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Van Der Kant R, Jonker CTH, Wijdeven RH, *et al.*: **Characterization of the mammalian CORVET and HOPS complexes and their modular restructuring for endosome specificity**. *J Biol Chem*. 2015; **290**(51): 30280–90. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Spang A: **Membrane Tethering Complexes in the Endosomal System**. *Front Cell Dev Biol*. 2016; **4**: 35. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Balderhaar HJK, Ungermann C: **CORVET and HOPS tethering complexes - coordinators of endosome and lysosome fusion**. *J Cell Sci*. 2013; **126**(Pt 6): 1307–16. [PubMed Abstract](#) | [Publisher Full Text](#)
- Perini ED, Schaefer R, Stöter M, *et al.*: **Mammalian CORVET Is Required for Fusion and Conversion of Distinct Early Endosome Subpopulations**. *Traffic*. 2014; **15**(12): 1366–89. [PubMed Abstract](#) | [Publisher Full Text](#)
- Wartosch L, Günesdogan U, Graham SC, *et al.*: **Recruitment of VPS33A to HOPS by VPS16 Is Required for Lysosome Fusion with Endosomes and Autophagosomes**. *Traffic*. 2015; **16**(7): 727–42. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Liang C, Lee JS, Inn KS, *et al.*: **Beclin1-binding UVRAG targets the class C Vps complex to coordinate autophagosome maturation and endocytic trafficking**. *Nat Cell Biol*. 2008; **10**(7): 776–87. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Jiang P, Nishimura T, Sakamaki Y, *et al.*: **The HOPS complex mediates autophagosome-lysosome fusion through interaction with syntaxin 17**. *Mol Biol Cell*. 2014; **25**(8): 1327–37. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Klinger CM, Klute MJ, Dacks JB: **Comparative Genomic Analysis of Multi-Subunit Tethering Complexes Demonstrates an Ancient Pan-Eukaryotic Complement and Sculpting in Apicomplexa**. *PLoS One*. 2013; **8**(9): e76278. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lachmann J, Glaubke E, Moore PS, *et al.*: **The Vps39-like TRAP1 is an effector of Rab5 and likely the missing Vps3 subunit of human CORVET**. *Cell Logist*. 2014; **4**(4): e970840. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Van Der Kant R, Fish A, Janssen L, *et al.*: **Late endosomal transport and tethering are coupled processes controlled by RILP and the cholesterol sensor ORP1L**. *J Cell Sci*. 2013; **126**(Pt 15): 3462–74. [PubMed Abstract](#) | [Publisher Full Text](#)
- Rink J, Ghigo E, Kalaidzidis Y, *et al.*: **Rab conversion as a mechanism of progression from early to late endosomes**. *Cell*. 2005; **122**(5): 735–49. [PubMed Abstract](#) | [Publisher Full Text](#)
- Guo Z, Johnston W, Kovtun O, *et al.*: **Subunit organisation of *in vitro* reconstituted HOPS and CORVET multisubunit membrane tethering complexes**. *PLoS One*. 2013; **8**(12): e81534. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hunter MR, Scourfield EJ, Emmott E, *et al.*: **VPS18 recruits VPS41 to the human HOPS complex via a RING-RING interaction**. *Biochem J*. 2017; **474**(21): 3615–26. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Graham SC, Wartosch L, Gray SR, *et al.*: **Structural basis of Vps33A recruitment to the human HOPS complex by Vps16**. *Proc Natl Acad Sci U S A*. 2013; **110**(33): 13345–50. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Saurin AJ, Borden KLB, Boddy MN, *et al.*: **Does this have a familiar RING?** *Trends Biochem Sci*. 1996; **21**(6): 208–14. [PubMed Abstract](#)
- Borden KLB, Freemont PS: **The RING finger domain: A recent example of a sequence-structure family**. *Curr Opin Struct Biol*. 1996; **6**(3): 395–401. [PubMed Abstract](#) | [Publisher Full Text](#)
- Joazeiro CAP, Weissman AM: **RING Finger Proteins: Mediators of Ubiquitin Ligase Activity**. *Cell*. 2000; **102**(5): 549–52. [PubMed Abstract](#) | [Publisher Full Text](#)
- Yang F, Hsu P, Lee SD, *et al.*: **The C terminus of Pcf11 forms a novel zinc-finger structure that plays an essential role in mRNA 3'-end processing**. *RNA*. 2017; **23**(1): 98–107. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Teo H, Perisic O, González B, *et al.*: **ESCRT-II, an endosome-associated complex required for protein sorting: Crystal structure and interactions with ESCRT-III and membranes**. *Dev Cell*. 2004; **7**(4): 559–69. [PubMed Abstract](#) | [Publisher Full Text](#)
- Gasteiger E, Hoogland C, Gattiker A, *et al.*: **Protein Identification and Analysis Tools on the ExPASy Server**. Walker JM editor. *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press; 2005; 571–607. [Publisher Full Text](#)
- Winter G, Waterman DG, Parkhurst JM, *et al.*: **DIALS: implementation and evaluation of a new integration package**. *Acta Crystallogr D Struct Biol*. 2018;



- 74(Pt 2): 85–97.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Winn MD, Ballard CC, Cowtan KD, *et al.*: **Overview of the CCP4 suite and current developments.** *Acta Crystallogr D Biol Crystallogr.* 2011; **67**(Pt 4): 235–42.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. Evans P: **Scaling and assessment of data quality.** *Acta Crystallogr D Biol Crystallogr.* 2006; **62**(Pt 1): 72–82.  
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Evans PR, Murshudov GN: **How good are my data and what is the resolution?** *Acta Crystallogr D Biol Crystallogr.* 2013; **69**(Pt 7): 1204–14.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Winter G: **xia2: an expert system for macromolecular crystallography data reduction.** *J Appl Crystallogr.* 2010; **43**(1): 186–90.  
[Publisher Full Text](#)
28. Skubák P, Pannu NS: **Automatic protein structure solution from weak X-ray data.** *Nat Commun.* 2013; **4**: 2777.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Sheldrick GM: **A short history of SHELX.** *Acta Crystallogr A.* 2008; **64**(Pt 1): 112–22.  
[PubMed Abstract](#) | [Publisher Full Text](#)
30. Cowtan K: **Recent developments in classical density modification.** *Acta Crystallogr D Biol Crystallogr.* 2010; **66**(Pt 4): 470–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Cowtan K: **Completion of autobuilt protein models using a database of protein fragments.** *Acta Crystallogr D Biol Crystallogr.* 2012; **68**(Pt 4): 328–35.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Cowtan K: **The Buccaneer software for automated model building. 1. Tracing protein chains.** *Acta Crystallogr D Biol Crystallogr.* 2006; **62**(Pt 9): 1002–11.  
[PubMed Abstract](#) | [Publisher Full Text](#)
33. Murshudov GN, Vagin AA, Dodson EJ: **Refinement of Macromolecular Structures by the Maximum-Likelihood Method.** *Acta Crystallogr D Biol Crystallogr.* 1997; **53**(Pt 3): 240–55.  
[PubMed Abstract](#) | [Publisher Full Text](#)
34. Emsley P, Lohkamp B, Scott WG, *et al.*: **Features and development of Coot.** *Acta Crystallogr D Biol Crystallogr.* 2010; **66**(Pt 4): 486–501.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. Croll TI: **ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps.** *Acta Crystallogr D Struct Biol.* 2018; **74**(Pt 6): 519–30.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Brünger AT: **Free R value: cross-validation in crystallography.** *Methods Enzymol.* 1997; **277**: 366–96.  
[PubMed Abstract](#) | [Publisher Full Text](#)
37. Bricogne G, Blanc E, Brandl M, *et al.*: **BUSTER version 2.10.3.** Cambridge, United Kingdom: Global Phasing Ltd; 2017.  
[Reference Source](#)
38. Touw WG, Van Beusekom B, Evers JMG, *et al.*: **Validation and correction of Zn–Cys, His complexes.** *Acta Crystallogr D Struct Biol.* 2016; **72**(Pt 10): 1110–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
39. Williams CJ, Headd JJ, Moriarty NW, *et al.*: **MolProbity: More and better reference data for improved all-atom structure validation.** *Protein Sci.* 2018; **27**(1): 293–315.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Shrodinger LLC: **PyMOL 2.4.0a0 Open-Source.** 2015.
41. Jeanmougin M: **Inkscape version 1.0.** 2020.  
[Reference Source](#)
42. Roy A, Kucukural A, Zhang Y: **I-TASSER: a unified platform for automated protein structure and function prediction.** *Nature Protocols.* 2010; **5**(4): 725–38.  
[Publisher Full Text](#)
43. Butt BG, Graham SC: **Apollo: Theoretical model of the VPS39 zinc finger domain.** [Dataset]. 2020.  
<http://www.doi.org/10.17863/CAM.54503>
44. Murray JW, Garman EF, Ravelli RBG: **X-ray absorption by macromolecular crystals: the effects of wavelength and crystal composition on absorbed dose.** *J Appl Cryst.* 2004; **37**(4): 513–22.  
[Publisher Full Text](#)
45. Butt B, Scourfield E, Graham S: **Crystallographic diffraction data for structure of the VPS39 C-terminal domain.** [Dataset]. 2020.  
<http://www.doi.org/10.17863/CAM.53867>
46. Butt BG, Owen DJ, Jeffries CM, *et al.*: **Insights into herpesvirus assembly from the structure of the pUL7:pUL51 complex.** *eLife.* 2020; **9**: e53789.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
47. Hatherley D, Graham SC, Turner J, *et al.*: **Paired Receptor Specificity Explained by Structures of Signal Regulatory Proteins Alone and Complexed with CD47.** *Molecular Cell.* 2008; **31**(2): 266–77.  
[PubMed Abstract](#) | [Publisher Full Text](#)
48. Carson M, Johnson DH, McDonald H, *et al.*: **His-tag impact on structure.** *Acta Crystallogr D Biol Crystallogr.* 2007; **63**(3): 295–301.  
[PubMed Abstract](#) | [Publisher Full Text](#)
49. Tang J, Kang SG, Saven JG, *et al.*: **Characterization of the Cofactor-Induced Folding Mechanism of a Zinc-Binding Peptide Using Computationally Designed Mutants.** *J Mol Biol.* 2009; **389**(1): 90–102.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
50. Lo Surdo P, Walsh MA, Sollazzo M: **A novel ADP- and zinc-binding fold from function-directed *in vitro* evolution.** *Nat Struct Mol Biol.* 2004; **11**(4): 382–3.  
[Publisher Full Text](#)
51. Keefe AD, Szostak JW: **Functional proteins from a random-sequence library.** *Nature.* 2001; **410**(6829): 715–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
52. Sharpe BK, Liew CK, Kwan AH, *et al.*: **Assessment of the robustness of a serendipitous zinc binding fold: Mutagenesis and protein grafting.** *Structure.* 2005; **13**(2): 257–66.  
[PubMed Abstract](#) | [Publisher Full Text](#)
53. Cho GS, Szostak JW: **Directed Evolution of ATP Binding Proteins from a Zinc Finger Domain by Using mRNA Display.** *Chem Biol.* 2006; **13**(2): 139–47.  
[PubMed Abstract](#) | [Publisher Full Text](#)
54. Greisman HA, Pabo CO: **A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites.** *Science.* 1997; **275**(5300): 657–61.  
[PubMed Abstract](#) | [Publisher Full Text](#)
55. Liu Q, Segal DJ, Ghiara JB, *et al.*: **Design of polydactyl zinc-finger proteins for unique addressing within complex genomes.** *Proc Natl Acad Sci U S A.* 1997; **94**(11): 5525–30.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
56. Sharpe BK, Matthews JM, Kwan AHY, *et al.*: **A new zinc binding fold underlines the versatility of zinc binding modules in protein evolution.** *Structure.* 2002; **10**(5): 639–48.  
[PubMed Abstract](#) | [Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status:  

---

## Version 1

Reviewer Report 23 July 2020

<https://doi.org/10.21956/wellcomeopenres.17642.r39468>

© 2020 Jackson L et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Lauren Jackson** 

Department of Biological Sciences, Vanderbilt University, Nashville, TN, 37235, USA

**Boyang Xie** 

Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

This article reports a non-native X-ray crystallography structure of the small C-terminal zinc finger domain of VPS39, which is hypothesized to bind HOPS tethering complex subunits or other cellular partners. The observed structure adopts an antiparallel b-hairpin structure that in turn forms eight-stranded b-barrels in the crystal. The authors anticipated four residues (Cys841, Cys844, His863, Cys866) would coordinate the zinc ion, but instead the zinc ion is coordinated by non-canonical residues, including two histidines from the His<sub>6</sub> affinity tag. The structure is further stabilized by an intramolecular disulfide bond formed between Cys841 and Cys866. The authors think refolding of the VPS39 domain happened during crystallization; they suggest Cys residues oxidized over an extended 30-month incubation period, which prevented them from competing with the MetHis<sub>6</sub> to bind zinc ion. The authors propose His<sub>6</sub> affinity tags should be used with caution in zinc-binding proteins, and suggest that non-native folds may be promising scaffolds in therapeutic protein design. This study is well-documented and well-presented. We suggest clarification regarding a few minor points in the final version.

### Figure 1

1. We suggest the authors include a sequence alignment between Pcf11 and VPS39 zinc finger domains with key Zn-binding residues marked, since authors use Pcf11 as an expected model for VPS39. This would help the reader follow their logic with a clear visual representation of Cys/His residues predicted to bind zinc.
2. In Figure 1C, the higher bands were presumed to be an SDS-resistant VPS39 dimer. It's possible both native and non-native folds already existed at that point. Were all fractions used for crystallization trials, or did the authors use only fractions containing the single band?
3. Which column was used in Figure 1C? Are standards available to support VPS39 molecular mass?

### Methods

1. The authors mentioned snap-freezing purified VPS39 for storage, but did not specify whether fresh or frozen protein was used in crystallization trials. Could the freeze-thaw cycle affect the protein fold? The authors might comment on whether fresh or frozen protein was used for crystallization set up.
2. Did authors include any zinc in purification buffers? Is this necessary for these proteins, or do the protein just acquire from solution during expression?

**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**

Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**

Yes

**Are all the source data underlying the results available to ensure full reproducibility?**

Yes

**Are the conclusions drawn adequately supported by the results?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Structural biology, biochemistry, biophysics, membrane trafficking

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 07 July 2020

<https://doi.org/10.21956/wellcomeopenres.17642.r39359>

© 2020 Kümmel D. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Daniel Kümmel** 

Institute of Biochemistry, University of Münster, Münster, Germany

This article reports an interesting finding about the crystallization of a small domain in a non-native fold. In their efforts to understand the architecture of the HOPS tethering complex, the authors attempted the crystallization of a predicted zinc-binding RING finger domain of the complex subunit Vps39. The determined structure does not show a RING finger fold, but reveals that the protein adopts a beta-hairpin

structure instead and forms a tetrameric beta-barrel structure in the crystal. This conformation is stabilized by a zinc ion, which does not interact with the expected native coordinating residues but non-canonical residues, including two histidines form the affinity tag used for purification. The structure is further stabilized by a disulfide bridge. The authors conclude that aging of the protein during a long (30 month) crystallization experiment and the presence of a Met-6xHis affinity tag led to refolding of the protein into a non-native structure in the crystallization set-up. They propose that the accidentally created beta-barrel might be a useful template for future protein engineering applications.

The results of the study are well documented and the conclusions are clearly presented. I suggest including one additional aspect in the manuscript, which is to discuss if the recombinant protein might not have been properly folded from the beginning. The following points should be considered:

- SDS-PAGE analysis of SEC fractions showed the presence of an SDS-resistant species of Vps39 CTD during purification. This might arise from the beta-structures observed in the crystal already being present at this point.
- The used construct contains only 35 residues, which is extremely small. Although it contains all predicted zinc-coordinating sites, it might not comprise the full, stable domain.
- A structure-based sequence alignment of the predicted structure of Vps39 CTD (Fig 1B), the observed structure and the structure of Pcf11 zinc finger would be interesting in this context.

Minor point: the spelling of TRIS should be corrected

**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**

Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**

Not applicable

**Are all the source data underlying the results available to ensure full reproducibility?**

Yes

**Are the conclusions drawn adequately supported by the results?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** structural biology, membrane trafficking

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---