



# Crystal structure of the TLDC domain of human NCOA7-AS

Mary Arnaud-Arnould, Marine Tauziet, Olivier Moncorgé, Caroline Goujon and Mickaël Blaise\*

IRIM, CNRS, 34293 Montpellier, France. \*Correspondence e-mail: mickael.blaise@irim.cnrs.fr

Received 12 May 2021

Accepted 2 July 2021

Edited by M. W. Bowler, European Molecular Biology Laboratory, France

**Keywords:** viral restriction factors; oxidation resistance; TLDC; human NCOA7-AS.

**PDB reference:** human TLDC domain, 7obp

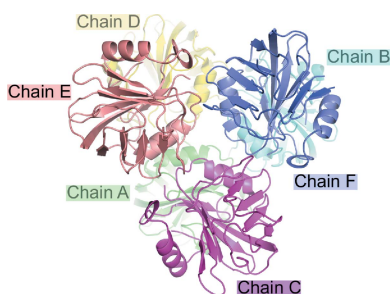
**Supporting information:** this article has supporting information at journals.iucr.org/f

The TLDC [Tre2/Bub2/Cdc16 (TBC), lysin motif (LysM), domain catalytic] domain is associated with oxidation-resistance related functions and is well conserved among eukaryotes. Seven proteins possess a TLDC domain in humans, notably proteins belonging to the oxidation resistance protein (OXR), nuclear receptor coactivator 7 (NCOA7) and TBC1 domain family member 24 (TBC1D24) families. Although the mechanism is unknown, a protective role of TLDC proteins against oxidative stress, notably in the brain, has been demonstrated. Neurobiological disorders caused by mutations in the TLDC domain have also been reported. The human *NCOA7* gene encodes several mRNA isoforms; among these, isoform 4, named *NCOA7-AS*, is up-regulated by type 1 interferon in response to viral infection. NCOA7 and NCOA7-AS both interact with several subunits of the vacuolar proton pump V-ATPase, which leads to increased acidification of the endolysosomal system and consequently impairs infection by viruses that enter their host cells through the endosomal pathway, such as influenza A virus and hepatitis C virus. Similarly to full-length NCOA7, NCOA7-AS possesses a TLDC domain in its C-terminus. Structures of TLDC domains have been reported from zebrafish and fly but not from humans. Here, the expression, purification and crystallization of the TLDC domain from NCOA7 and NCOA7-AS is reported. The crystal structure solved at 1.8 Å resolution is compared with previously solved three-dimensional structures of TLDC domains.

## 1. Introduction

Nuclear receptor coactivator 7 (NCOA7) belongs to the TLDC [Tre2/Bub2/Cdc16 (TBC), lysin motif (LysM), domain catalytic] domain-containing family of proteins. In human and mouse, seven TLDC domain-containing proteins have been reported, among which are oxidation resistance (OXR) proteins 1 and 2 and NCOA7 short and long [alternative start (AS) and full-length (FL), respectively] isoforms (Volkert *et al.*, 2000; Durand *et al.*, 2007; Finelli & Oliver, 2017). NCOA7-FL associates with the estrogen receptor and has been reported to translocate to the nucleus upon estradiol treatment, where it was suggested to act as a transcriptional coregulator (Shao *et al.*, 2002). In addition, TLDC domain-containing proteins have been shown to play a protective role against oxidative stress, notably in the brain, through an unknown mechanism (Finelli *et al.*, 2016; Finelli & Oliver, 2017). The short isoform of NCOA7, NCOA7-AS, does not seem to share this property and is uniquely regulated by type 1 interferon (IFN) via an internal promoter (Yu *et al.*, 2015).

Additional functions have recently been attributed to NCOA7-AS, which plays a significant role in the IFN-induced control of influenza A virus (IAV). NCOA7-AS notably impairs IAV replication in pulmonary cells through the



OPEN ACCESS

regulation of V-ATPase activity (Doyle *et al.*, 2018). As previously shown for NCOA7-FL (Merkulova *et al.*, 2015), NCOA7-AS interacts with several subunits of the vacuolar V-ATPase, the proton pump responsible for endosomal acidification (Doyle *et al.*, 2018). NCOA7-AS and NCOA7-FL mostly share the TLDC domain, pinpointing a probable role for this domain in interaction with the V-ATPase. Through an as yet unelucidated mechanism, this interaction leads to greater acidification of the endolysosomal system, which increases antigen degradation and is detrimental to IAV, possibly by irreversibly affecting the ability of haemagglutinin to allow fusion with the host cells (Doyle *et al.*, 2018). In line with this finding, the *NCOA7* locus has been shown to be important for regulation of V-ATPase function. Interestingly, NCOA7-FL has been shown to interact with the vacuolar V-ATPase in the brain, which enables the correct assembly and transport activity of the proton pump (Castroflorio *et al.*, 2021). Similar to the NCOA7 isoforms, the OXR1 TLDC protein has also been identified as a V-ATPase partner, suggesting that regulation of the V-ATPase could be a common feature of TLDC family members (Merkulova *et al.*, 2015). Despite these findings, the function of the TLDC domain is not yet clearly established and it is not known whether this domain is solely responsible for direct interaction with the V-ATPase.

Mutations in genes encoding TLDC domain-containing proteins have been reported in several human diseases (Finelli & Oliver, 2017). Missense mutations in the TLDC domain found at the C-terminus of the TBC1D24 protein trigger multiple phenotypes, but seem to be particularly linked to epilepsy (Falace *et al.*, 2010; Lüthy *et al.*, 2019). Via its TBC domain in the N-terminus, TBC1D24 regulates synaptic vesicle trafficking, which is exerted through the capacity of this domain to interact with small Rab GTPases (Frasa *et al.*, 2012). However, the function of the TLDC domain of TBC1D24 remains elusive. Little structural information has been gathered on proteins harboring a TLDC domain. The crystal structure of the Skywalker/TBC domain of TBC1D24 has been solved (Fischer *et al.*, 2016) and two studies have reported crystal structures of the OXR2 and TBC1D24 TLDC domains from *Danio rerio* (Blaise *et al.*, 2012) and *Drosophila melanogaster* (Lüthy *et al.*, 2019). The determination of the three-dimensional structure of the TLDC domain from a fly orthologue of human TBC1D24 enabled the pathological mutations linked to human epilepsy to be mapped, and it was proposed that some of these mutations could impair the stability of the protein (Lüthy *et al.*, 2019), illustrating that structural data on the TLDC domain could be of help in deciphering the impact of genetic mutations. However, to date, no structural information is available for any human TLDC domain.

In this context, and because of the important roles and functions of TLDC domain-containing proteins in various cellular processes and because pathological mutations are found in this domain in human diseases, we engaged in structural studies of the human TLDC domain. Here, we report the crystal structure of the TLDC domain of the human NCOA7-AS protein at high resolution.

## 2. Materials and methods

### 2.1. Gene cloning

The TLDC coding sequence of *NCOA7* (encoding amino acids 54–219) was amplified from pRRL.sin.cPPT.SFFV/IRES-puro.WPRE.*NCOA7* variant 6 (Doyle *et al.*, 2018) using the primers 5'-CGGGGTACCGAGAATCTGTACTTCCA GGAATGCGGCCCCACAGCGCGC-3' and 5'-AATTA ATTTACTCGAGTCAATCAAATGCCACACCTCCAG-3'. The PCR fragment was digested and cloned into KpnI/XhoI-digested pET-30 Ek/LIC expression vector (Novagen). The insert is in frame with an S-tag and a His-tag, and a sequence encoding the Tobacco etch virus (TEV) protease cleavage site was inserted upstream of the TLDC coding region to enable tag removal during the purification process.

### 2.2. Protein expression and purification

The recombinant plasmid pET-30 Ek/LIC::TLDC was transformed into the *Escherichia coli* BL21 (DE3) strain resistant to phage T1 (New England Biolabs, Evry, France) carrying pRARE2. One colony was used to inoculate an overnight culture of 500 ml LB medium supplemented with kanamycin (50 µg ml<sup>-1</sup>) and chloramphenicol (34 µg ml<sup>-1</sup>). This culture was diluted in 10 l LB medium supplemented with the two antibiotics. The cells were grown at 289 K to an optical density at 600 nm of 0.8, and protein expression was then induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). The culture was grown overnight at 289 K. The cells were harvested by centrifugation at 8200g for 20 min and were resuspended in 100 ml buffer A (50 mM Tris-HCl pH 8, 400 mM NaCl, 5 mM β-mercaptoethanol, 40 mM imidazole, 1 mM benzamidine). The cells were disrupted by sonication and cell debris was removed by centrifugation at 28 000g for 60 min. The supernatant was loaded at 277 K onto Ni-NTA agarose beads previously equilibrated with buffer A. The beads were washed twice with buffer B (50 mM Tris-HCl pH 8, 1 M NaCl, 5 mM β-mercaptoethanol, 40 mM imidazole, 1 mM benzamidine) and elution was performed with buffer E (50 mM Tris-HCl pH 8, 400 mM NaCl, 5 mM β-mercaptoethanol, 500 mM imidazole). The eluted protein was incubated with His-tagged TEV protease purified in our laboratory in a 1:100(w:w) ratio; the cleavage reaction was performed during dialysis (dialysis-bag cutoff 12–15 kDa) against 1 l dialysis buffer D (50 mM Tris-HCl pH 8, 100 mM NaCl, 5 mM β-mercaptoethanol) overnight at 277 K. After dialysis, the proteins were centrifuged for 20 min at 28 000g and the supernatant was again loaded at 277 K onto Ni-NTA agarose beads equilibrated with buffer D. The TLDC domain without the tag was collected in the flowthrough, concentrated to 5 mg ml<sup>-1</sup> using a Vivaspinn column (10 kDa cutoff), loaded onto a size-exclusion chromatography column (Superdex 75 10/300 GL, GE Healthcare) and eluted with buffer D. Following this protocol, 1 mg highly pure protein as judged by a Coomassie Blue-stained denaturing gel was obtained from 1 l culture. Macromolecule-production information is summarized in Table 1.

**Table 1**  
Macromolecule-production information.

Source organism	<i>H. sapiens</i>
DNA source	pRRL.sin.cPPT.SFFV/IRES-puro. WPRE.NCOA7 variant 6†
Forward primer‡	5'-CGGGTACC <b>GAGAATCTGTACTTCCAG</b> GGAAATGCGGCCACACAGCGCGC-3'
Reverse primer§	5'-AATTAATTTACTCGAGTCAATCAAATG CCCACACTCCAG-3'
Cloning vector	pET-30 Ek/LIC
Expression vector	pET-30 Ek/LIC
Expression host	<i>E. coli</i> BL21 (DE3) transformed with the pRARE2 plasmid
Complete amino-acid sequence of the construct produced	MHHHHHSSGLVPRGSGMKETA <sup>AAK</sup> FERQH MDSPDLGTENLYFQGM <sup>R</sup> PHSALLENMHI EQLARRLPARVQGY <sup>P</sup> WR <sup>L</sup> AYSTLEHGTS LKTLYRKSASLDS <sup>P</sup> VLLVIKMDNQIFG AYATH <sup>P</sup> FKFSDHY <sup>G</sup> TGETFLYTFSPHF KVFKW <sup>S</sup> GENSYFINGDIS <sup>S</sup> LELGGGGGR FGLWLDADLYHGRS <sup>N</sup> SCSTFNNDILSKK EDFIVQDLE <sup>V</sup> WAFD
Complete amino-acid sequence of the construct after TEV cleavage	GMRPHSALLENMHIEQLARRLPARVQGY <sup>P</sup> W RLAYSTLEHGTS <sup>L</sup> KTLYRKSASLDS <sup>P</sup> VL LVIKMDNQIFGAYATH <sup>P</sup> FKFSDHY <sup>G</sup> T GETFLYTFSPHF <sup>K</sup> VFKW <sup>S</sup> GENSYFINGD ISSLELGGGGGRFGLWLDADLYHGRS <sup>N</sup> S CSTFNNDILSKKEDFIVQDLE <sup>V</sup> WAFD

† Doyle *et al.* (2018). ‡ Letters in bold indicate the KpnI restriction site and those in italics indicate the sequence encoding the Tobacco etch virus protease cleavage site. § Letters in bold indicate the XhoI restriction site.

### 2.3. Crystallization

Initial crystallization screening was performed at two protein concentrations: 4.2 and 10.9 mg ml<sup>-1</sup>. The vapor-diffusion method was performed in sitting drops by mixing 0.6 µl protein solution with 0.6 µl reservoir solution using 96-well Swissci MRC plates (Molecular Dimensions, Suffolk, UK) at 291 K. The commercial Structure 1 + 2 (Molecular Dimensions), Index, SaltRx and PEGRx screens (Hampton Research) were assessed. Several hits were obtained and crystal optimization was performed in Swissci 48-Well MRC Maxi Optimization Plates. The best diffraction crystals derived from these optimizations were obtained by mixing 1.5 µl protein solution at 10.9 mg ml<sup>-1</sup> with 1.5 µl reservoir solution consisting of 0.1 M sodium acetate pH 4.5, 32% PEG 300 (Table 2). The crystals were cryocooled in liquid nitrogen without any cryoprotection prior to data collection.

### 2.4. Data collection and processing

X-ray data were collected on the ID30A-1/MASSIF-1 beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. The data set was recorded on a PILATUS3 2M detector (Dectris) at a wavelength of 0.965 Å (12.842 keV) and a crystal-to-detector distance of 190.5 mm. A total of 934 images were collected with an exposure time of 0.097 s, a rotation range of 0.15° and full beam transmission. Data were processed, scaled and merged with *XDS* (Kabsch, 2010) and the data-collection statistics are given in Table 3.

### 2.5. Structure solution and refinement

The structure was solved by molecular replacement performed with *Phaser* (McCoy *et al.*, 2007) from the *Phenix*

**Table 2**  
Crystallization.

Method	Vapor diffusion in sitting drops
Plate type	Swissci 48-Well MRC Maxi Optimization Plates
Temperature (K)	291
Protein concentration (mg ml <sup>-1</sup> )	10.9
Buffer composition of protein solution	50 mM Tris-HCl pH 8, 100 mM NaCl, 5 mM β-mercaptoethanol
Composition of reservoir solution	0.1 M sodium acetate pH 4.5, 32% PEG 300
Volume and ratio of drop	1.5 µl protein solution + 1.5 µl reservoir solution
Volume of reservoir (µl)	250

**Table 3**  
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	ID30A-1, ESRF
Wavelength (Å)	0.965
Temperature (K)	100
Detector	PILATUS3 2M
Crystal-to-detector distance (mm)	190.5
Rotation range per image (°)	0.15
Total rotation range (°)	140
Exposure time per image (s)	0.097
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub>
<i>a</i> , <i>b</i> , <i>c</i> (Å)	68.41, 107.06, 146.44
α, β, γ (°)	90, 90, 90
Mosaicity (°)	0.048
Resolution range (Å)	44.4–1.8 (1.86–1.80)
Total No. of reflections	518696 (47788)
No. of unique reflections	100055 (9863)
Completeness (%)	99.79 (99.78)
Multiplicity	5.2 (4.8)
<i>I</i> /σ( <i>I</i> )	10.10 (1.44)
Wilson <i>B</i> factor (Å <sup>2</sup> )	19.8
<i>R</i> <sub>meas</sub>	0.147 (1.14)
CC <sub>1/2</sub>	0.997 (0.58)

package (Lieschner *et al.*, 2019) using the TLDc structure from zebrafish (PDB entry 4acj; Blaise *et al.*, 2012) as a search model. *Coot* (Emsley *et al.*, 2010) was used for manual rebuilding, while structure refinement and validation were performed with the *Phenix* package. Because of the high resolution of the data, NCS restraints were not applied during refinement. The statistics for structure refinement are displayed in Table 4. Figures were prepared with *PyMOL* (<http://www.pymol.org>).

## 3. Results and discussion

As NCOA7-AS plays an important function, notably as an interferon-induced antiviral inhibitor, we aimed to solve its crystal structure. The 219-amino-acid-long NCOA7-AS can be divided into two parts: an N-terminal domain (residues 1–53) followed by the TLDc domain (residues 54–219). We have performed bioinformatic analyses that predicted the first 53 amino acids to be mainly unfolded, with only the presence of two β-strands formed by residues 10–15 and 24–28. Despite numerous efforts, we have so far been unable to purify full-length NCOA7-AS expressed in *E. coli* to homogeneity. NCOA7-AS has a tendency to aggregate and to form large

**Table 4**

Structure refinement.

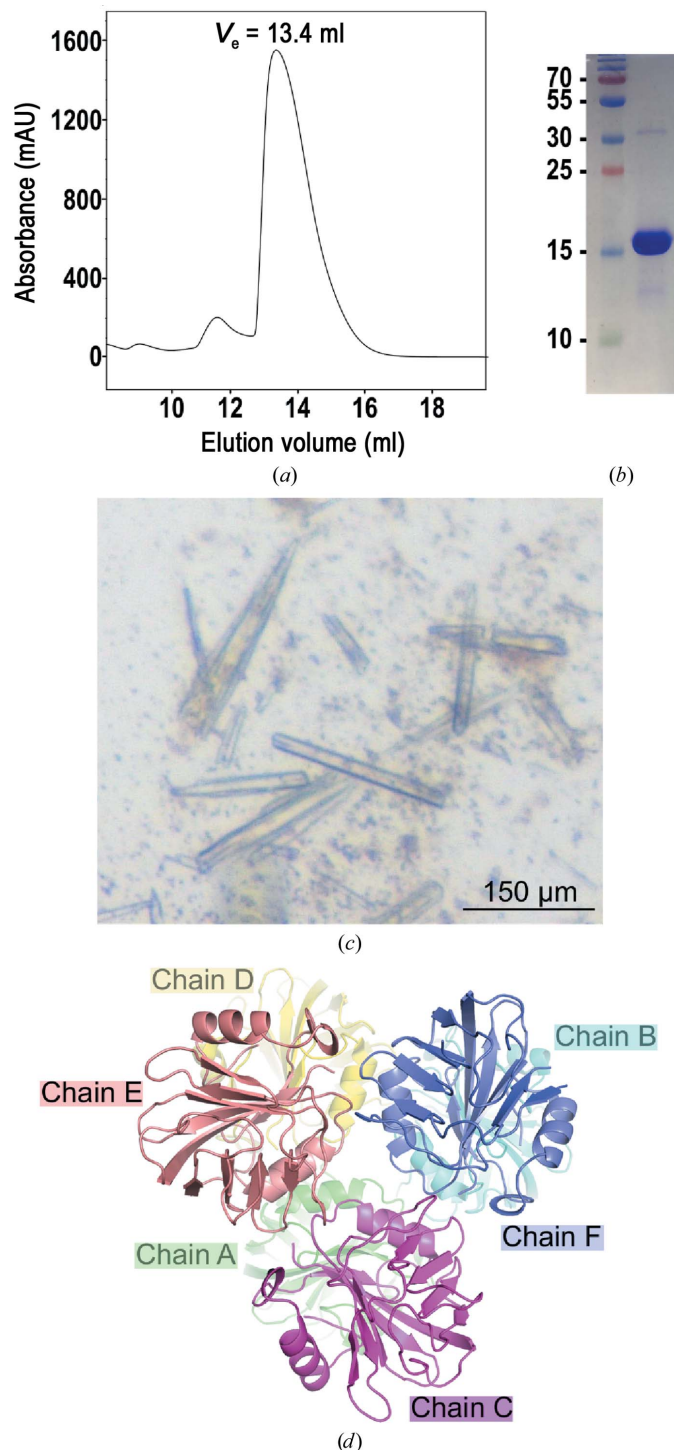
Values in parentheses are for the outer shell.

Resolution range (Å)	44.4–1.8 (1.86–1.80)
Reflections used in refinement	100033 (9861)
Reflections used for $R_{\text{free}}$	2000 (197)
$R_{\text{work}}$	0.176 (0.273)
$R_{\text{free}}$	0.221 (0.334)
No. of non-H atoms	
Total	9256
Macromolecules	8157
Solvent	1099
Protein residues	984
R.m.s.d., bond lengths (Å)	0.005
R.m.s.d., angles (°)	0.80
Ramachandran favored (%)	97.20
Ramachandran allowed (%)	2.80
Ramachandran outliers (%)	0
Rotamer outliers (%)	0.46
Clashscore	2.44
Average $B$ factor (Å <sup>2</sup> )	
Overall	27.88
Macromolecule	26.83
Solvent	35.64
PDB code	7obp

oligomers as judged by size-exclusion chromatography (SEC). We therefore alternatively expressed and purified a truncated form of NCOA7-AS corresponding to the TLDC domain, hereafter referred to as TLDC<sub>HS</sub>, using *E. coli* as an expression host (Table 2). TLDC<sub>HS</sub> could be purified using a three-step chromatography procedure as a very pure and homogeneous material as attested by SEC (Figs. 1*a* and 1*b*).

We could crystallize the domain under several conditions using commercial screens. Crystallization-condition optimization led to rod-shaped crystals with a length of about 100–150 µm (Fig. 1*c*), which were obtained in 0.1 M sodium acetate pH 4.5, 32% PEG 300 (Table 2). A full X-ray data set could be collected and processed to a resolution of 1.8 Å. The crystals belonged to the orthorhombic space group  $P2_12_12_1$ , with unit-cell parameters as indicated in Table 3. The Matthews coefficient ( $V_M$ ) of 2.2 Å<sup>3</sup> Da<sup>-1</sup> assumes the presence of 44.5% solvent and six molecules of TLDC in the asymmetric unit (Fig. 1*d*).

The TLDC<sub>HS</sub> structure was solved by molecular replacement using the TLDC domain (PDB entry 4acj) from *D. rerio* (TLDC<sub>Dr</sub>) as a search model, which shares 62% sequence identity with its human homologue. Six molecules of TLDC<sub>HS</sub> were found in the asymmetric unit, as expected from the  $V_M$ . The structure was manually rebuilt and refined to  $R_{\text{work}}$  and  $R_{\text{free}}$  values of 0.176 and 0.221, respectively, with rather good geometry, as indicated in Table 4. Most residues could be rebuilt for the six monomers, except for the first two amino acids of chain *A* and the first three residues at the N-terminus for chains *B*, *C*, *D*, *F* and *E*. Gly175 was not modeled in chain *A* as well as the glycine stretch ranging from 174 to 176 in chain *F*. Residues 146–148 were also disordered in chain *F* but were well ordered in the other chains. Analysis of the crystal packing with the *PISA* server (Krissinel & Henrick, 2007) revealed the potential existence of a stable complex made of three TLDC<sub>HS</sub> monomers within the asymmetric unit (Fig. 1*d*). Nonetheless, our SEC analysis demonstrated that the protein


**Figure 1**

Purification and crystallization of the human TLDC domain. (a) Elution profile of the purified TLDC domain on a Superdex 75 10/300 GL Increase column. The elution volume of 13.4 ml attests to the presence of a monomer in solution. (b) Coomassie Blue-stained SDS polyacrylamide gel electrophoresis of TLDC performed after the last step of purification by size-exclusion chromatography (10 µg protein; right lane). The left lane contains molecular-mass markers (labelled in kDa). (c) Crystals of the TLDC domain obtained in a sitting drop using PEG 300 as a precipitant. Crystals reached their final size (about 100–200 µm) within two days. (d) Asymmetric unit composition. Six monomers are present in the asymmetric unit, forming two superposed stable trimers as predicted by the *PISA* server (<https://www.ebi.ac.uk/pdbe/pisa/>). Chains *A*, *B* and *D* and chains *C*, *E* and *F* form the two assemblies.





**Table 5**

Comparison of the pathological mutations and associated syndromes found in the TLDC domain from human TBC1D24 with the TLDC domain from *D. melanogaster* TBC1D24 (red) and the TLDC domain from human NCOA7-AS (blue).

The residues that differ between the TLDC domains are highlighted in bold.

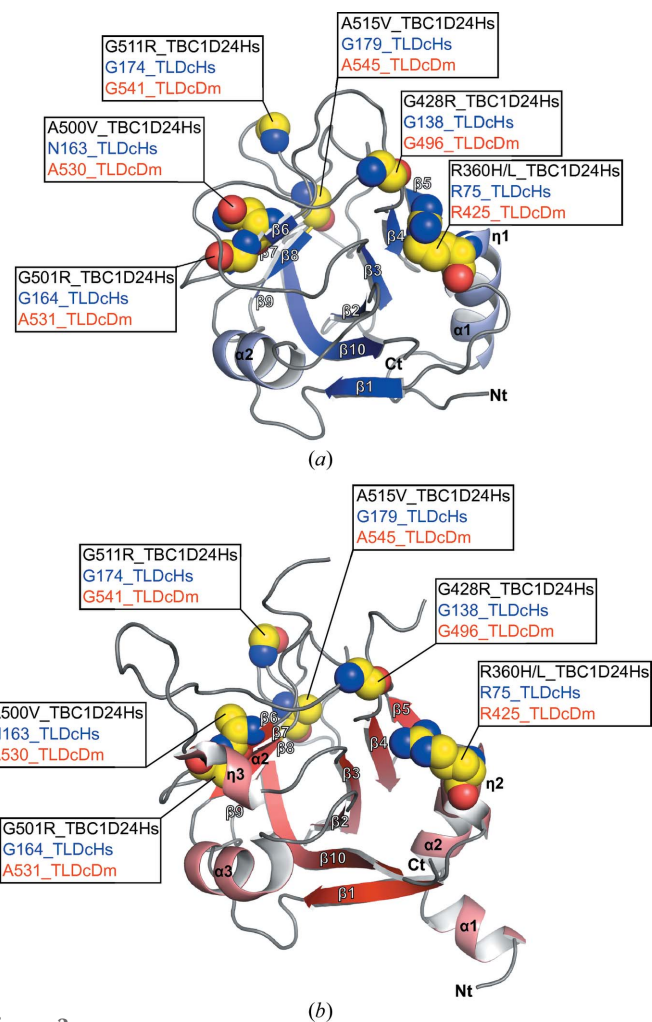
Pathological mutations in the TLDC domain of human TBC1D24 and associated diseases	Residue in TBC1D24	Residue in TLDC <sub>Dm</sub>	Corresponding residue in NCOA7-AS TLDC <sub>Hs</sub>	References
<b>Arg360His/Leu</b> RE-EID (Rolandic epilepsy–writer’s cramp–exercise induced dystonia), progressive myoclonus epilepsy	<b>Arg425</b>		<b>Arg75</b>	Lüthy <i>et al.</i> (2019), Muona <i>et al.</i> (2015)
<b>Gly428Arg</b> RE-EID, DOORS (deafness, onychodystrophy, osteodystrophy and mental retardation syndrome)	<b>Gly496</b>		<b>Gly138</b>	Lüthy <i>et al.</i> (2019), Atli <i>et al.</i> (2018)
<b>Ala500Val</b> RE-EID, focal motor seizures involving the face (with F229S or S473Rfs*43)	<b>Ala530</b>		<b>Asn163</b>	Lüthy <i>et al.</i> (2019), Balestrini <i>et al.</i> (2016), Uzunhan & Uyanik (2020)
<b>Gly501Arg</b> RE-EID	<b>Ala531</b>		<b>Gly164</b>	Lüthy <i>et al.</i> (2019)
<b>Gly511Arg</b> RE-EID	<b>Gly541</b>		<b>Gly174</b>	Lüthy <i>et al.</i> (2019)
<b>Ala515Val</b> RE-EID, familial infantile myoclonic epilepsy, impaired neurite growth and length (with Asp147His)	<b>Ala545</b>		<b>Gly179</b>	Lüthy <i>et al.</i> (2019), Falace <i>et al.</i> (2010)

behaves only as a monomer in solution (Fig. 1a), as also described for the TLDC domain from zebrafish (Blaise *et al.*, 2012).

The overall structure of TLDC<sub>Hs</sub> is globular and consists of two  $\alpha$ -helices in the N-terminus and ten  $\beta$ -strands forming two antiparallel  $\beta$ -sheets. The two sheets, organized as a central pseudo-orthogonal  $\beta$ -sandwich, are made by strands  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$ ,  $\beta_5$  and  $\beta_{10}$  and strands  $\beta_6$ ,  $\beta_7$ ,  $\beta_8$  and  $\beta_9$ , respectively. The N-terminal part of  $\beta_{10}$  interacts with  $\beta_9$  and to this extent contributes to forming the second  $\beta$ -sandwich. The fold is similar to those of the two previously solved TLDC structures from *D. rerio* OXR2 (TLDC<sub>Dr</sub>) and *D. melanogaster* TBC1D24 (TLDC<sub>Dm</sub>). TLDC<sub>Hs</sub> shares 62% sequence identity with TLDC<sub>Dr</sub> (Fig. 2a) and the two structures display an r.m.s.d. of 0.7 Å when superposed over 164 C $\alpha$  atoms. Few structural differences could be observed. TLDC<sub>Hs</sub> does not possess an  $\alpha$ -helix after strand  $\beta_1$  as seen in TLDC<sub>Dr</sub> (Fig. 2b). TLDC<sub>Hs</sub> and TLDC<sub>Dm</sub> are more distant as they present only 38% sequence identity (Fig. 2a), and superimposition of the two structures leads to an r.m.s.d. of 1.1 Å over 144 residues. The N-terminal sequences are not well conserved in the three proteins compared with the rest of the sequence, which is reflected by a few differences at the three-dimensional level. TLDC<sub>Dm</sub> possesses one extra helix ( $\alpha_2$ ) in the N-terminus that is not seen in TLDC<sub>Hs</sub> or TLDC<sub>Dr</sub>. Finally, noticeable differences are found as two extended loops between  $\beta_3$  and  $\beta_4$  and  $\beta_5$  and  $\beta_6$  in TLDC<sub>Dm</sub>.

Several missense mutations have been reported in the TLDC domain from TBC1D24 (Table 5) that are linked to human diseases (Falace *et al.*, 2010; Balestrini *et al.*, 2016; Wang *et al.*, 2019; Lüthy *et al.*, 2019; Muona *et al.*, 2015; Uzunhan & Uyanik, 2020; Atli *et al.*, 2018). The three-dimensional structure of TLDC<sub>Hs</sub> therefore offers the possibility to map these mutations onto the domain. We performed mapping of these mutations onto TLDC<sub>Hs</sub> (Figs. 2a and 3) and assessed the conservation of the residues with missense mutations mapped onto TLDC<sub>Dm</sub> from fly TBC1D24 as described previously by Lüthy *et al.* (2019). The pathological Arg360His or Arg360Leu mutation found in the human TBC1D24 protein is situated in  $3_{10}$ -helix  $\eta_1$ ; the equivalent residue in TLDC<sub>Hs</sub> from NCOA7-AS is Arg75. Arg75 is involved in a salt-bridge interaction with the side chain of the

highly conserved Glu139 (not shown), and mutation(s) breaking this interaction might therefore destabilize the long loop spanning between strands  $\beta_4$  and  $\beta_5$  (Figs. 2a and 3). The



**Figure 3** Structural mapping of the pathological mutations found in the human TLDC domain from TBC1D24. The mutations are reported in the crystal structures of TLDC<sub>Hs</sub> from NCOA7-AS (a) and TLDC<sub>Dm</sub> from fly TBC1D24 (b). The human pathological missense mutations reported in TBC1D24 are indicated in black, while the corresponding residues are in blue for the TLDC<sub>Hs</sub> structure and in red for the TLDC<sub>Dm</sub> structure. Residues corresponding to the mutations are shown as yellow spheres.

reported Gly428Arg mutation corresponds to Gly138 in human TLDC and is situated just before strand  $\beta 4$ . The Ala500 residue mutated to Val in human TBC1D24 is not conserved in TLDC<sub>HS</sub> from NCOA7-AS, where Asn163 is instead found. Despite this lack of residue conservation, both amino acids are part of strand  $\beta 6$ . Gly501 in TBC1D24 and the equivalent Gly164 in TLDC<sub>HS</sub> are also part of strand  $\beta 6$ . Mutation(s) in a structured region such as strand  $\beta 6$  will probably affect the folding of this strand. The Gly511 residue mutated to Arg matches Gly174 in TLDC<sub>HS</sub> and is situated in a loop formed by a five-glycine stretch between strands  $\beta 7$  and  $\beta 8$ . Finally, Ala515, which has been reported to be mutated to Val in several studies, is not conserved in TLDC<sub>HS</sub> as Gly179 is instead found. Nonetheless, both residues are at the beginning of strand  $\beta 8$ , and these mutations could destabilize the integrity of this strand. The missense mutations found in the TLDC domain of human TBC1D24 which are linked to neurological disorders are all found at structurally conserved positions. It is indeed clear that the Gly511Arg and Gly501Arg mutations as well as the Arg360His/Leu, Ala500Val and Ala515Val mutations could disturb the folding of the TLDC domain and potentially destabilize the protein because of the nature of the amino-acid substitution and/or its position in the structurally conserved regions. This analysis is therefore in strong agreement with the previously proposed effect on protein stability of the TLDC domain induced by pathological mutations found in TBC1D24 (Lüthy *et al.*, 2019).

To conclude, this structural work strongly attests to the high protein similarity between TLDC domains from vertebrates as well as, albeit slightly more distantly, that from an arthropod. Although this was expected from the high sequence identity that is shared between these proteins, reporting the first crystal structure of a human TLDC domain is notably of interest because of the versatile roles of TLDC proteins and also because pathological mutations affect this domain.

### Acknowledgements

We thank the staff at the MASSIF-1 beamline at ESRF for support during data collection as well as the staff at the PXIII beamline at the Swiss Light Source and J. McKellar for critical reading of the manuscript.

### Funding information

This work was funded by the National Research Agency (ANR-20-CE15-0010-02-CAIPIRINAS to CG and MB) and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Starting Grant, grant agreement No. 759226 to CG), the ATIP-Avenir programme (to CG) and institutional funds from the Centre National de la Recherche Scientifique (CNRS) and Montpellier University (to CG and MB).

### References

Atli, E., Gurkan, H., Ulusal, S., Karal, Y., Atli, E. I. & Tozkir, H. (2018). *Clin. Dysmorphol.* **27**, 1–3.

Balestrini, S., Milh, M., Castiglioni, C., Lüthy, K., Finelli, M. J., Verstreken, P., Cardon, A., Stražišar, B. G., Holder, J. L., Lesca, G., Mancardi, M. M., Poulat, A. L., Repetto, G. M., Banka, S., Bilo, L., Birkeland, L. E., Bosch, F., Brockmann, K., Cross, J. H., Doummar, D., Félix, T. M., Giuliano, F., Hori, M., Hüning, I., Kayserili, H., Kini, U., Lees, M. M., Meenakshi, G., Mewasingh, L., Pagnamenta, A. T., Peluso, S., Mey, A., Rice, G. M., Rosenfeld, J. A., Taylor, J. C., Troester, M. M., Stanley, C. M., Ville, D., Walkiewicz, M., Falace, A., Fassio, A., Lemke, J. R., Biskup, S., Tardif, J., Ajeawung, N. F., Tolun, A., Corbett, M., Gecz, J., Afawi, Z., Howell, K. B., Oliver, K. L., Berkovic, S. F., Scheffer, I. E., de Falco, F. A., Oliver, P. L., Striano, P., Zara, F., Campeau, P. M. & Sisodiya, S. M. (2016). *Neurology*, **87**, 77–85.

Blaise, M., Alsarraf, H. M. A. B., Wong, J. E. M. M., Midtgaard, S. R., Laroche, F., Schack, L., Spink, H., Stougaard, J. & Thirup, S. (2012). *Proteins*, **80**, 1694–1698.

Castroflorio, E., den Hoed, J., Svistunova, D., Finelli, M. J., Cebrian-Serrano, A., Corrochano, S., Bassett, A. R., Davies, B. & Oliver, P. L. (2021). *Cell. Mol. Life Sci.* **78**, 3503–3524.

Doyle, T., Moncorgé, O., Bonaventura, B., Pollpeter, D., Lussignol, M., Tauziet, M., Apollonia, L., Catanese, M.-T., Goujon, C. & Malim, M. H. (2018). *Nat. Microbiol.* **3**, 1369–1376.

Durand, M., Kolpak, A., Farrell, T., Elliott, N. A., Shao, W., Brown, M. & Volkert, M. R. (2007). *BMC Cell Biol.* **8**, 13.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.

Falace, A., Filipello, F., La Padula, V., Vanni, N., Madia, F., De Pietri Tonelli, D., de Falco, F. A., Striano, P., Dagna Bricarelli, F., Minetti, C., Benfenati, F., Fassio, A. & Zara, F. (2010). *Am. J. Hum. Genet.* **87**, 365–370.

Finelli, M. J. & Oliver, P. L. (2017). *Mamm. Genome*, **28**, 395–406.

Finelli, M. J., Sanchez-Pulido, L., Liu, K. X., Davies, K. E. & Oliver, P. L. (2016). *J. Biol. Chem.* **291**, 2751–2763.

Fischer, B., Lüthy, K., Paesmans, J., De Koninck, C., Maes, I., Swerts, J., Kuenen, S., Uytterhoeven, V., Verstreken, P. & Versionées, W. (2016). *Nat. Struct. Mol. Biol.* **23**, 965–973.

Frasa, M. A. M., Koessmeier, K. T., Ahmadian, M. R. & Braga, V. M. M. (2012). *Nat. Rev. Mol. Cell Biol.* **13**, 67–73.

Kabsch, W. (2010). *Acta Cryst. D* **66**, 133–144.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst. D* **75**, 861–877.

Lüthy, K., Mei, D., Fischer, B., De Fusco, M., Swerts, J., Paesmans, J., Parrini, E., Lubarr, N., Meijer, I. A., Mackenzie, K. M., Lee, W.-T., Cittaro, D., Aridon, P., Schoovaerts, N., Versées, W., Verstreken, P., Casari, G. & Guerrini, R. (2019). *Brain*, **142**, 2319–2335.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Merkulova, M., Păunescu, T. G., Azroyan, A., Marshansky, V., Breton, S. & Brown, D. (2015). *Sci. Rep.* **5**, 14827.

Muona, M., Berkovic, S. F., Dibbens, L. M., Oliver, K. L., Maljevic, S., Bayly, M. A., Joensuu, T., Canafoglia, L., Franceschetti, S., Michelucci, R., Markkinen, S., Heron, S. E., Hildebrand, M. S., Andermann, E., Andermann, F., Gambardella, A., Tinuper, P., Licchetta, L., Scheffer, I. E., Criscuolo, C., Filla, A., Ferlazzo, E., Ahmad, J., Ahmad, A., Baykan, B., Said, E., Topcu, M., Riguzzi, P., King, M. D., Ozkara, C., Andrade, D. M., Engelsens, B. A., Crespel, A., Lindenau, M., Lohmann, E., Saletti, V., Massano, J., Privitera, M., Espay, A. J., Kauffmann, B., Duchowny, M., Møller, R. S., Straussberg, R., Afawi, Z., Ben-Zeev, B., Samocha, K. E., Daly, M. J., Petrou, S., Lerche, H., Palotie, A. & Lehesjoki, A.-E. (2015). *Nat. Genet.* **47**, 39–46.

- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Shao, W., Halachmi, S. & Brown, M. (2002). *Mol. Cell. Biol.* **22**, 3358–3372.
- Uzunhan, T. A. & Uyanik, B. (2020). *Clin. Neurol. Neurosurg.* **196**, 106080.
- Volkert, M. R., Elliott, N. A. & Housman, D. E. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 14530–14535.
- Wang, J., Rousseau, J., Kim, E., Ehresmann, S., Cheng, Y.-T., Duraine, L., Zuo, Z., Park, Y.-J., Li-Kroeger, D., Bi, W., Wong, L.-J., Rosenfeld, J., Gleeson, J., Faqeih, E., Alkuraya, F. S., Wierenga, K. J., Chen, J., Afenjar, A., Nava, C., Doummar, D., Keren, B., Juusola, J., Grompe, M., Bellen, H. J. & Campeau, P. M. (2019). *Am. J. Hum. Genet.* **105**, 1237–1253.
- Yu, L., Croze, E., Yamaguchi, K. D., Tran, T., Reder, A. T., Litvak, V. & Volkert, M. R. (2015). *J. Interferon Cytokine Res.* **35**, 186–199.