Crystal structure of the TLDc domain of human NCOA7-AS
Mary Arnaud-Arnould, Marine Tauziet, Olivier Moncorgé, Caroline Goujon, Mickaël Blaise

To cite this version:
Mary Arnaud-Arnould, Marine Tauziet, Olivier Moncorgé, Caroline Goujon, Mickaël Blaise. Crystal structure of the TLDc domain of human NCOA7-AS. Acta crystallographica Section F: Structural biology communications [2014-..], International Union of Crystallography 2021, 77, 10.1107/s2053230x21006853. hal-03306175

HAL Id: hal-03306175
https://hal.archives-ouvertes.fr/hal-03306175
Submitted on 28 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Crystal structure of the TLDc domain of human NCOA7-AS

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

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

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
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 human epilepsy (Falace et al., 2010; Lu¨ thy 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., 2021). The Skywalker/TBC 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 (Lu¨ thy 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 (Lu¨ thy 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’-CGGGTACCGAGAATCTGTACTTCCA GGGAATGCAGCCCCACAGGCCGC-3’ and 5’-AAATTA ATTTACTCGAGTCAATCAAATGCCCACACCTTCCAG-3’. The PCR fragment was digested and cloned into KpnI/XhoI-digested pET-30 Ekc/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 Ekc/LIC::TLDc was transformed into the Escherichia coli BL21 (DE3) strain carrying pRARE2. One colony was used to inoculate an overnight culture of 500 ml LB medium supplemented with kanamycin (50 µg ml⁻¹) and chloramphenicol (34 µg ml⁻¹). 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⁻¹ using a Vivaspin 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 | DNA source | Forward primer† | Reverse primer§ |
|-----------------|------------|-----------------|-----------------|
| H. sapiens      | pRRL.sin.cPETESFFV/IRESPuro. | 5’-GGGGGATCCAGAATCCTGACTTCCGAGAATGGGCGGGCCACAGCCGGC-3’ | 5’-AATTAATTTACTCGAGTCATCAAAATTCCCCACCTCAG-3’ |
| Cloning vector | pET-30 Ek/LIC | E. coli BL21 (DE3) transformed with the pRARE2 plasmid |

| Expression vector | Expression host |
|-------------------|-----------------|
| pET-30 Ek/LIC     | E. coli BL21 (DE3) |

| Complete amino-acid sequence of the construct produced |
|------------------------------------------------------|
| MHHHHHHSSGLVPRGGGKETAAAKFERQH MDSFLGTONLYQQRMPHARS |
| Selfassembly in sitting drops by mixing 0.6 M protein solution at 10.9 mg ml⁻¹ |

| Complete amino-acid sequence of the construct after TEV cleavage |
|----------------------------------------------------------------|
| GFP/HSALENHMLEQQLARRLPLAQRGYPW |
| Initial crystallization screening was performed at two protein concentrations: 4.2 and 10.9 mg ml⁻¹. The vapor-diffusion method was performed in sitting drops by mixing 0.6 M protein solution with 0.6 M reservoir solution using 96-well Swissi 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 Swissi 48-Well MRC Maxi Optimization Plates. The best diffraction crystals derived from these optimizations were obtained by mixing 1.5 M protein solution at 10.9 mg ml⁻¹ with 1.5 M 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. |

2.3. Crystallization

Initial crystallization screening was performed at two protein concentrations: 4.2 and 10.9 mg ml⁻¹. The vapor-diffusion method was performed in sitting drops by mixing 0.6 M protein solution with 0.6 M reservoir solution using 96-well Swissi 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 Swissi 48-Well MRC Maxi Optimization Plates. The best diffraction crystals derived from these optimizations were obtained by mixing 1.5 M protein solution at 10.9 mg ml⁻¹ with 1.5 M 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.

2.5. Structure solution and refinement

The structure was solved by molecular replacement performed with Phaser (McCoy et al., 2007) from the Phenix package (Liebschner 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
We therefore alternatively expressed and purified a truncated form of NCOA7-AS corresponding to the TLDc domain, hereafter referred to as TLDcHs, using *E. coli* as an expression host (Table 2). TLDcHs could be purified using a three-step chromatography procedure as a very pure and homogeneous material as attested by SEC (Figs. 1a and 1b).

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. 1c), 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 *P*2₁2₁2₁, with unit-cell parameters as indicated in Table 3. The Matthews coefficient (*V_M*) of 2.2 Å³ Da⁻¹ assumes the presence of 44.5% solvent and six molecules of TLDc in the asymmetric unit (Fig. 1d).

The TLDcHs structure was solved by molecular replacement using the TLDc domain (PDB entry 4acj) from *D. rerio* (TLDcDr) as a search model, which shares 62% sequence identity with its human homologue. Six molecules of TLDcHs were found in the asymmetric unit, as expected from the *V_M*. The structure was manually rebuilt and refined to *R*_work and *R*_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 TLDcHs monomers within the asymmetric unit (Fig. 1d).

Nonetheless, our SEC analysis demonstrated that the protein oligomers as judged by size-exclusion chromatography (SEC). The 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*_free | 2000 (197) |
| *R*_work | 0.176 (0.273) |
| *R*_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 (Å²) |
| Overall | 27.88 |
| Macromolecule | 26.83 |
| Solvent | 35.64 |

Nonetheless, our SEC analysis demonstrated that the protein 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 TLDcHs, using *E. coli* as an expression host (Table 2). TLDcHs could be purified using a three-step chromatography procedure as a very pure and homogeneous material as attested by SEC (Figs. 1a and 1b). 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. 1c), 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 *P*2₁2₁2₁, with unit-cell parameters as indicated in Table 3. The Matthews coefficient (*V_M*) of 2.2 Å³ Da⁻¹ assumes the presence of 44.5% solvent and six molecules of TLDc in the asymmetric unit (Fig. 1d).

The TLDcHs structure was solved by molecular replacement using the TLDc domain (PDB entry 4acj) from *D. rerio* (TLDcDr) as a search model, which shares 62% sequence identity with its human homologue. Six molecules of TLDcHs were found in the asymmetric unit, as expected from the *V_M*. The structure was manually rebuilt and refined to *R*_work and *R*_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 TLDcHs monomers within the asymmetric unit (Fig. 1d).

Nonetheless, our SEC analysis demonstrated that the protein 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 TLDcHs, using *E. coli* as an expression host (Table 2). TLDcHs could be purified using a three-step chromatography procedure as a very pure and homogeneous material as attested by SEC (Figs. 1a and 1b).

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. 1c), 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 *P*2₁2₁2₁, with unit-cell parameters as indicated in Table 3. The Matthews coefficient (*V_M*) of 2.2 Å³ Da⁻¹ assumes the presence of 44.5% solvent and six molecules of TLDc in the asymmetric unit (Fig. 1d).

The TLDcHs structure was solved by molecular replacement using the TLDc domain (PDB entry 4acj) from *D. rerio* (TLDcDr) as a search model, which shares 62% sequence identity with its human homologue. Six molecules of TLDcHs were found in the asymmetric unit, as expected from the *V_M*. The structure was manually rebuilt and refined to *R*_work and *R*_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 TLDcHs monomers within the asymmetric unit (Fig. 1d).

Nonetheless, our SEC analysis demonstrated that the protein 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 TLDcHs, using *E. coli* as an expression host (Table 2). TLDcHs could be purified using a three-step chromatography procedure as a very pure and homogeneous material as attested by SEC (Figs. 1a and 1b).

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. 1c), 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 *P*2₁2₁2₁, with unit-cell parameters as indicated in Table 3. The Matthews coefficient (*V_M*) of 2.2 Å³ Da⁻¹ assumes the presence of 44.5% solvent and six molecules of TLDc in the asymmetric unit (Fig. 1d).
Figure 2
Protein sequences and structural comparisons of the TLDc domains. (a) The multiple sequence alignment was performed with ENDscript (Robert & Gouet, 2014) and adjusted manually. The secondary structures (α, α-helix; β, β-strand; η, 3_10-helix; TT, turn) of the three TLDc structures extracted from the crystal structures are indicated above the alignment. The green spheres below the alignment indicate the position of missense mutations that are found in the TLDc domain of TBC1D24 and are associated with human diseases. (b) Structural comparison of the three TLDc domains from D. rerio (PDB entry 4acj; left; brown), H. sapiens (PDB entry 7obp; middle; blue) and D. melanogaster (PDB entry 6r82; right; red).
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 TLDcHs is globular and consists of two α-helices in the N-terminus and ten β-strands forming two antiparallel β-sheets. The two sheets, organized as a central pseudo-orthogonal β-sandwich, are made by strands β1, β2, β3, β4, β5 and β10 and strands β6, β7, β8 and β9, respectively. The N-terminal part of β10 interacts with β9 and to this extent contributes to forming the second β-sandwich. The fold is similar to that of the two previously solved TLDc structures from D. rerio OXR2 (TLDcDr) and D. melanogaster TBC1D24 (TLDcDm). TLDcHs shares 62% sequence identity with TLDcDr (Fig. 2a) and the two structures display an r.m.s.d. of 0.7 Å when superposed over 144 Cα atoms. Few structural differences could be observed. TLDcDr does not possess an α-helix after strand β1 as seen in TLDcDm (Fig. 2b). TLDcHs and TLDcDm 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. TLDcDm possesses one extra helix (α2) in the N-terminus that is not seen in TLDcHs or TLDcDr. Finally, noticeable differences are found as two extended loops between β3 and β4 and β5 and β6 in TLDcDm.

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 TLDcHs therefore offers the possibility to map these mutations onto the domain. We performed mapping of these mutations onto TLDcHs (Figs. 2a and 3) and assessed the conservation of the residues with missense mutations mapped onto TLDcDm 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 3α-helix η1; the equivalent residue in TLDcHs 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 β4 and β5 (Figs. 2a and 3). The
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 TLDcHs 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 TLDcHs 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 TLDcHs, 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 TLDcHs 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 of the amino-acid stretch 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).

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. & Tozkar, H. (2018). Clin. Dysmorphol. 27, 1–3.
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., Faqeh, 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.