Crystal structure of the TLDc domain of oxidation resistance protein 2 from zebrafish

Blaise, M.; Alsaraf, J.E.; Wong, J.E.; Midtgaard, S.R.; Laroche, F.; Spaink, H.P.; ... ; Thirup, S.

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Crystal structure of the TLDc domain of oxidation resistance protein 2 from zebrafish

Mickaël Blaise,[1*] Husam M. A. B. Alsarraf,[1] Jaslyn E. M. M. Wong,[1] Søren Roi Midtgaard,[2] Fabrice Laroche,[1,3] Lotte Schack,[1] Herman Spaink,[1,3] Jens Stougaard,[1] and Søren Thirup[1]

[1] Department of Molecular Biology and Genetics, Centre for Carbohydrate Recognition and Signalling, Aarhus University, Aarhus, Denmark

[2] Nanobioscience Group, Faculty of LIFE Sciences, University of Copenhagen, Copenhagen, Denmark

[3] Institute of Biology, Leiden University, Leiden, The Netherlands

ABSTRACT

The oxidation resistance proteins (OXR) help to protect eukaryotes from reactive oxygen species. The sole C-terminal domain of the OXR, named TLDc is sufficient to perform this function. However, the mechanism by which oxidation resistance occurs is poorly understood. We present here the crystal structure of the TLDc domain of the oxidation resistance protein 2 from zebrafish. The structure was determined by X-ray crystallography to atomic resolution (0.97 Å) and adopts an overall globular shape. Two antiparallel β-sheets form a central β-sandwich, surrounded by two helices and two one-turn helices. The fold shares low structural similarity to known structures.

Proteins 2012; 80:1694–1698.

Key words: oxidation resistance; Ncoa7; ERAP140; TBC1D24; TLDc; reactive oxygen species.

INTRODUCTION

Reactive oxygen species (ROS) are highly reactive byproducts of oxygen produced during normal cellular metabolism, such as respiration. Due to their high reactivity, ROS can react with lipids, proteins or nucleic acids and trigger damage to the cells.[1] ROS can, for example, modify DNA and initiate tumerogenesis. Cells have developed numerous strategies to detoxify ROS and to keep ROS concentration at low level. Numerous enzymes are indeed involved in ROS processing; among them are catalase, superoxide dismutase, and peroxidase. Small compounds such as vitamin C or glutathione also play important roles in ROS protection.[2]

In 2000, Volkert *et al.*[3] established an *Escherichia coli* based genetic screen to identify potential new human genes involved in ROS protection. In this study, the authors identified a gene encoding a novel protein, named oxidation resistance protein (OXR). Expression of the human OXR1 protein reversed an oxidative mutator phenotype when expressed in *E. coli*. Furthermore, the role of the OXR1 protein has also been confirmed in other eukaryotic species. The yeast OXR1 protein protects cells from oxidative stress and as observed for its human homolog, gene expression is upregulated under oxidative stress.[4] More recently, the involvement of OXR1 in oxidative stress protection was demonstrated in mosquito and mouse neurons.[5,6] Interestingly, the sole C-terminal (C-t) domain of the OXR1 protein was sufficient to confer protection against ROS species.[6,7]

Additional Supporting Information may be found in the online version of this article.

Grant sponsors: Danish National Research Foundation and Danscatt

*Correspondence to: Mickaël Blaise, Department of Molecular Biology and Genetics, Centre for Carbohydrate Recognition and Signalling, Gustav Wieds vej 10c, Aarhus University, Aarhus Denmark. E-mail: mick@mb.au.dk

Received 16 December 2011; Revised 23 January 2012; Accepted 25 January 2012
Published online 6 February 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24050

The OXR1 C-t domain was named TLDc[8] (TBC, LysM, Domain catalytic) because this domain is very often associated to either TBC (Tre2/Bub2/Cdc16) domain found in Rab GTPAse-activating protein or LysM (Lysin Motif) domain that is involved in carbohydrate binding. Genes encoding only the TLDc domain have also been found.[8] The TLDc domain is present in all eukaryotes, and the primary sequence is highly conserved among species [Fig. 1(A)]. The TLDc domain has been identified in the NCOA7 protein, an OXR1 homolog. NCOA7 is able to protect against oxidative damages,[7] but the NCOA7 cytoplasmic[7] localization differs from that of OXR1, which is mitochondrial.[4] NCOA7 interacts with the estrogen receptor and relocalizes to the nucleus on estrogen treatment.[7,9] The C-terminal domain from NCOA7 is sufficient for oxidative protection, and a role of NCOA7 in detoxification of oxidized estrogen was therefore proposed.[7]

We have recently identified a NCOA7 homolog named oxidation resistance protein 2 (OXR2, Genbank:JQ649325) in the zebrafish genome. As for the human NCOA7, several spliced variants have been identified for OXR2 (Laroche *et al.*, in preparation). According to the SMART server,[10] the longest splice variant encodes the following domains from N-t to C-t: a predicted unstructured region of 95 amino acids followed by a LysM domain (residues 96:139), a GRAM domain (residues 145:164), a coiled coil (residues 434:463), and the TLDc domain (residues 635:801).

The involvement of the OXR protein family and the TLDc in oxidative damage protection has been clearly demonstrated. Nevertheless, the precise function of the TLDc and its three-dimensional structure remain unknown. As a part of a study to identify the function, of the zebrafish LysM containing proteins, we have investigated the structure of the OXR2 protein. We have recently reported the crystallization and preliminary structural studies of the TLDc domain of OXR2 from zebrafish.[11] We present here the crystal structure of the TLDc domain of OXR2 at 0.97 Å resolution.

MATERIAL AND METHODS

Protein expression and purification

The TLDc domain of OXR2 was expressed, purified, and crystallized as described previously.[11] Briefly, the TLDc gene region encoding for amino acid 636–801 of OXR2 was cloned into the pET-41 Ek/LIC expression vector (Novagen). The protein was expressed with a N-t tag, containing Glutathione-S-Transferase, 6xHis, and S-tags. A Tobacco Etch Virus protease (TEV) cleavage site was inserted after the three tags, which on cleavage leaves the protein with an additional N-t glycine. *E. coli* BL21 codon plus RIL strain was used for expression. The protein was purified as follows: first a nickel affinity purification, second a nickel affinity purification after cleavage of the N-t tag by the TEV protease, and a third step, by size exclusion chromatography on a 24 mL Superdex 75 10/300GL (GE Healthcare). The selenomethionine substituted protein was purified under similar conditions.[11]

Crystallization and data collection

The protein was crystallized using the sitting drop method at 4°C, in 100 m*M* Na-HEPES pH 8, 1.8 *M* $(NH_4)_2SO_4$, by mixing 1.5 µL of reservoir solution with 1.5 µL of a protein solution at 8.7 mg/mL. The selenomethionine-substituted protein was crystallized in the same condition but at a concentration of 7.0 mg/mL. Crystals were soaked in 100 m*M* Na-HEPES pH 8, 1.8 *M* $(NH_4)_2SO_4$ and 10% glycerol before flash cooling in liquid nitrogen. A native dataset was collected on the I911-3 beamline at a wavelength of 1 Å, at Max-lab, Lund, Sweden. In addition, a single-wavelength anomalous dispersion (SAD) dataset was collected at a wavelength of 0.978 Å, at the X06DA beamline of the Swiss Light Source, Villigen, Switzerland. Data were processed with the XDS program[12] (Table I).

Phasing and structure refinement

The structure was determined by SAD and all six Se sites were located using the program AUTOSOL from the PHENIX package.[13] A first model was obtained by combining automatic and manual building using, respectively, the AUTOBUILD[14] program from the PHENIX package and COOT.[15] The resulting partial model was further used as a search model for molecular replacement against the high-resolution native dataset, using PHASER[16] from the PHENIX package. The model was refined to 0.97 Å resolution with the PHENIX package: after a first rigid body refinement, five steps of refinement and rebuilding using PHENIX and COOT were performed. Each refinement step was composed of five macrocycles, and included, refinement of individual atomic coordinates using gradient driven minimization (LBFGS) and individual isotropic *B*-factors. Furthermore, water molecules were placed automatically by PHENIX. When R_{free} reached $\sim$ 20%, anisotropic *B*-factor refinement was used for both protein and water molecules. Nine steps of LBFGS and individual atomic *B*-factors refinement, and rebuilding were then performed. Alternate conformations were included during these steps, and their occupancies were refined. At this stage, riding hydrogens were included in the refinement, and water molecules were manually checked. Finally, two additional steps of refinement (limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS), *B*-factors and occupancy) were performed, during which weights between X-ray target and stereochemistry were automatically optimized by PHENIX. Translation, libration and screw motion (TLS) were not included in the refinement strategy as it was increasing the *R*-values and was not improving the refinement. The final model possesses 25 side chains in alternative conformations. The structure quality was assessed with MOLPROBITY[17] and POLYGON from the PHENIX package (Table I).

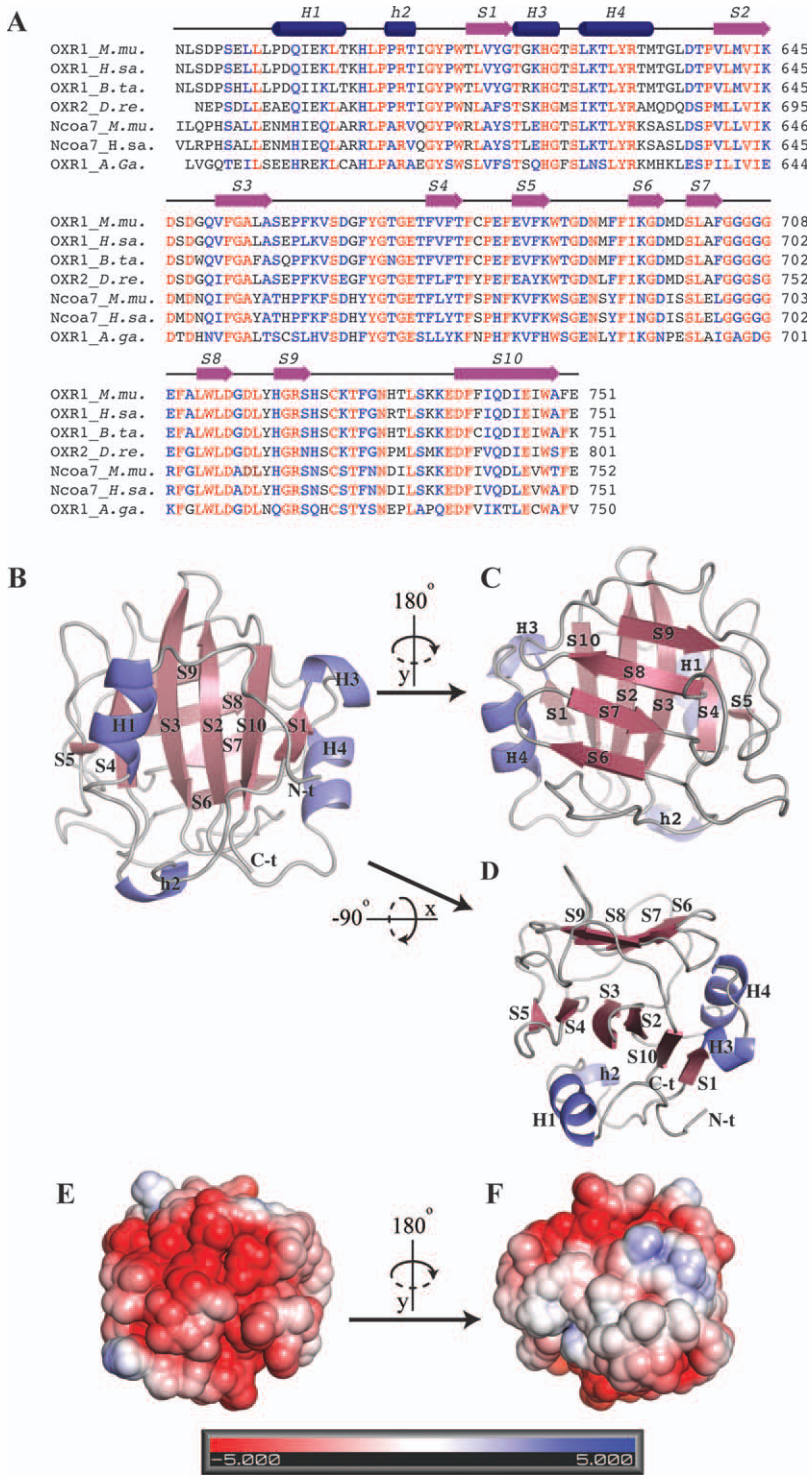

Figure 1

Overall structure of the TLDc domain. **A:** Protein sequence alignment of the TLDc domain of OXR1, NCOA7, and OXR2 proteins from, *Mus musculus* (*M.mu.*), *Homo sapiens* (*H.sa.*), *Bos taurus* (*B.ta.*), Zebrafish (*D.re.*), and *Anophele gambiae* (*A.ga.*). The TLDc secondary structure is indicated on top of the protein sequence: blue arrows for strands and magenta cylinders for helices. The sequence alignment color code is as follows, nonconserved residues are black, semiconserved residues are blue, and strictly conserved residues are red. **B, C, D:** Overall structure of the TLDc domain, the structure is colored according to the secondary structure, magenta for strands, blue for helices, and gray for coils. S1 to S10 and H1 to H4 indicate, respectively, the strands and helices numbering; h2 and H3 refer, respectively, to a single turn of 3_{10}-helix type and a single turn of α-helix. **E, F:** Electrostatic potential of the TLDc structure calculated with *Pymol* with the APBS plugin. The units are given in kT/e. All the figures were generated with *Pymol* (www.pymol.org).

Table I
Data Collection and Refinement Statistics

Data collection	Native	Se-peak
Beamline	Max-lab I911-3	SLS X06DA
Wavelength (Å)	1.0	0.978
Space group	$P2_12_12$	$P2_12_12$
Cell dimensions		
a, *b*, *c* (Å)	65.62, 69.11, 36.27	65.65, 69.34, 36.39
Resolution (Å)	19.4–0.97 (0.98–0.97)[a]	50–1.1 (1.16–1.1)
R_{meas} (%)	5.8 (38.1)	6.3 (46.1)
$\langle I/\sigma\ (I)\rangle$	25.6 (3.9)	19.9 (3.2)
Completeness (%)	99.8 (96.9)	98.8 (91.8)
Multiplicity	8.9 (3.6)	6.74 (3.3)
Refinement		
Resolution (Å)	19.4–0.97	
No reflections	98,078	
R_{work}/R_{free} (%)	11.8/13.8	
No atoms		
Protein	2833	
Water	266	
B-factors		
Protein overall	10.5	
Water	23.7	
Root-mean square deviations		
Bond lengths (Å)	0.010	
Bond angles (°)	1.39	
Ramachandran plot (%)		
In core/allowed	100	

[a]The values in parenthesis are for the last resolution shell.

Data deposition

Atomic coordinates and structure factors have been deposited to the Protein Data Bank under accession number 4acj.

RESULTS AND DISCUSSION

Overall structure description

The refined model of the asymmetric unit contains one monomer composed of 167 residues, that is, the 166 residues of the TLDc domain and an additional glycine and 279 water molecules. The glycine residue in N-t is the remaining from the TEV cleavage site. The residues numbering used in the following description corresponds to the numbering of the OXR2 full-length sequence. Clear electron density can be observed for all the residues except the Glu714 side chain. Additional electron density on the carboxylate group of Glu721 can be seen, suggesting a modification of this residue (Supporting Information Fig. S1). We could not however identify this modification. The overall structure of the TLDc domain is globular. The secondary structure was assigned by the PDBsum server.[18] The domain is composed of a β-sandwich formed by two antiparallel β-sheets surrounded by four helices. Each β-sheet is composed of six and four strands. The first sheet is formed by strands S1, S10, S2, S3, S4, and S5, where S10 is placed between S1 and S2 [Fig. 1(A–C)]. Strands S6 to S9 form the second sheet. The two sheets organize as a pseudo-orthogonal β-sandwich and interact with each other only by hydrophobic interactions. The N-t (residues 635:684) is formed by an unstructured region (residues 635:645) followed by helix 1 (H1, residues 644:651) linked by five residues to a single turn of 3_{10}-helix (h2, residues 656:658). Strand 1 (S1, residues 664–668) is intercalated between h2 and H4 (residues; 676:683) by a single turn of α-helix (H3; residues 669:672). H4 is therefore between S1 and S2 (residues 690:696). With the exception of strands S5 and S6, each strand is separated by β-turns and β-hairpins.

As shown in Figure 1, the TLDc surface potential is negative on one side [Fig. 1(E)] but is mainly hydrophobic on the opposite side with only a small electropositive patch [Fig. 1(F)]. A search for TLDc related structure using the DALI server[19] leads only to poor similarities. The best hit (*Z*-score = 4.5; root-mean-square deviation = 3.3) is the structure of the *p*-coumaric decarboxylase from *Lactobacillus plantarum*[20] (PDB: 2W2A). The two structures display indeed a similar central β-sandwich (Fig. S2). However, none of the proposed catalytic residues of the *p*-coumaric decarboxylase[20] are conserved in the TLDc structure (not shown). This excludes that the TLDc domain has similar function.

Biological significance of the TLDc structure

Several studies have shown that the OXR protein family is involved in oxidation resistance. However, none of these studies suggested a direct role of these proteins in ROS processing, and no catalytic activity has been demonstrated for the TLDc domain. Nonetheless, Volkert and coworkers[3,4,7] have shown that the OXR gene expression is upregulated on H_2O_2 treatment. Thus, we have investigated if the TLDc domain may have a role in decomposing H_2O_2, as do catalase or peroxidase enzymes. However, neither the TLDc domain nor the full-length OXR2 protein displays detectable activity (data not shown). Similar observation was recently reported for the OXR1 protein.[6] Alternatively, OXR1 was proposed to act as a scavenger of ROS compounds and Cys704 residue (Cys776 in OXR2) was suggested as an essential residue for this activity.[6] A recent report has also shown, that the missense mutation Asp147His and Ala509Val in the human TBC1D24 protein is responsible for familial infantile myoclonic epilepsy.[21] The TBC1D24 protein modulates the ARF6 protein function, which is involved in the control of dendritic branching and axon extension.[22] The TBC1D24 protein possesses in N-t, a TBC domain and in C-t, a TLDc. The Ala509Val mutation localized in the TLDc is sufficient to deregulate the dendritic branching.[21] In the TLDc structure presented here, the equivalent residue Gly761 positioned at the beginning of strand S8, is in the vicinity of the reactive Cys. The Ala509Val mutation would trigger a steric hindrance with the side chain of Cys that might disturb the reactivity of the Cys residue and therefore its ROS scavenging function (Fig. 2).

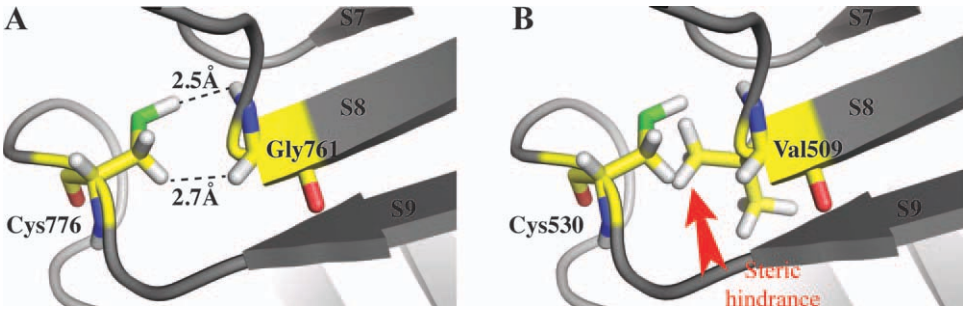

Figure 2

Zoom view of the Cys776 region. **A:** Region of the Cys776 in the crystal structure of the TLDc of OXR2, Cys776, and Gly761 are colored in yellow, the black dashed lines indicate the distance between the two residues. **B:** Model of the mutation in the TLDc reported by Falace *et al.*[21] The presence of the Val509 side chain clashes with the Cys530 ones as indicated by the red arrow. The numbering corresponds to the one of the *Homo sapiens* TBC1D24 protein sequence.

To conclude, recent studies have clearly demonstrated the involvement of the OXR proteins and the TLDc domain in oxidation resistance. The structure of the TLDc does not present close structural similarity to any other protein. Nonetheless, the structure sheds light on how a missense mutation might affect the function of the TLDc domain. Biochemical investigations on the OXR protein family are still needed to characterize further the functions of these proteins and to determine if the TLDc domain possesses other function than ROS scavenger.

ACKNOWLEDGMENTS

The authors thank C. Olesen, R. Kidmose, and the staffs at SLS and MAX-lab for help during data collection and L. Arleth and K. Mortensen for fruitful discussion.

REFERENCES

1. Valko M, Rhodes CJ, Moncol J, Izakovic M, Mazur M. Free radicals, metals and antioxidants in oxidative stress-induced cancer. Chem Biol Interact 2006;160:1–40.
2. Halliwell B. Antioxidant defence mechanisms: from the beginning to the end (of the beginning). Free Radic Res 1999;31:261–272.
3. Volkert MR, Elliott NA, Housman DE. Functional genomics reveals a family of eukaryotic oxidation protection genes. Proc Natl Acad Sci USA 2000;97:14530–14535.
4. Elliott NA, Volkert MR. Stress induction and mitochondrial localization of Oxr1 proteins in yeast and humans. Mol Cell Biol 2004;24:3180–3187.
5. Jaramillo-Gutierrez G, Molina-Cruz A, Kumar S, Barillas-Mury C. The Anopheles gambiae oxidation resistance 1 (OXR1) gene regulates expression of enzymes that detoxify reactive oxygen species. PLoS One 2010;5:e11168.
6. Oliver PL, Finelli MJ, Edwards B, Bitoun E, Butts DL, Becker EB, Cheeseman MT, Davies B, Davies KE. Oxr1 is essential for protection against oxidative stress-induced neurodegeneration. PLoS Genet 2011;7:e1002338.
7. Durand M, Kolpak A, Farrell T, Elliott NA, Shao W, Brown M, Volkert MR. The OXR domain defines a conserved family of eukaryotic oxidation resistance proteins. BMC Cell Biol 2007;8:13.
8. Doerks T, Copley RR, Schultz J, Ponting CP, Bork P. Systematic identification of novel protein domain families associated with nuclear functions. Genome Res 2002;12:47–56.
9. Shao W, Halachmi S, Brown M. ERAP140, a conserved tissue-specific nuclear receptor coactivator. Mol Cell Biol 2002;22:3358–3372.
10. Letunic I, Doerks T, Bork P. SMART 6: recent updates and new developments. Nucleic Acids Res 2009;37:D229–D232.
11. Alsarraf HMAB, Laroche F, Spaink HP, Thirup S, Blaise M. Purification, crystallization and preliminary crystallographic studies of the TLDc domain of oxidation resistance protein 2 from zebrafish. Acta Crystallogr Sect F Struct Biol Cryst Commun 2011;67:1253–1256.
12. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr 2010;66:125–132.
13. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 2010;66:213–221.
14. Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, Hung LW. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr D Biol Crystallogr 2009;65:582–601.
15. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010;66:486–501.
16. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40:658–674.
17. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 2010;66:12–21.
18. Laskowski RA. PDBsum new things. Nucleic Acids Res 2009;37:355–359.
19. Holm L, Rosenström P. Dali server: conservation mapping in 3D. Nucleic Acids Res 2010;38:545–549.
20. Rodríguez H, Angulo I, de Las Rivas B, Campillo N, Páez JA, Muñoz R, Mancheño JM. p-Coumaric acid decarboxylase from *Lactobacillus plantarum*: structural insights into the active site and decarboxylation catalytic mechanism. Proteins 2010;78:1662–1676.
21. Falace A, Filipello F, La Padula V, Vanni N, Madia F, De Pietri Tonelli D, de Falco FA, Striano P, Dagna Bricarelli F, Minetti C, Benfenati F, Fassio A, Zara F. TBC1D24, an ARF6-interacting protein, is mutated in familial infantile myoclonic epilepsy. Am J Hum Genet 2010;87:365–370.
22. Jaworski J. ARF6 in the nervous system. Eur J Cell Biol 2007;86:513–524.