Journal of Structural Biology 212 (2020) 107576

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Human ARMT1 structure and substrate specificity indicates that it is a DUF89 family damage-control phosphatase

Taylor N. Dennis[a], Nikola Kenjić[a], Amrik S. Kang[a], Jonathan D. Lowenson[b], Jay S. Kirkwood[c], Steven G. Clarke[b], J. Jefferson P. Perry[a,d,*]

[a] Department of Biochemistry, University of California, Riverside, Riverside, CA 92521, USA
[b] Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA
[c] Metabolomics Core Facility, Institute for Integrative Genome Biology, University of California, Riverside, Riverside, CA 92521, USA
[d] The School of Biotechnology, Amrita University, Kollam 690525 Kerala, India

ARTICLE INFO

Keywords:
Metabolite repair
DUF89
C6ORF211
Phosphatase
Glycation
ARMT1

ABSTRACT

Metabolite damage control is a critical but poorly defined aspect of cellular biochemistry, which likely involves many of the so far functionally uncharacterized protein domain (domains of unknown function; DUFs). We have determined the crystal structure of the human DUF89 protein product of the *C6ORF211* gene to 1.85 Å. The crystal structure shows that the protein contains a core α-β-α fold with an active site-bound metal ion and α-helical bundle N-terminal cap, which are both conserved features of subfamily III DUF89 domains. The biochemical activities of the human protein are conserved with those of a previously characterized budding yeast homolog, where an *in vitro* phosphatase activity is supported by divalent cations that include Co^{2+}, Ni^{2+}, Mn^{2+} or Mg^{2+}. Full steady-state kinetics parameters of human DUF89 using a standard PNPP phosphatase assay revealed a six times higher catalytic efficiency in presence of Co^{2+} compared to Mg^{2+}. The human enzyme targets a number of phosphate substrates similar to the budding yeast homolog, while it lacks a previously indicated methyltransferase activity. The highest activity on substrate was observed with fructose-1-phosphate, a potent glycating agent, and thus human DUF89 phosphatase activity may also play a role in limiting the buildup of phospho-glycan species and their related damaged metabolites.

A large numbers of domains of unknown function (DUFs) have been revealed through recent genomic sequencing efforts, and in the 31.0 release of the Pfam database they totaled more than 20 percent of the 16,712 known protein domain families (Bateman et al., 2010). Over 800 of the DUFs are conserved between bacteria and eukaryotes, and approximately 300 of these occur in archeal species (Goodacre et al., 2014). This widespread taxonomic distribution is suggestive of important roles in the cell, and therefore defining DUF functions could provide critical new insights into cellular metabolism.

One DUF family likely to have key cellular functions, due to its conservation from specific bacterial species through to higher eukaryotes, is the DUF89 family (Pfam01937). A landmark study on recombinant DUF89 proteins that included specific bacterial, archeal, fungi and eukaryotic sequences revealed a conserved, novel metal-dependent phosphatase activity (Huang et al., 2016). This catalytic activity was defined as targeting reactive and/or damaged phosphoesters (Huang et al., 2016). The DUF89 family was further divided into three subfamilies, namely I, II and III, with subfamily I members occurring in specific anaerobic or microaerophilic bacteria and archaea, and having activities directed towards nucleotides. Subfamily II proteins belong to two eukaryotic subgroups, the first of which is observed in plant species with activity against sugar phosphates. The second subset consists of a DUF89 domain fused onto the C-terminus of a pantothenate kinase (PANK) domain (Tilton et al., 2006; Hortnagel et al., 2003), and these genes are present in certain fungi, and in plants and animals. The human and Arabidopsis PANK-DUF89 protein homologs were demonstrated to target oxidized forms of phosphopantetheine, and the plant protein was also active against 4′-phosphopantothenate (Huang et al., 2016). Subfamily III is the most prevalent of the subfamilies, being present in some bacterial species through to higher eukaryotes, and the few members of this family characterized so far have activity against sugar phosphate substrates (Huang et al., 2016).

The human genome contains two DUF89 containing sequences, the subfamily II DUF89-PANK fusion and a largely uncharacterized product of the *C6ORF211* gene (HGNC designation ARMT1), with sequence similarity to subfamily III members. Interestingly, a recent study

* Corresponding author at: Department of Biochemistry, University of California, Riverside, Riverside, CA 92521, USA.
E-mail address: jeff.perry@ucr.edu (J.J.P. Perry).

https://doi.org/10.1016/j.jsb.2020.107576
Received 1 November 2019; Received in revised form 7 July 2020; Accepted 10 July 2020
Available online 15 July 2020

identified a mutation within the PANK-DUF89 gene (PANK4), which was a novel pathogenic variant for congenital cataracts (Sun et al., 2019), and thereby provides the first evidence of a human DUF89 domain impacting health. Initial studies on *C6ORF211* had noted C6orf211 protein to be co-expressed with the estrogen receptor, and siRNA knockdown of C6orf211 expression was observed to reduce breast cancer cell proliferation independent of estrogen levels (Dunbier et al., 2011). C6orf211 protein has also been indicated to have potential roles in DNA damage responses, and it was suggested to have a potential methyltransferase activity on PCNA in a study completed before the wider DUF89 family was biochemically analyzed (Perry et al., 2015).

Here, we present the first detailed structural and biochemical characterizations of the human *C6or211* DUF89 protein product, revealing a conservation of both structure and biochemical activities from lower to high eukaryotes. The crystal structure of the human protein at 1.85 Å reveals a conserved DUF89 domain α-β-α fold that includes key conserved structural motifs, such as a metal-bound active site and an N-terminal helical bundle cap region, which belong to subfamily III members including a previously characterized yeast homolog. A number of different divalent metal cations were observed to support *in vitro* phosphatase activity against several distinct substrates, with the greatest levels observed against phospho-sugars, and with the highest activity observed against fructose-1-phosphate. Fructose-1-phosphate is known to be potent glycating agent, and this suggests that human DUF89, similar to yeast DUF89, could play a potential role as a damage-control phosphatase limiting the potentially damaging buildup of phospho-glycan species in human cells.

1. Materials and methods

1.1. Protein expression and purification

Human *C6ORF211* D291A mutant in a plasmid, a gift from the Hoelz lab, was cloned into the pET SUMO vector [Invitrogen Champion™], and quick-change PCR was conducted to generate a wild type (WT) construct. WT hDUF89 protein (Uniprot: Q9H993) was expressed in BL21 (DE3) Codon Plus RIL cells, and protein expression at 18 °C for 15 h was induced with 0.4 mM final concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG). The D291A mutant was expressed using the same protocol, in BL21 (DE3) *E. coli* cells. Selenomethionine labelled (SeMet) inactive D291A mutant protein was expressed in B834 (DE3) methionine auxotroph cells (Novagen). Cells were cooled to 18 °C before being supplemented with selenomethionine for 30 min, thereafter, 0.4 mM IPTG was added to induce expression for 15 h. Wild type, D291A mutant, and SeMet hDUF89 proteins were purified using the same protocol: the 6xHis-Smt3-hDUF89 fusion protein was purified by Ni-NTA affinity chromatography (HisTrap FF; GE Health Sciences). The 6xHis-Smt3 tag was subsequently removed by recombinant ubiquitin-like protease 1 (Ulp1) at a ratio of 1 μg Ulp1:1 mg 6xHis-Smt3-hDUF89 for 15 h at 8 °C, followed by ion-exchange (HiTrap Q FF; GE Health Sciences) and size-exclusion (Superdex 200; GE Health Sciences) chromatography steps. Purified hDUF89 protein was concentrated to 13 mg/ml in 20 mM Tris-HCl pH 7.5, 120 mM NaCl, 1 mM β-mercaptoethanol, and 0.1% sodium azide storage buffer.

1.2. Crystallization, data collection and structural refinement

Purified human hDUF89 wild-type, D291A mutant, and SeMet-substituted proteins were crystallized in 18% polyethylene glycol MW 4000, 200 mM $MgCl_2$, and 100 mM Tris-HCl pH 8.5 at 4 °C. The SeMet and D291A crystals were flash-frozen directly, whereas WT crystals used 30% xylitol (w/v) as *cryo*-protection. Data was obtained for both WT and D291A crystals at the SIBYLS synchrotron beamline BL12.3.1, Advanced Light Source, Lawrence Berkeley National Laboratories, Berkeley, CA. SAD-based x-ray diffraction data for the SeMet hDUF89 D291A mutant crystal was collected at BL 9.0.3, Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Stanford, CA. To obtain single-wavelength anomalous signal, a fluorescence scan was conducted and determined the peak wavelength to be 0.979 Å. Diffraction data were indexed, integrated, and scaled using HKL2000 (Otwinowski and Minor, 1997). Phasing information from the SeMet-based data was obtained using the AutoSol extension of PHENIX (Adams et al., 2010), with the single-wavelength anomalous diffraction data and an iTasser (Yang et al., 2015) generated homology model based on the yeast homolog (3PT1). Crystals belonged to the $C222_1$ space group, with two molecules in the asymmetric unit. Model building and refinement of the three crystal structures was completed using PHENIX (Adams et al., 2010), and Coot (Emsley and Cowtan, 2004). The WT hDUF89 crystal structure has been deposited in the PDB with code 6UMQ and the D291A hDUF89 structure was deposited with PDB code 6UMR.

1.3. Methylation assay

Triplicate mixtures containing 2 μg of enzyme (DUF89, DUF89 D291A, or PCMT1) or no enzyme, 10 μg of PCNA, and 20 pmol of S-adenosyl[*methyl*-^{3}H]methionine (PerkinElmer, 81.9 Ci/mmol) in a final volume of 50 μL of 5 mM sodium HEPES, pH 7.5, 50 mM NaCl were incubated for 30 min at 37 °C. ^{3}H-methyl esters formed were hydrolyzed by adding 100 μL of 0.2 M NaOH on ice. After vortexing, 100 μL of the mixture was added to filter paper wedged into the neck of a 20 mL scintillation vial containing 5 mL of scintillation cocktail (Safety-Solve, RPI). Vials were immediately capped and incubated at room temperature for 2 h allowing ^{3}H-methanol formed to diffuse into the scintillation cocktail. After removing the filter paper, the vials were counted in a Beckman LS6500 scintillation counter. A Methylation assay with base-treated PCNA, was also performed, using the same conditions, except that PCNA that had been previously treated in a 20 μg/μl solution for 30 min at 37 °C in 0.1 M NaOH and then neutralized.

1.4. Phosphatase assay

pNPP (para-nitrophenyl phosphate) phosphatase assay monitors removal of phosphate by direct detection of product formation, para-nitrophenol (pNP), with λ_{max} at 405 nm. The assay was performed at RT using a Shimadzu UV-2600 spectrophotometer. Reaction rate have been acquired by monitoring the change at 405 nm after introduction of reaction mixture to 0.4 μM DUF89. Reaction mixture was composed 200 mM HEPES, 120 mM NaCl, 25 mM pNPP, pH 7.5 and substrate. Substrates were prepared by dissolving the $CoCl_2$ or $MgCl_2$ in water to 5 mM and serially diluting this solution to make 2500 μM, 1250 μM, 625 μM, 312.5 μM, 156.2 μM, 78.1 μM, 39 μM, 19.5 μM, 9.8 μM, 4.9 μM and 2.4 μM substrate. A standard curve of pNP concentrations vs absorbance at 405 nm was used to convert the change in absorbance to the amount of product generated. Obtained rates were plotted against substrate concentration and the data was fit to Michaelis-Menten equation. Triplicate measurements were conducted for all analyses, and the final kinetic parameters have been reported as average of these values.

1.5. Phosphatase steady state kinetic assay in presence of pNPP as a varied substrate

This experiment has been done in same manner as indicated above with few changes. Each reaction was initiated by addition of the reaction mixture into cuvette containing 0.4 μM enzyme. Reaction mixture was composed of 200 mM HEPES, 120 mM NaCl, and 5 mM metal ($CoCl_2$ or $MgCl_2$) at pH 7.5. Different substrate concentrations have been made by serial dilution of 30 mM pNPP to make 22.5 mM, 16.9 mM, 12.6 mM, 9.5 mM, 7.1 mM, 5.3 mM, 4 mM, 3 mM, 2.2 mM,

1.7 mM, 1.3 mM and 0.95 mM pNPP. Obtained rates were plotted against the substrate concentration and data was fit to Michaelis-Menten equation. Triplicate measurements were conducted for all analyses and the final kinetic parameters have been reported as average of these values.

1.6. Phosphometabolite activity assay

Using malachite green detection method, hDUF89 protein at 100 nM concentration was added to phosphometabolite substrates at 100 μM concentration in a reaction buffer of 20 mM HEPES pH 7.4, 120 mM NaCl, 5 mM of varying metal ion, 0.5 mM EDTA in a total reaction volume of 100 μL, and was incubated at 37 °C for one hour. The reaction mixture was then diluted to 1000 μL with H_2O, and 250 μL of malachite green reagent was added to quench the reaction. The malachite green reagent solution was composed of 14.95 mM malachite green oxalate, 1.19 μM ammonium molybdate, > 2.3 M H_2SO_4, and 0.17% Tween-20. This quenched solution was then incubated 30 min at room temperature for color development, and absorbance was measured at 620 nm using a UV-2600 Shimadzu spectrophotometer. Results were measured in triplicate and controls were run in the absence of enzyme and/or substrate. A second phospho-metabolite activity assay, using the same malachite green detection conditions and with 5 mM $CoCl_2$, was conducted against an array of phospho-metabolites, and triplicate measurements were conducted for all analyses.

1.7. Liquid chromatography-mass spectrometry (LC-MS) analyses

Separation and detection of fructose-1-phosphate and fructose was performed on a TQ-XS triple quadrupole mass spectrometer (Waters) coupled to an I-class UPLC system (Waters). Separations were carried out on a ZIC-pHILIC column (2.1 × 150 mm, 5 μM) (EMD Millipore). The mobile phases were (A) water with 15 mM ammonium bicarbonate adjusted to pH 9.6 with ammonium hydroxide and (B) acetonitrile. The flow rate was 200 μL/min and the column was held at 50° C. The injection volume was 1 μL. The gradient was as follows: 0 min, 90% B; 1.5 min, 90% B; 16 min, 20% B; 18 min, 20% B; 20 min, 90% B; 28 min, 90% B. The MS was operated in selected reaction monitoring mode. For fructose, the precursor ion was 179 m/z and the product ion was 71 m/z (collision energy (CE) 14 eV). For fructose-1-phosphate, the precursor ion was 259 m/z and the product ion was 97 m/z (CE 15 eV). Source and desolvation temperatures were 150 °C and 500 °C, respectively. Desolvation gas was set to 1000 L/h and cone gas to 150 L/h. Collision gas was set to 0.15 mL/min. All gases were nitrogen except the collision gas, which was argon. Capillary voltage was 2 kV in negative ion mode.

2. Results

2.1. Human DUF89 crystal structure

The human DUF89 (hDUF89) crystal structure was determined to 1.85 Å resolution, with the single-wavelength anomalous diffraction (SAD) method used for phasing with selenomethioine labeled protein crystals. The final model produced had R_{work} of 19.64% and R_{free} of 22.40% (Table 1; Fig. 1**a**). Two molecules of hDUF89 are present in the asymmetric unit, having 180° rotational symmetry along the axis of interaction between the N-lobe of each molecule (Fig. 1**a**). Residues 1–190 and 213–441 are readily observable in the electron density maps of each molecule in the crystallographic asymmetric unit, while residues 191–213 are disordered in both molecules. hDUF89 protein consists of 18 α-helices and 7 β-strands that form two regions within the protein structure, an N-lobe that is composed of an α-helical bundle and this is followed by an α-β-α three-layered sandwich (Fig. 1**b**). The two molecules in the asymmetric unit interact via their N-terminal α-helical lobe (Fig. 1**a**), and this interaction includes the coordination of two metal ions, likely to be magnesium as judged from the electron density maps and the high concentration of this metal in the crystallization solution. The first Mg^{2+} is bound by residues D253, N254 and D291 (Fig. 1**c**) and the second is coordinated in between the two molecules of the asymmetric unit, by water molecules that also form hydrogen bonds to the D86 and E93 side chains of one dimer, and the Q96 side chain and main chain oxygens of L99 and N100 of the other (Fig. 1**d**). Analysis of hDUF89 by gel filtration chromatography indicates that it is a monomer in solution with estimated mass of 53 KDa, verses actual mass of 51 KDa (data not shown), suggesting that the presence of the magnesium ions is promoting packing interactions within the crystal. The largest pocket in the hDUF89 structure contains a metal-ion that is coordinated by the side chains of D253 and N254 of the β1-α10 loop, and the D291 side chain present in the β2-α11 loop, in addition to three water molecules (Fig. 1**c**). The crystal structure of the hDUF89 D291A mutant protein was determined to 2.2 Å, with an R_{work} of 22.06% and an R_{free} of 25.66%, and the alpha carbon backbone overlays with WT hDUF89 at 0.2 Å RMSD, and shows no significant changes within secondary structure elements. The electron density maps revealed that the D291A mutation does not chelate a metal ion, in addition to a slight rotation of the N254 side chain and altered water coordination in this region of the active site (Fig. 1**e**). The residues of the active site of the human protein are structurally conserved with the ScDUF89 homolog (Fig. 1**f**), where key active site residues of ScDUF89 were revealed by its complex with fructose 6-phosphate (F-6-P) substrate.

The core α-β-α sandwich domain of hDUF89 starts at α6 and contains 13 α-helices, all 7 β-strands, and includes a cis-proline at position 132. The central seven-stranded β-sheet has a topology similar to the Rossmann fold, except that strand 0 & 6 are in an anti-parallel

Table 1
Highest resolution shell is shown in parenthesis.

	SeMet	DUF89	DUF89D291A
Data collection			
Space group	$C222_1$	$C222_1$	$C222_1$
a, b, c (Å)	91.05, 195.35, 115.01	90.14, 194.47, 114.20	91.26, 194.69, 114.96
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	26.75–1.98 (2.03–1.98)	66.49–1.85 (1.92–1.85)	99.99–2.21 (2.29–2.21)
R_{merge} (%)	15.5 (63.1)	11.8 (21.5)	13.4 (48.4)
I/σ (I)	15.4 (2.47)	8.4 (5.2)	8.1 (3.4)
Completeness (%)	99.15 (99.72)	98.54 (99.27)	98.66 (97.48)
Redundancy	4.5 (4.5)	6.1 (6.2)	6.0 (6.9)
Refinement			
No. reflections		85,552 (8418)	50,993 (4987)
R_{work}/R_{free} †		0.1964/0.2240	0.2206/0.2566
No. atoms		7471	7342
Protein		6892	6894
Ligand/ion		4	2
Water		575	446
B-factors		16.26	21.99
Protein		15.78	21.93
Ligand/ion		10.39	24.64
Water		22.03	22.95
r.m.s deviations			
Bond lengths (Å)		0.014	0.010
Bond angles (°)		1.22	1.07
Ramachandran favoured (%)		97.81	97.94
allowed (%)		1.95	1.82
outlier (%)		0.24	0.24
Rotamer outlier (%)		0.26	1.19

Rmerge = $\Sigma hkl \Sigma i |Ii(hkl) - \langle I(hkl)\rangle| / \Sigma hkl \Sigma i Ii(hkl)$, where Ii(hkl) is the intensity of the ith observation of reflection hkl and ⟨I(hkl)⟩ is the average over all observations of reflection hkl.

Rfactor = $\Sigma hkl \| Fo| - |Fc\| / \Sigma hkl |Fo|$ for all data excluding the 10% that comprised the Rfree used for cross-validation.

† These values are higher due to the presence of pseudosymmetry and tNCS.

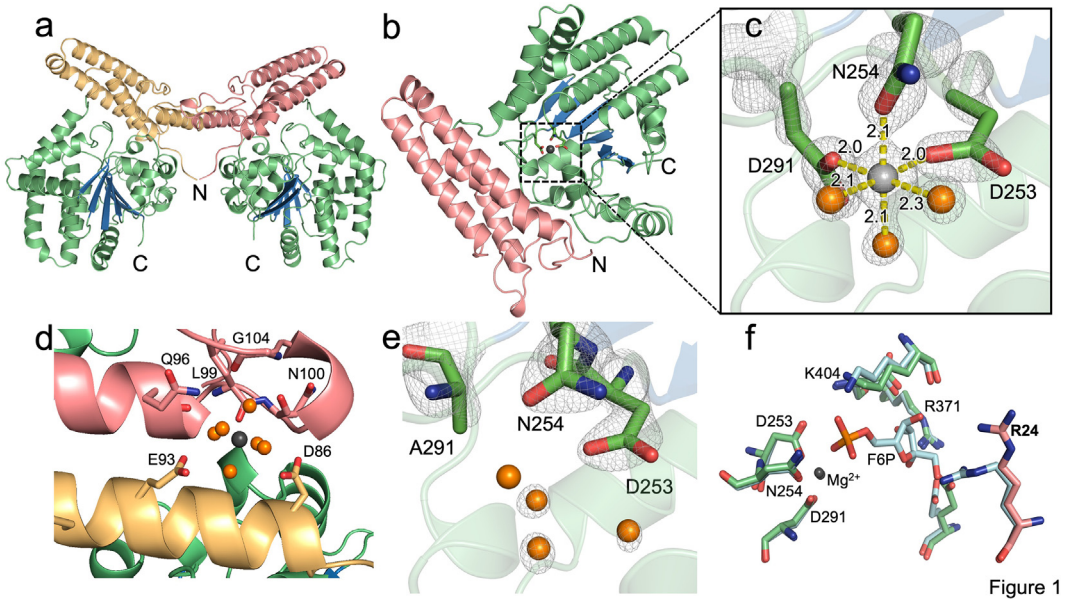


Fig. 1. hDUF89 Crystal structure. (a) Cartoon representation of hDUF89 dimer within the asymmetric unit, with N-terminal lobe region of one unit depicted in salmon, and the other subunit in light orange. (b) hDUF89 monomer with N-terminal lobe region colored in salmon, and the α-β-α sandwich-like catalytic domain with loops and α-helices in green and β-strands in blue. (c) hDUF89 active site, magnesium ion (gray sphere) is coordinated in an octahedral geometry by the side chains of residues D253, N254 and D291, and three water molecules (orange). Hydrogen bonds are colored with yellow dashed lines and distances are in Å, and the omit map contoured at 3.0σ is represented as a wire mesh. (d) Magnesium coordination in the dimer interface is via water molecules, and where these waters form hydrogen bonds to the D86 and E93 side chains of one dimer, and the Q96 side chain and main chain oxygens of L99 and N100 of the other. (e) The active site of the D291A DUF89 mutant, with the omit map contoured at 2.0σ and represented as a wire mesh, revealing loss of metal binding. (f) Structural superimposition of the active site of hDUF89, with carbons atoms colored in green, and ScDUF89 with F6P-bound, with carbon atoms colored in cyan (PDB code 3PT1). The depicted amino acids are key to substrate hydrogen bonding and/or metal-coordination, and are as follows, with the yeast homolog in parenthesis: R24 (R23), E117 (E110), D253 (D254), N254 (N255), D291 (D292), D367 (D384), R371 (R388), K404 (K423). The R24 side chain of hDUF89 is highlighted with salmon colored carbon atoms and it is in a different orientation to the equivalent residue in yeast, R23, and this likely due to the absence of F-6-P substrate in the hDUF89 structure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

orientation in hDUF89. A search for structural homologues using the DALI server (Holm and Laakso, 2016) revealed that the *S. cerevisiae* homolog, ScDUF89 (PDB: 3PT1) is the most similar to hDUF89, with the Cα atoms superimposing within 1.31 Å RMSD, and the two sequences share 31% sequence identity from 405 aligned residues. ScDUF89 belongs to subfamily III of the larger DUF89 family, and members of this subset were previously noted to contain conserved DNxG and RTxK motifs that are also present in hDUF89 (Fig. 2). hDUF89 metal binding residues D253 and N254 belong to this DNxG motif and D291 belongs to a PΦXVSD motif, where Φ represents a bulky hydrophobic residue, which we observe are conserved as F/W/Y in an alignment of divergent DUF89 subfamily III members, from human, budding yeast, *C. elegans*, xenopus (*X. tropicalis*), zebra fish, and the western honey bee (Fig. 2). In the ScDUF89 structure in complex with F-6-P, PDB code 3PT1, the K424 of the side chain of the conserved RTxK motif binds the phosphate group of substrate, with K404 as the corresponding residue in hDUF89. We also observed two further sequence motifs that are well conserved in divergent DUF89 subfamily III members (Fig. 2). The first being hDUF89 residues 361–378, L(I/L/V)(I/L/V)FKGDLNYRKL(V/T)GDR(K/D/N)W, belonging to part of the β-strand 4, the β-4-α-16 loop, and α-helix 16. The N-terminal region of this β-4-α-16 loop forms the base of the active site. The conserved second motif, R(I/W/L)PXI(I/L/V) is the N-terminal cap region, being residues 24–29 in hDUF89 (Fig. 2). hDUF89 R24 of this motif resides in α-helix 1, and this R24 side chain is observed in distinct orientations between the human and yeast homolog structures. In the absence of substrate, the hDUF89 R24 side chain is orientated away from the active site, and in the apo-ScDUF89 structure (PDB code: 5F13) the equivalent residue, R23, is not readily observed in the electron density maps. However, in the presence of F-6-P substrate (PDB code: 3PT1), ScDUF89 R23 side chain rotates 45°, as compared to the hDUF89 structure, to move approximately 4.5 Å towards the substrate and form a hydrogen bond to the hydroxyl group on carbon 1 of the sugar-phosphate substrate (Fig. 1f). Soaking and co-crystallization studies on hDUF89 with F-6-P under these crystallization conditions were not successful, potentially due to this change in conformation of R24.

Crystallographic analyses of hDUF89 reveal that key features are conserved between the monomeric structures of the human and yeast DUF89 subfamily III proteins. These features include a metal binding site, an α-helical bundle N-terminal cap region, and some differences in the topology of central seven-stranded β-sheet as compared to the Rossmann fold. While the function of N-terminal cap remains to be defined, this region is noted to contain an arginine side chain that interacts with F-6-P substrate in yeast (Huang et al., 2016), and this residue is conserved in the hDUF89 cap. Also similar to ScDUF89 is an octahedral coordination of a metal ion in the hDUF89 active site, likely Mg^{2+}, and that hDUF89 metallohydrolyase activity can be supported by one of a number of divalent cations. The catalytic mechanism of the subfamily III members has been proposed to likely follow a haloacid dehalogenase (HAD)-like phosphatase reaction scheme (Huang et al., 2016), and this superfamily contains phosphofructose phosphatases including YqaB that is a bacterial fructose-1-phosphate phosphatase (Kuznetsova et al., 2006).

2.2. Characterization of catalytic activities

DUF89 had been initially indicated to have a methyltransferase activity for modifying glutamyl residues of PCNA (Perry et al., 2015), before a phosphatase activity was described for a DUF89 subfamily III member. We incubated PCNA with hDUF89 or a D291A mutant form that does not chelate an active site metal ion with ^{3}H-methyl-labeled *S*-adenosylmethionine. Methyl ester products were detected as ^{3}H-

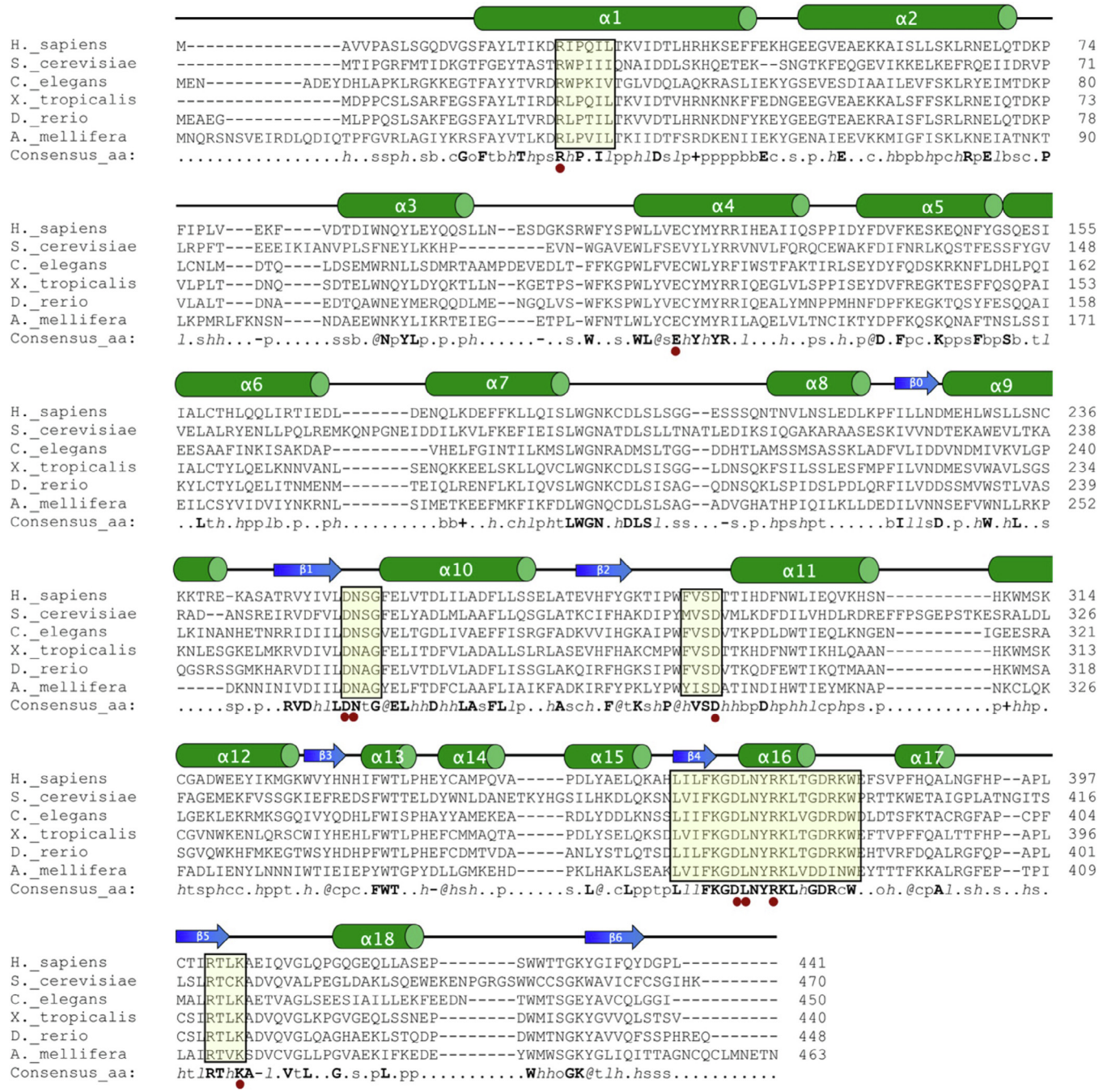


Fig. 2. Sequence alignment of DUF89 homologs. Gene sequences were obtained from GENBANK or SEED using the following ascension codes: *H. sapiens* (fig|9606.3.peg.27930), *S. cerevisiae* (fig|559292.3.peg.4170), *C. elegans* (fig|6239.3.peg.22181), *X. tropicalis* (NP_001004883.1), *D. rerio* (NP_001014353.1) *and A. mellifera* (XP_623407.3). The secondary-structural elements of the human DUF89 structure are depicted above the alignment, α-helices are represented by green cylinders and β-strands are represented by blue arrows. Highly conserved motifs are contained in yellow boxes, and burgundy-colored circles highlight conserved active-site amino acids. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

methanol in a vapor-phase diffusion assay after ester hydrolysis in base (Lowenson et al., 2001). With PCNA as a substrate, we detected no increase in methyl ester formation with wildtype hDUF89 over the mutant form or with a no enzyme control (Fig. 3**a**). To control for the possibility that PCNA was already methyl esterified, we also assayed a base-treated preparation of the protein. Again, we detected no increase in methyl ester formation with wildtype hDUF89 over the mutant protein or the no enzyme control (Fig. 3**b**). For a positive control, we showed that the L-isoaspartyl protein repair methyltransferase PCMT1 did catalyze an increase in methyl ester formation under these conditions, particularly when PCNA was first base treated to enhance isomerization (Fig. 3**a**, **b**). These results indicate that hDUF89 does not catalyze methyl ester formation with PCNA.

The yeast homolog to hDUF89 was previously observed to have a phosphatase activity against phospho-sugars that included F-6-P and fructose-1-phosphate (F-1-P). A classical malachite green-based phosphatase activity was conducted with recombinant hDUF89 and F-6-P substrate, in the presence of distinct divalent metal ions at 5 mM concentration. Phosphatase activity against F-6-P was observed, which enabled the definition of metal ion preference for the enzyme; hDUF89 catalysis was supported to the greatest extent by Co^{2+} followed by Ni^{2+}, Mg^{2+}, and Mn^{2+}, in order of progressively lower levels of

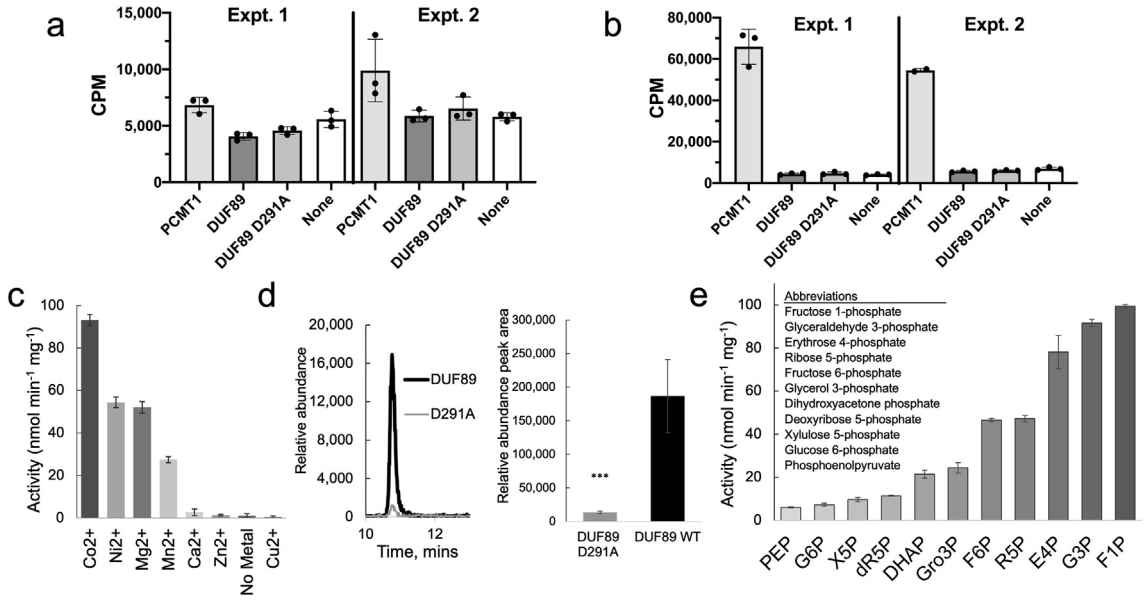


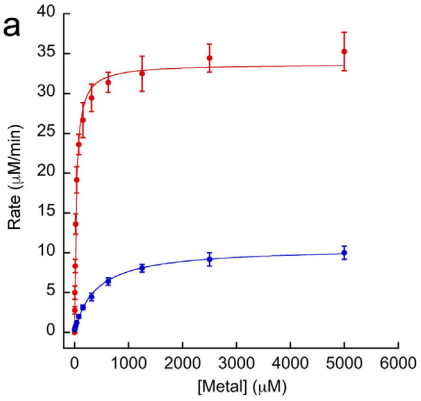
Fig. 3. hDUF89 Activity Assays. (a) Methylation assay with PCNA conducted in triplicate with DUF89, DUF89 D291A, PCMT1 or no enzyme control. Results from two independent experiments are shown, and error bars indicate the standard deviation. (b) Methylation assay with base-treated PCNA, and conducted as in (a). (c) Phosphometabolite activity assay using malachite green detection method, hDUF89 protein was incubated with different bivalent metals, and results measured were in triplicate are presented in a bar graph with units of nmol min^{-1} mg^{-1}. Controls were run in the absence of enzyme and/or substrate. (d) Liquid chromatography-mass spectrometry (LC–MS) analyses, separation and detection of fructose-1-phosphate and fructose. Left panel, representative extracted ion chromatograms showing the presence of fructose, after incubation with WT hDUF89 or hDUF89 D291A mutant, right panel, relative abundance of fructose, after incubation with WT hDUF89 or hDUF89 D291A mutant. (e) Phospho-metabolite activity assay using malachite green detection method and conditions as in (b) and with 5 mM $CoCl_2$, against an array of phospho-metabolites, with the results from the top 11 substrates being depicted, and triplicate measurements were conducted for all analyses.

activity (Fig. 3**c**). This hDUF89 metal-dependent phosphatase activity was largely eliminated by the addition of 5 mM EDTA to the reaction mix, and the D291A hDUF89 mutation lacking a catalytic metal in the crystal structure did not support significant levels of activity (data not shown). F-1-P phosphatase activity was characterized by a mass spectrometry assay, where hDUF89 and the hDUF89 D291A mutant, were each incubated with F-1-P. Following incubation, extraction, and LC-MS/MS analysis, the WT hDUF89 sample contained fructose product and only trace amounts of F-1-P substrate, while the D291A mutant contained near-starting levels of F-1-P substrate and a lack of fructose product (Fig. 3**d**). A malachite green-based screen of phospho-metabolites in the presence of $CoCl_2$ revealed a preference for F-1-P among tested substrates (Fig. 3**e**). Higher activity of hDUF89 with Co^{2+} co-factor over Mg^{2+} was further confirmed by using a steady-state kinetic assay in the presence of pNPP, a scientific standard commonly used in the colorimetric detection of phosphatase activity. Titrating increasing amounts of Co^{2+} and Mg^{2+} at concentrations much higher than that of the protein, we find that the concentration of Co^{2+} to achieve half-maximum velocity is twelve times lower than that of Mg^{2+} (31 ± 4 μM for Co^{2+} versus 396 ± 76 μM for Mg^{2+}) and the enzyme can achieve a three times higher maximum velocity (Fig. 4**a**). Using 5 mM of the metal co-factor, well in excess of the concentration necessary to achieve the half-maximum velocity, we titrated the artificial substrate pNPP and found that the K_M in the presence of Co^{2+} is 3140 ± 265 μM and in presence of Mg^{2+} is 5731 ± 530 μM (Fig. 4**b**). We also find that the catalytic rate in the presence of Co^{2+} (88 ± 5 min^{-1}) is three times higher compared to Mg^{2+} (28 ± 1 min^{-1}). Thus, the presence of Co^{2+} in comparison to Mg^{2+} provides about six times higher catalytic efficiency (k_{cat}/K_M) for the artificial substrate pNPP (Fig. 4**b**). While this analysis helps further confirm the metal-dependent phosphatase activity, we would expect Mg^{2+} to be the more physiologically relevant co-factor.

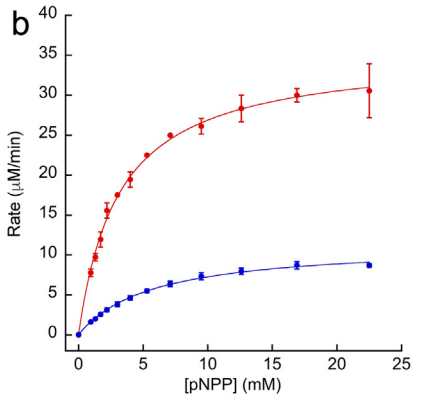
3. Conclusions

Crystallographic analyses of hDUF89 reveal that key features are conserved between the human and yeast DUF89 subfamily III proteins, including a metal binding site, an α-helical bundle N-terminal cap region, and some differences in the topology of central seven-stranded β-sheet as compared to the Rossmann fold. Taken together, these features separate DUF89 subfamily III proteins from more remote structural neighbors that include methyltransferases (Perry et al., 2015). Catalytic activity is also conserved between the human and yeast DUF89 proteins, with a broad selection of phospho-sugars being substrates, which belong to pathways that include the pentose phosphate pathway and glycolysis. Highest activity was also noted on F-1-P for hDUF89, and in humans, liver aldolase B is known to act on F-1-P, but hDUF89 is structurally dissimilar to this enzyme and it lacks the aldolase's activity on fructose 1,6-diphosphate (Dalby et al., 2001), indicating distinct cellular roles.

A damage pre-emption function in metabolite control for DUF89 subfamily III proteins has been suggested (Huang et al., 2016) due to the DNA damage response phenotypes in human (Perry et al., 2015) and budding yeast cells (Gasch et al., 2001), and because F-1-P is a non-canonical yeast metabolite that is toxic at high levels. It may be of interest to further define hDUF89 *in cellulo* roles that could additionally include directed overflow and/or preventing the trapping of phosphate on sugar substrates (Galperin et al., 2006; Linster et al., 2013; Sun et al., 2017; Beaudoin and Hanson, 2016; Bommer et al., 2019; Reaves et al., 2013), and thereby allow for high flux through key metabolic pathways such as through the glycolytic pathway (Bommer et al., 2019; Collard et al., 2016). hDUF89 and yeast DUF89 activities on phospho-fructose species and on erythrose 4-phosphate and ribose-5-phoshpate, may also pre-empt of unwanted glycation events. Notably, fructose and fructose phosphates are relatively potent glycating agents, as compared to glucose and glucose phosphates (Levi and Werman, 2003). Also, glycation is the first step in generating an advanced glycation product (AGE) (Fu et al., 1994; Bucala and Cerami, 1992), where AGEs have been strongly



	Co^{2+}	Mg^{2+}
V_{max} (μM/min)	34 ± 2	11 ± 1
$[M^{2+}]$ at ½ V_{max} (μM)	31 ± 4	396 ± 76
k_{cat} (min^{-1})	85 ± 4	27 ± 3



	Co^{2+}	Mg^{2+}
V_{max} (μM/min)	35 ± 2	11 ± 1
K_M (μM)	3140 ± 265	5731 ± 530
k_{cat} (min^{-1})	88 ± 5	28 ± 1

Fig. 4. Steady-state kinetics of hDUF89 in presence of Co^{2+} (red) and Mg^{2+} (blue) using a pNPP phosphatase assay. (a) Concentration of divalent metal ion to achieve half-maximum velocity ($[M^{2+}]$ at $1/2V_{max}$). (b) Kinetic parameters determined by using pNPP as a varied substrate. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

implicated in aging pathophysiology and many age-related diseases (Ajith and Vinodkumar, 2016; Singh et al., 2001; Yan et al., 2008). Thus, molecular-based definitions of DUF proteins are likely to provide unique biochemical insights to greatly impact our understanding of cellular metabolism and potentially of human health.

CRediT authorship contribution statement

Taylor N. Dennis: Investigation, Visualization. **Nikola Kenjić:** Investigation, Visualization, Writing - review & editing. **Amrik S. Kang:** Investigation. **Jonathan D. Lowenson:** Investigation, Visualization. **Jay S. Kirkwood:** Investigation, Formal analysis. **Steven G. Clarke:** Supervision, Writing - review & editing, Resources. **J. Jefferson P. Perry:** Conceptualization, Writing - original draft, Funding acquisition, Resources, Project administration, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Dr. Anna Travesa Centrich and Jennifer Jossart for critical reading of the manuscript. We gratefully acknowledge the support of the SSRL and ALS beamline staff for crystallographic data collection support. This work was supported in part by DOD CDMRP grant W81XWH-19-1-0327 and a UCR Metabolomics Seed grant to JJPP, and a grant from the National Science Foundation (MCB-1714569) to SGC. The work was performed at the University of California, Riverside and at the University of California, Los Angeles.

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66 (2), 213–221 PubMed PMID: 20124702; PMCID: PMC2815670.

Ajith, T.A., Vinodkumar, P., 2016. Advanced glycation end products: association with the pathogenesis of diseases and the current therapeutic advances. Curr. Clin. Pharmacol. 11 (2), 118–127 PubMed PMID: 27164965.

Bateman, A., Coggill, P., Finn, R.D., 2010. DUFs: families in search of function. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66 (Pt 10), 1148–1152 PubMed PMID: 20944204; PMCID: PMC2954198.

Beaudoin, G.A., Hanson, A.D., 2016. A guardian angel phosphatase for mainline carbon metabolism. Trends Biochem. Sci. 41 (11), 893–894 PubMed PMID: 27544441.

Bommer, G.T., Van Schaftingen, E., Veiga-da-Cunha, M., 2019. Metabolite repair enzymes control metabolic damage in glycolysis. Trends Biochem. Sci PubMed PMID: 31473074.

Bucala, R., Cerami, A., 1992. Advanced glycosylation: chemistry, biology, and implications for diabetes and aging. Adv. Pharmacol. 23, 1–34 PubMed PMID: 1540533.

Collard, F., Baldin, F., Gerin, I., Bolsee, J., Noel, G., Graff, J., Veiga-da-Cunha, M., Stroobant, V., Vertommen, D., Houddane, A., Rider, M.H., Linster, C.L., Van Schaftingen, E., Bommer, G.T., 2016. A conserved phosphatase destroys toxic glycolytic side products in mammals and yeast. Nat. Chem. Biol. 12 (8), 601–607 2104. PubMed PMID: 27294321.

Dalby, A.R., Tolan, D.R., Littlechild, J.A., 2001. The structure of human liver fructose-1,6-bisphosphate aldolase. Acta Crystallogr. D Biol. Crystallogr. 57 (Pt 11), 1526–1533 PubMed PMID: 11679716.

Dunbier, A.K., Anderson, H., Ghazoui, Z., Lopez-Knowles, E., Pancholi, S., Ribas, R., Drury, S., Sidhu, K., Leary, A., Martin, L.A., Dowsett, M., 2011. ESR1 is co-expressed with closely adjacent uncharacterised genes spanning a breast cancer susceptibility locus at 6q25.1. PLoS Genet. 7 (4) e1001382. PubMed PMID: 21552322; PMCID: PMC3084198 as an expert witness for them.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60 (Pt 12 Pt 1), 2126–2132 PubMed PMID: 15572765.

Fu, M.X., Wells-Knecht, K.J., Blackledge, J.A., Lyons, T.J., Thorpe, S.R., Baynes, J.W., 1994. Glycation, glycoxidation, and cross-linking of collagen by glucose. Kinetics, mechanisms, and inhibition of late stages of the Maillard reaction. Diabetes 43 (5), 676–683 PubMed PMID: 8168645.

Galperin, M.Y., Moroz, O.V., Wilson, K.S., Murzin, A.G., 2006. House cleaning, a part of good housekeeping. Mol. Microbiol. 59 (1), 5–19 PubMed PMID: 16359314.

Gasch, A.P., Huang, M., Metzner, S., Botstein, D., Elledge, S.J., Brown, P.O., Walter, P., 2001. Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR Homolog Mec1p. MBoC 12 (10), 2987–3003 PubMed PMID: 11598186; PMCID: PMC60150.

Goodacre, N.F., Gerloff, D.L., Uetz, P., 2014. Protein domains of unknown function are essential in bacteria. mBio 5 (1) e00744-13. PubMed PMID: 24381303; PMCID: PMC3884060.

Holm, L., Laakso, L.M., 2016. Dali server update. Nucl. Acids Res. 44 (W1), W351–W355 PubMed PMID: 27131377; PMCID: PMC4987910.

Hortnagel, K., Prokisch, H., Meitinger, T., 2003. An isoform of hPANK2, deficient in pantothenate kinase-associated neurodegeneration, localizes to mitochondria. Hum. Mol. Genet. 12 (3), 321–327 PubMed PMID: 12554685.

Huang, L., Khusnutdinova, A., Nocek, B., Brown, G., Xu, X., Cui, H., Petit, P., Flick, R., Zallot, R., Balmant, K., Ziemak, M.J., Shanklin, J., de Crécy-Lagard, V., Fiehn, O., Gregory III, J.F., Joachimiak, A., Savchenko, A., Yakunin, A.F., Hanson, A.D., 2016. A family of metal-dependent phosphatases implicated in metabolite damage-control. Nat. Chem. Biol. 12 (8), 621–627 PubMed PMID: 27322068; PMCID: PMC7001580.

Kuznetsova, E., Proudfoot, M., Gonzalez, C.F., Brown, G., Omelchenko, M.V., Borozan, I., Carmel, L., Wolf, Y.I., Mori, H., Savchenko, A.V., Arrowsmith, C.H., Koonin, E.V., Edwards, A.M., Yakunin, A.F., 2006. Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. J. Biol. Chem. 281 (47), 36149–36161 PubMed PMID: 16990279.

Levi, B., Werman, M.J., 2003. Fructose and related phosphate derivatives impose DNA damage and apoptosis in L5178Y mouse lymphoma cells. J. Nutr. Biochem. 14 (1), 49–60 PubMed PMID: 12559477.

Linster, C.L., Van Schaftingen, E., Hanson, A.D., 2013. Metabolite damage and its repair or pre-emption. Nat. Chem. Biol. 9 (2), 72–80 PubMed PMID: 23334546.

Lowenson, J.D., Kim, E., Young, S.G., Clarke, S., 2001. Limited accumulation of damaged proteins in l-isoaspartyl (D-aspartyl) O-methyltransferase-deficient mice. J. Biol. Chem. 276 (23), 20695–20702 PubMed PMID: 11279164.

Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326 PubMed PMID: 27754618.

Perry, J.J., Ballard, G., Albert, A., Dobrolecki, L., Malkas, L., Hoelz, D., 2015. Human C6orf211 Encodes Armt1, a protein carboxyl methyltransferase that targets PCNA and is linked to the DNA damage response. Cell Rep. 10 (8), 1288–1296 PubMed PMID: 25732820; PMCID: PMC4350021.

Reaves, M.L., Young, B.D., Hosios, A.M., Xu, Y.F., Rabinowitz, J.D., 2013. Pyrimidine homeostasis is accomplished by directed overflow metabolism. Nature 500 (7461), 237–241 PubMed PMID: 23903661; PMCID: PMC4470420.

Singh, R., Barden, A., Mori, T., Beilin, L., 2001. Advanced glycation end-products: a review. Diabetologia 44 (2), 129–146 PubMed PMID: 11270668.

Sun, M., Chen, C., Hou, S., Li, X., Wang, H., Zhou, J., Chen, X., Liu, P., Kijlstra, A., Lin, S., Ye, J., 2019. A novel mutation of PANK4 causes autosomal dominant congenital posterior cataract. Hum. Mutat. 40 (4), 380–391 PubMed PMID: 30585370.

Sun, J., Jeffryes, J.G., Henry, C.S., Bruner, S.D., Hanson, A.D., 2017. Metabolite damage and repair in metabolic engineering design. Metab Eng. 44, 150–159 PubMed PMID: 29030275.

Tilton, G.B., Wedemeyer, W.J., Browse, J., Ohlrogge, J., 2006. Plant coenzyme A biosynthesis: characterization of two pantothenate kinases from Arabidopsis. Plant Mol. Biol. 61 (4–5), 629–642 PubMed PMID: 16897480.

Yan, S.F., Ramasamy, R., Schmidt, A.M., 2008. Mechanisms of disease: advanced glycation end-products and their receptor in inflammation and diabetes complications. Nat. Clin. Pract. Endocrinol. Metab. 4 (5), 285–293 PubMed PMID: 18332897.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2015. The I-TASSER Suite: protein structure and function prediction. Nat Methods 12 (1), 7–8 PubMed PMID: 25549265; PMCID: PMC4428668.