Biophysical characterization of the human K0513 protein

Ndivhuwo Nemukondeni, Data curation, Formal analysis, Methodology, Investigation, Visualization, Writing – original draft¹, Afolake Arowolo, Data curation, Formal analysis, Methodology, Resources, Visualization, Writing – original draft, Writing – review & editing², Addmore Shonhai, Formal analysis, Funding acquisition, Resources, Visualization, Writing – original draft, Writing – review & editing¹, Tawanda Zininga, Conceptualization, Formal analysis, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing³, and Adelle Burger, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing^{1*}

¹Department of Biochemistry, School of Mathematical & Natural Sciences, University of Venda, Private Bag X5050, Thohoyandou, 0950, South Africa

²Department of medicine, University of Cape Town, Observatory, 7925, South Africa

³Department of Biochemistry, Stellenbosch University, Stellenbosch, 7602, South Africa

*Corresponding author

Email: adelle.burger@univen.ac.za / adelle.burger@gmail.com (AB)

ABSTRACT

Glioblastoma multiforme (GBM) is an aggressive grade IV primary malignant tumour which accounts for 78 % of all brain tumours. *K0513* is a GBM biomarker that is upregulated in the invasive phenotype. *K0513* is expressed ubiquitously and is reportedly enriched in the cerebral cortex of the brain. *K0513* is further implicated in signalling pathways involving neuroplasticity, cytoskeletal regulation and apoptosis. The protein encoded by *K0513* is a prospective biomarker for pancreatic cancer prognosis. However, the gene product of *K0153* is not well characterised. This study focused on structure-function characterisation of human *K0513* protein. To this end, we employed bioinformatics analysis and biophysical approaches to characterize the protein. *In silico* structural characterisation of the human *K0513* protein suggests the presence of a SET binding factor 2 (SBF2) domain and a transmembrane region. The SBF2 domain is found in the Myotubularin-related protein 13 (MTMR13), which may function as a nucleotide exchange factor for the RAS-associated GTPase, Rab28. *K0513* was predicted to interact with RAS-associated GTPase, Rab3a. Secondary structure prediction revealed *K0513* to be predominantly α -helical in nature. The predicted three-dimensional model of *K0513* showed a globular fold, suggesting that the protein is water-soluble. *K0513* was heterologously expressed in *E. coli* XL1-Blue cells and subsequent purification yielded 80 % soluble protein. Biophysical characterisation using tryptophan-based fluorescence spectroscopy and limited proteolysis showed the conformation of *K0513* is mostly unperturbed in the presence of nucleotides. Interestingly, *K0513* was detected in lung carcinoma, fibrosarcoma and cervical adenocarcinoma cells, supporting its possible role in cancer signalling pathways.

KEYWORDS

K0513; recombinant protein; SBF2/MTMR13; GTP binding; GEF; Rab3a

1. INTRODUCTION

Glioblastoma multiforme (GBM) is one of the most lethal adult brain tumours. GBM is a fast-growing and nonmetastatic cancer that may invade tissues of the brain and spine. GBM patients have a very poor prognosis with a median survival of ~10 months [1], a 5,1 % survival rate over five years, and a 37,2 % survival rate over a single year [2]. GBM cells are highly complex and heterogeneous, and display diverse gene expression profiles. Various diagnostic and therapeutic glioblastoma biomarkers are identified by gene expression profiling. Extensive alterations of certain genes amongst the various types of glioblastomas are known to disrupt key signalling pathways [3, 4]. One such gene is the novel *K0513*. *K0513* was reported to be upregulated in an invasive form of GBM (iGBM) [5]. *K0513* is ubiquitously expressed and enriched in the cerebral region of the brain [6], and an upregulation of this gene is correlated with poor disease prognosis in iGBM and pancreatic cancer [5,7]. In addition, an altered expression of *K0513*, due to atypical methylation patterns, is associated with non-survival of neuroblastoma patients [8]. The Human Protein Atlas (Pathology Atlas) reports that the K0513 protein is expressed in 20 different cancers [9]. Although K0513 is not well characterized, it is suggested that K0513 is associated with tumour progression [10]. K0513 is also reported to interact with proteins involved in signalling pathways of neuroplasticity, apoptosis and cytoskeletal regulation, such as kidney and brain expressed protein (KIBRA) or WW domain-containing protein 1 (WWC1); Hematopoietic Cell-Specific Lyn Substrate 1 (HCLS1)-associated protein X-1 (HAX-1), and integrator complex subunit 4 (INTS4), respectively [6]. K0513 seems to be associated with aggressive cancer forms including iGBM, pancreatic cancer and neuroblastoma, and because of this, its interactome would be important as an indicator of its functional pathway in cancer progression. Thus, there is a need to characterize K0513. Based on structure-function features, the association of K0513 with aggressive cancers could serve as a direct or indirect biomarker of iGBM and other invasive cancers and hence, aid in the design or repurposing of novel or existing drugs respectively. The paucity of

information on the structure-function features of this apparently important molecule makes its characterization a priority.

2. MATERIALS AND METHODS

2.1. Materials

Unless otherwise stated, the materials used in this study were procured from Merck KGaA and Sigma-Aldrich (Darmstadt, Germany), Melford Laboratories (Suffolk, UK), and VWR International (Pennsylvania, USA). Mouse raised monoclonal anti-polyhistidine (His) antibodies, and goat anti-mouse IgG horseradish peroxidase (HRP) conjugated antibodies were purchased from ThermoFisher Scientific (Massachusetts, USA). Rabbit raised polyclonal anti-K0513 antibodies, and goat anti-rabbit IgG HRP conjugated antibodies were purchased from Abcam (Cambridge, UK; AB121430).

2.2. Bioinformatics analysis of human K0513

Amino acid sequences of K0513 (K0513a: NCBI accession number: NP_001273494.1), and isoforms (K0513b: NCBI accession number: NP_001273495.1; K0513c: NCBI accession number: PN_001273695.1) and homologues from *Pan troglodytes* (PtK0513, NCBI accession number: XP_003952958.1), *Xenopus tropicalis* (XtK0513, NCBI accession number: NP_001096454.1), *Canis lupus* (ClK0513, NCBI accession number: XP_005620695.1), *Bos taurus* (BtK0513, NCBI accession number: XP_005218484.1), *Mus musculus* (MmK0513, NCBI accession number: NP_001157232.1), *Rattus norvegicus* (RnK0513, NCBI accession number: NP_001100906.1), *Gallus gallus* (GgK0513, NCBI accession number: XP_001235240.2), *Danio rerio* (DrK0513, NCBI accession number: NP_001002193.1), and *Macaca mulatta* (MamK0513, NCBI accession number: XP_001113080.1), and 40 orthologues (NCBI accession numbers shown in Fig 1) were retrieved from the National Centre for Biotechnology Information [11]. Phylogenetic tree analysis was used to establish orthology and a multiple sequence alignment was generated using the COBALT Multiple

Alignment Tool and Neighbour Joining analysis was performed. The rooted tree was generated using a maximum sequence difference of 0,7, and a Kimura protein distance model. To the best of our knowledge, none of the homologues nor orthologues have been characterized. The multiple sequence alignment of K0513 and homologues was generated using Clustal Omega [12] and shaded using Boxshade. K0513 protein secondary structure was predicted by generating a multiple sequence alignment using PROMALS3D [13]. The analysis was confirmed using PHYRE² [14], which predicts the percentage composition of alpha-helices, beta-sheets, and loops. Furthermore, the hydrophobicity and amphipathicity of K0513 was predicted using Biotools Web-based Hydropathy, Amphipathicity and Topology (WHAT) [15] to predict whether the protein makes contact with a membrane. The amino acid composition of K0513 was computed using PEPSTAT module integrated in the EMBOSS software [16]. Physicochemical prediction of the protein was conducted using the Expasy's ProtParam server [17]. A three-dimensional structure of K0513 was generated using PHYRE² and the resulting PDB file was visualized using Biovia Discovery studio [18]. An online Integrated Interaction Database (IID) was launched [19] using the K0513 sequence as identifier (UNIPROT ID; O60268) to generate the predicted interactome.

2.3. Construction of human K0513 expression vector

The complementary DNA (cDNA) sequence coding for the human K0513 isoform a protein was codon harmonised using the OptimumGeneTM algorithm for bacterial system expression. This cDNA was subsequently amplified by PCR using primers that introduced *Nco*I and *Bgl*II restriction sites at 5' and 3' end of the amplified gene respectively . The 1241 bp DNA fragment comprising *K0513* was subcloned into the expression vector pQE60 between the *Nco*I and *Bgl*II restriction sites to generate a C-terminally 6xHis tagged protein-. The pQE60-K0513 construct was synthesized by GenScript (Piscataway, USA). The C-terminal His-tagged pQE60-K0513 construct was verified by sequencing and restriction digestion analysis.

2.4. Overexpression and purification of recombinant K0513 protein

The generated plasmid construct pQE60-K0513 coding for the expression of K0513 protein was transformed into chemically competent *E. coli* XL1-Blue cells (a kind gift from Prof. A. Shonhai). Recombinant K0513 protein was expressed and purified using a previously described protocol [20]. The purification of the 6 x Histidine C-terminally tagged recombinant K0513 protein was conducted using nickel-affinity chromatography under native conditions. The recombinant protein was dialyzed overnight using an Amicon Ultra 15 mL spin dialysis column 10 000 MWCO (Merck, Germany). Purified protein was quantified using Bradford's assay. Protein expression and purification was evaluated by SDS-PAGE analysis. The presence of recombinant K0513 was confirmed by Western blotting using mouse raised monoclonal anti-His antibodies (1 : 5000 dilution), and goat raised horseradish peroxidase (HRP) conjugated anti-mouse IgG antibodies (1 : 5000 dilution) [ThermoFisher Scientific]. The presence of K0513 was validated using polyclonal peptide specific anti-K0513 antibodies (1 : 500 dilution) [Abcam, UK], and HRP-conjugated goat anti-rabbit IgG secondary antibodies (1 : 3000 dilution) [ThermoFisher Scientific]. Peptide specific antibody against K0513, with N-terminal peptide segment from residue 7 - 123, (VGSLIDFGPEAPTSSPLEAPPPVLQDGDGSLGDGASESETTESADSENDMGESPSHPSW DQDRRSSNESFSSNQSTESTQDEETLALRDFMRGYVEKIFSGGEDLDQEEKAKFGGEY), was purchased from Abcam (AB121430, Abcam, Cambridge, United Kingdom). The presence of the secondary antibody was confirmed using the PierceTM ECL Plus Western blotting substrate (ThermoScientific, USA) and viewed using the ChemiDoc (BioRad, USA).

2.5. Tertiary structure analysis of recombinant K0513 protein

2.5.1. Tryptophan-based fluorescence spectroscopy

The tertiary structural conformation of recombinant K0513 protein was analysed using tryptophan-based fluorescence spectroscopy as previously described [21]. Briefly, varying concentrations (0,1 - 0,5 μ M) of recombinant K0513 protein was suspended in Tris-Buffered

Saline (TBS: 50 mM Tris-HCl [pH 7,6], 150 mM NaCl) and incubated for 30 minutes at 20 °C. Fluorescence spectra were monitored using the JASCO-FP-850 Spectrofluorometer (Jasco, UK). The fluorescence emission spectra were monitored after initial excitation at 295 nm between 280 nm to 500 nm, at a scan speed of 1000 nm/min on a 10 nm emission bandwidth. The spectral measurements were averaged for three accumulations after baseline subtraction of buffer without protein to remove background emissions. In order to investigate the effect of nucleotides on the tertiary structure of K0513, the assay was repeated in the presence of different concentrations of nucleotides (0,2 – 1,0 mM) of GDP/GTP or (1,0 – 5,0 mM) ATP/ADP.

2.5.2. Limited proteolysis

The conformational changes of K0513 that was expressed in *E. coli* XL1-Blue cells in the absence and presence of nucleotides using limited proteolysis was investigated. The assay was conducted as previously described [21]. Purified K0513 (6,42 µM) was incubated with 3,21 nM trypsin, using an enzyme to substrate ratio of 1:2000, at 37°C for 160 minutes, in the absence or of 5 mM ATP/ADP or 1 mM GTP/GDP. Proteolytic digestion of K0513 was analysed using 12 % SDS-PAGE.

2.7. Cell culture and Western blotting

All cell lines except the Human retinal pigment epithelial (RPE-1) were cultured and maintained at 37 °C in a humidified atmosphere of 5 % CO₂ in air in Dulbecco's Modified Eagle's Medium (DMEM) + GlutMAXTM-1 containing 10 % Fetal calf serum (FCS) and 1% Penicillin-Streptomycin (10 000 U/ml) (Gibco Life Technologies, Grand Island, NY, USA). The RPE-1 cell line was cultured with 50 %DMEM/50% F10 Nutrient Medium supplemented with 10 % FBS, 0.01 mg/ml Hygromycin B and 1 % Penicillin-Streptomycin. Whole cell lysates from all cell lines were prepared using a modified method described by Holden and Horton, 2009 with enhance radioimmunoprecipitation buffer [E-RIPA, 50 mM HEPES pH 7.4, 150 mM NaCl, 0.5 % Sodium deoxycholate, 1 % Sodium dodecylsulfate

(SDS), 100 mM Dithiothreitol (DTT), 1U/ml Benzonase (Sigma), 1 mM EDTA, 1 mM PMSF and 0.1 % protease inhibitor cocktail (Roche Applied Science, Germany)]. The nutrient medium on cells was aspirated, washed twice with ice-cold phosphate buffered saline (PBS) pH 7.4 and lysed in E-RIPA buffer directly on the cell culture plate for 1 h at 4 °C [22]. The total protein content of each cell lysates was estimated using the BCA protein quantification kit (Pierce, ThermoScientific, USA). An equal amount of protein concentration was resolved on an SDS-PAGE in triplicates per cell line and electroblotted onto a nitrocellulose membrane. Immunoprobings of the blots were performed using anti-K0513 (AB121430, Abcam, Cambridge, United Kingdom) and GAPDH (endogenous control) antibody (14C10, Cell signaling, Danvers, USA). Chemiluminescence signals were visualized with the Clarity Max™ Western ECL substrate (Bio-Rad, Italy) on the Azure c400 Geldoc system (Azure Biosystems, Inc, Dublin, USA).

3. RESULTS

3.1. K0513 is predicted to contain an SBF2 domain and a transmembrane region

Based on bioinformatics predictions, there are four variant transcripts of gene *K0513*, (variants 1 – 4; S1 Table) and two predicted variants (variants X1 and X2). *K0513* variants 1 and 2 code for the longer K0513 protein isoform a [K0513a; variant 2 differs in the 5' untranslated region (UTR) compared to variant 1]. *K0513* variant 3 codes for isoform b, which is shorter but has the same N- and C-termini as isoform a (K0513b; variant 3 differs in the 5' UTR and lacks an alternate in-frame exon compared to variant 1). *K0513* variant 4 codes for isoform c, which is shorter than isoform a and has a distinct C-terminus (K0513c; variant 4 differs in both the 5' and 3' UTRs and has multiple coding region differences, compared to variant 1). The human K0513a protein sequence is well conserved, and is found in eight homologues and 320 orthologues, based on data from NCBI. Phylogenetic tree analysis was used to establish orthology between the human K0513a and K0513 from eukaryotes including primates, rodents, ungulates, bony fishes and bats, and to compare their phylogenetic relationships (Fig

1). To the best of our knowledge, none of the K0513 protein homologues nor orthologues have been characterized. Not surprisingly, human K0513 clustered together phylogenetically with K0513 from primates, forming a monophyletic clade with chimpanzee K0513 (*Pan troglodytes*). Human K0513a was also found to cluster closely with K0513 from rodents and placentals (Fig 1).

Fig 1. Phylogenetic analysis of homologues and orthologues of human K0513.

K0513 amino acid accession numbers were obtained from NCBI. *Callorhinchus milii* K0513 (NCBI accession number: XP_007906114.1); *Rhincodon typus* K0513 (NCBI accession number: XP_020373474.1); *Latimeria chalumnae* (NCBI accession number: XP_005989186.1); *Clupea harengus* (NCBI accession number: XP_012688017.2); *Danio rerio* (NCBI accession number: NP_001002193.1); *Carassius auratus* (NCBI accession number: XP_026143180.1); *Ictalurus punctatus* (NCBI accession number: XP_017321347.1); *Astyanax mexicanus* (NCBI accession number: XP_007249268.2); *Oncorhynchus mykiss* (NCBI accession number: XP_021452347.1); *Salmo trutta* (NCBI accession number: XP_029552093.1); *Hippocampus comes* (NCBI accession number: XP_019733774.1); *Oryzias melastigma* (NCBI accession number: XP_024138114.1); *Betta splendens* (NCBI accession number: XP_029008500.1); *Lates calcarifer* (NCBI accession number: XP_018559566.1); *Nothobranchius furzeri* (NCBI accession number: XP_015820647.1); *Xenopus tropicalis* (NCBI accession number: NP_001096454.1); *Microcaecilia unicolor* (NCBI accession number: XP_030060415.1); *Apteryx rowi* (NCBI accession number: XP_025947786.1); *Aythya fuligula* (NCBI accession number: XP_032051570.1); *Gallus gallus* (NCBI accession number: XP_001235240.2); *Lonchura striata domestica* (NCBI accession number: XP_021393624.1); *Chelonoidis abingdonii* (NCBI accession number: XP_032649654.1); *Pogona vitticeps* (NCBI accession number: XP_020644015.1); *Tupaia chinensis* (NCBI accession number: XP_006152622.1); *Nannospalax galili* (NCBI accession number: XP_008826124.1); *Mesocricetus auratus* (NCBI accession number:

XP_021084646.1); *Mus musculus* (NCBI accession number: NP_758490.3); *Rattus*
norvegicus (NCBI accession number: NP_001100906.1); *Galeopterus variegatus* (NCBI
accession number: XP_008575913.1); *Callithrix jacchus* (NCBI accession number:
XP_017822472.1); *Macaca mulatta* (NCBI accession number: XP_014982200.1); *Macaca*
fascicularis (NCBI accession number: XP_005592743.1); *Homo sapiens* (NCBI accession
number: NP_001273495.1); *Pan troglodytes* (NCBI accession number: XP_009429617.1);
Camelus ferus (NCBI accession number: XP_032320364.1); *Capra hircus* (NCBI accession
number: XP_017917293.1); *Bos taurus* (NCBI accession number: XP_005218483.1); *Bos*
mutus (NCBI accession number: XP_005892521.1); *Equus caballus* (NCBI accession
number: XP_023493199.1); *Acinonyx jubatus* (NCBI accession number: XP_026932127.1);
Mustela putorius furo (NCBI accession number: XP_004753933.1); *Ursus maritimus* (NCBI
accession number: XP_008700908.1); *Canis lupus familiaris* (NCBI accession number:
XP_546795.30); *Canis lupus dingo* (NCBI accession number: XP_025282809.1);
Heterocephalus glaber (NCBI accession number: XP_021112276.1); *Pteropus vampyrus*
(NCBI accession number: XP_011368572.1); *Hipposideros armiger* (NCBI accession number:
XP_019490714.1); *Myotis brandtii* (NCBI accession number: XP_005862607.1); *Myotis*
lucifugus (NCBI accession number: XP_023608330.1). A multiple sequence alignment was
generated using the COBALT multiple alignment tool and Neighbour Joining analysis was
performed using COBOLT RID. The rooted tree was generated using a maximum sequence
difference of 0.7, and a Kimura protein distance model.

K0513 is reported to contain a conserved SBF2 domain based on data obtained from
UniProtKB, which is also reportedly present in about 52 % of the 320 orthologues of K0513.
The SBF2 domain is found in Myotubularin related protein 13 (MTMR13; also called
SBF2/SET binding factor 2), and is approximately 220 amino acids in length, based on data
obtained from PFam (pfam12335) [23]. The significance of the SBF domain family
(pfam12335) lies in its association with the Differentially Expressed in Normal and Neoplastic

cells (DENN) domains, which is typically comprised of three parts. The original DENN domain flanked by two divergent domains, the uDENN (upstream DENN; pfam03456), and dDENN (downstream DENN; pfam03455). Some DENN domain containing proteins have been reported to interact with GTPases from the Rab family, functioning as nucleotide exchange factors [24], and others have been reported to be involved in regulation of Mitogen-activated protein kinases (MAPKs) signalling pathways. The myotubularin family is a highly conserved group of ubiquitously expressed phosphatidylinositol 3-phosphatases [25]. The 208 kDa MTMR13 protein is composed of various domains including the tripartite DENN domain, an SBF2 domain, a glucosyltransferase/Rab-like GTPase activator/myotubularin (GRAM) domain, a pseudo-phosphatase domain and a pleckstrin homology (PH) domain. PH domains are involved in membrane association and usually bind phosphates [26,27,28]. The GRAM domain is known to mediate membrane attachment by binding to phosphoinositides [28]. MTMR13 also forms part of the 18 known DENN domain proteins encoded in the human genome [29].

A multiple sequence alignment of K0513a, K0513b, K0513c and SBF2-containing protein, MTMR13, showed that the K0513 protein sequence is associated with the dDENN and SBF2 domains of MTMR13 (Figs 2 and S1). K0513 is partially aligned to the cDENN domain of MTMR13 (Figs 2 and S1). MTMR13 has previously been reported to facilitate exchange of nucleotides for Rab21 and Rab28 [30]. Rab (Ras-related protein in brain) proteins represent the largest family of the small GTPases, with over 70 members in humans [31]. Rab GTPases are important regulators of intracellular membrane trafficking in all eukaryotes [32]. These GTPases are known to interact with effectors such as kinases, phosphatases and tubular-vesicular cargo [33]. Rab GTPases localize to certain intracellular membranes and recruit various Rab effectors including molecular motors, tethering factors and phospholipid modulators [34,35]. They have been described as “switch molecules” in their ability to alter between an active GTP-bound state and an inactive GDP-bound state. In its active state, the Rab GTPase will attach to the membrane and regulate its trafficking through secretory and

endocytic pathways, which is accomplished through interaction with effector proteins known for the connection of activated Rab proteins to a diversity of cellular events such as cytoskeletal dynamics, phosphatidylinositol signaling events and protein kinase-mediated signal transduction [36,37]. For Rab to localize to the membrane, prenylation at the C-terminal cysteine residues is required, and the only known function of prenylation is to confer a hydrophobic character onto a cytosolic protein, thus giving the recipient protein the physical capacity to make a steady attachment with a lipid bilayer [36]. During this process, Rab binds the Rab escort protein (REP), after which Rab geranylgeranyltransferase prenylates the Rab GTPase, which is then transported to the target membrane [38]. Rab is then activated for downstream signalling by a guanine exchange factor (GEF), which facilitates the exchange of guanine-5'-diphosphate (GDP) for guanine-5'-triphosphate (GTP) [35]. Whilst activated, Rab associates with various effector proteins that facilitate the movement of vesicles from donor membranes to target membranes [39]. Rab is then inactivated by binding GTPase activating proteins (GAPs), which hydrolyse GTP to GDP. Finally, inactive Rab is removed from the membrane by the GDP dissociation inhibitor (GDI), which solubilizes the GTPase, thereby preparing Rab for the next cycle of vesicular transport [40].

The DENN domain is found in a variety of signalling proteins involved in Rab-mediated processes or regulation of mitogen-activated protein kinase (MAPK) signalling pathways [24]. DENN domain proteins are GEFs for specific Rab proteins [30,41,42]. Studies have shown that the DENN domain by itself is adequate for guanine exchange function (GEF) activity [41,42]. MTMR13 promotes the exchange of GDP for GTP in the small GTPase proteins, Rab21 and Rab28 [29,43]. MTMR13 converts the inactive GDP-bound Rab28 into its active form, GTP-bound Rab28 protein [30]. The conserved SBF2 domain and cDENN domain in K0513 suggests that K0513 may potentially interact with a Rab protein, and secondly, may play a role in the nucleotide exchange of Rab proteins. The secondary structure of K0513 was predicted to be comprised of alpha helices and coils, with the helices taking up 70 % of the whole structure (Fig 2).

Fig 2. Schematic of the domain organization and multiple sequence alignment of K0513 isoforms and MTMR13 DENN and SBF2 domains.

K0513 was aligned with the known DENN domain containing GEF from the myotubularin family, MTMR13. Segments highlighted by green, orange and blue lines represent domains found in MTMR13 and their respective positions. The hypothetical positions of the cDENN, dDENN and SBF2 domains are highlighted by olive green, orange and blue, respectively. Prediction of the secondary structures of human K0513a was performed using PROMALS3D [13], where helices represent alpha helices, and black lines represent coils. The alignment was generated using Clustal Omega [12].

The amino acid composition and physiochemical properties determine the fundamental properties of a protein. K0513 has a molecular weight of 46,638 kDa and is composed of 411 residues; it has a theoretical isoelectric point of 4,7 (Fig 3). The primary structure analysis shows that K0513 is strongly dominated by polar residues, constituting 55.5 % of the protein thus making the protein hydrophilic in nature (Fig 3). Interestingly, the presence of several cysteine residues suggests that disulphide bridges may constitute a strong component regulating the folding of HsK0513, however, there are none predicted to occur based on DISULFIND (S2 Fig). The computed isoelectric point (pI) of human K0513 is 4.6994, suggesting that the molecule is acidic in nature. This is due in part to the high number of glutamic acid (Glu) residues (Fig 3). The secondary structure of human K0513 was further analysed for hydropathy, amphipacity and topology to predict whether K0513 has capability to bind to the cell membrane in relation to its role as a GEF for Rab, studies have shown that some GEFs are recruited to specific organelle membranes via their transmembrane regions, thereby localizing the GTP-bound Rab to the same membrane [43]. Residue 150 to 167 were predicted to be hydrophobic with the hydrophobicity score of 1.2 and the region is suggested to form part of the transmembrane anchor (Fig 3). Both the N-terminal and the C-terminal are

predicated to be cytosolic and anchored to the membrane through the transmembrane region (Fig 3).

Fig 3. Amino acid composition and physicochemical properties HsK0513.

The amino acid composition of K0513 was computed using PEPSTAT module integrated in the EMBOSS software (A). The amino acid properties of amino acids are summarized in (B). Physicochemical characterization of the protein was computed using the ExPASy's ProtParam server (C). Hydropathy plot showing the degree of hydrophobicity and hydrophilicity, the hydrophobic transmembrane region is highlighted in grey (D). The Hidden Markov Model for Topology Prediction (HMMTOP) showing helical-transmembrane residues and the position of the carboxyl-terminal and amino-terminal relative to the cell cytosol (E). The models were generated using Biotoools Web-based Hydropathy, Amphipathicity and Topology (WHAT) [15].

3.2. The SBF2 domain of K0513 appears to exhibit a coiled coil fold

The predicted three-dimensional structure of human K0513 exhibits a globular fold and the N-terminal loop is positioned towards the C-terminus (Fig 4A). The N-terminal and the C-terminal regions of K0513 are more hydrophobic as compared to the central region of the protein (Fig 4B). The K0513 homology model was generated from several templates using Phyre², for which the highest probability (58,1 %) for shared homology was the Glutathione S-transferase C-terminal domain-like template (Fig 4). The SBF2 domain of K0513 resembles a coiled coil (Fig 4A). GRAB protein has been shown to function as a physiological GEF for Rab3a, specifically through its coiled coil domain [44].

Considering that the three-dimensional structure of a protein may yield valuable information about its functional properties, the structural similarity between proteins may be a good

indicator of functional similarity. Using Phyre², some interesting proteins were predicted to partially align with regions of K0513 (S2 Table). Amongst these was lipid binding protein (PDB ID: c2jobA), which are normally molecules that may reversibly and non-covalently associate with lipids, and are often membrane-associated [45]. Another pair of interesting templates are regulator of g-protein signalling 7 (RGS7; PDB ID: c2d9jA) and RGS2 (PDB ID: c2af0A), which both act as G-protein activating proteins (GAPs) interacting directly with the activated G-protein and whose function is to hydrolyse GTP which in turn promotes inactivation of the G-protein (S2 Table) [46]. Could K0513's function possibly be related to these proteins, with respect to association with cellular membrane and facilitating nucleotide exchange of GTPases?

Fig 4. Three-dimensional homology model of K0513.

The 3D model of K0513 showing the SBF2 domain (yellow) and the phosphoserine residue represented in balls and sticks (pink) (A). The carboxyl-terminal is represented by red whereas the amino-terminal is represented with blue colour. The hydrophobicity surface plot showing hydrophobic (brown), amphipathic (grey) and hydrophilic residues (blue) (B); as well as the schematic presentation of K0513 structure (C) are shown. The model was generated by Phyre² [14] and visualized using Biovia Discovery studio [18].

3.3. Predicted interactome of human K0513

The protein-protein interaction prediction revealed 71 possible interactors of K0513 (S3 Table), interaction partners were generated through the online Integrated Interaction Database (IID) prediction. The prediction identified molecules that had previously been established to interact with K0513 through the yeast two-hybrid system, thereby validating the predictions through experimental data (Table 3) [6]. These included INTS4, which is involved

in the small nuclear RNAs (snRNA) U1 and U2 transcription; HAX1, which regulates reorganization of the cortical actin cytoskeleton, and KIBRA/WWC1, which plays a pivotal role in tumour suppression by restricting proliferation and promoting apoptosis (Table 3) [6]. The predicted interactome included various proteins implicated in GTPase activity. The most interesting of these was predicted interaction of Rab3a, a small GTPase that selectively binds secretory vesicles and switches between the active vesicle associated GTP-bound form and the inactive cytosolic GDP-bound form (Fig 5A; S3 Table) [47]. Another predicted interactor of K0513 is Syntaxin-binding protein 1 (STXBP1, also called MUNC18-1) [47], which is involved in the regulation of synaptic vesicle docking and fusion through specifically binding with GTPases (Fig 5A). Interestingly, Rab3A has been reported to interact directly with Munc18-1 in vesicle priming during exocytosis (Fig 5A; S3 Table) [47]. The protein, IQ motif and SEC7 domain-containing protein 3 (IQSEC3), is also an interesting predicted interactor of K0513, in particular in light of its role as a guanine exchange factor for the small GTPase ADP-ribosylation factor (Arf6), which plays a role in intracellular vesicular trafficking, and regulates the recycling of various types of cargo (Fig 5A; S3 Table) [48]. The last predicted interactor of K0513 involved with GTPase activities is Rho GDP-dissociation inhibitor 3 (ARHGDIG), which functions to inhibit the GDP/GTP exchange reaction of RhoB; RhoB is a small GTPase that regulates vesicle transport and actin reorganization (Fig 5A; S3 Table) [49].

Our *in silico* analysis of human K0513 would suggest the association of K0513 with Rab GTPases. We have seen that K0513 contains a conserved SBF2 domain, and a partially conserved DENN domain (Fig 2), K0513 exhibits regions of sequence similarity to proteins involved in regulation of GTPases (S2 Table), and is predicted to interact with a Rab GTPase, and proteins involved in their regulation (S3 Table). However, it is still unclear in what capacity K0513 may interact with the Rab GTPase. Whether K0513 hydrolyses GTP, thereby acting as a GAP protein; or dissociates a GAP from Rab GTPase, thereby acting as a GDF; or even directly associates with Rab GTPase, inducing a conformational change which facilitates the exchange of GTP for GDP, and thereby acting as a GEF, is unclear and remains to be

determined (Fig 5). We were therefore curious to investigate whether nucleotides could induce conformational changes K0513 through biophysical analysis to develop a better understanding.

Fig 5. Predicted interactome of K0513 and proposed roles of K0513 in the Rab3a signal transduction pathway.

(A) K0513 is predicted to interact with Rab3a, Mun18-1, and Munc18-1 has been shown to interact with Rab3a in processes facilitating vesicle priming for endocytosis; K0513 is predicted to interact with IQSEC3 which functions as a GEF for Arf6, which facilitates vesicular trafficking and cargo recycling; K0513 is predicted to interact with ARHGDIG1 which is the GDP dissociation inhibitor for RhoB, and facilitates vesicular transport and actin reorganization. (B) Newly synthesized Rab interacts with Rab escort protein (REP) in its GDP bound state, recruiting the small GTPase to Rab geranylgeranyl transferase (RabGGT) which prenylates Rab at C-terminal cysteine residues. Rab is then activated by a guanine exchange factor (GEF) which exchanges GDP for GTP and facilitates recruiting Rab to specific cellular membranes. GTP-bound Rab binds effector proteins that mediate various steps in vesicular trafficking. 2. Rab is inactivated by interacting with GTPase activating protein (GAP) through hydrolysis of GTP to yield GDP-bound Rab. 3. Prenylated Rab is removed from the membrane and solubilized by interacting with GDP dissociation inhibitor (GDI), which shields the nonpolar geranylgeranyl groups in the prenyl tails from the polar environment. 4. GDI dissociation factor (GDI) separates Rab from GDI and so allows Rab to be inserted into the target membrane, in preparation for the next cycle.

3.4. Overexpression and purification of recombinant K0513

The *E. coli* expressed recombinant K0513 protein migrated as a species of 47 kDa on SDS-PAGE, which correlated to the predicted molecular weight. K0513 protein was purified in its the native form using Nickel-affinity chromatography (Fig 6). The authenticity of the successfully expressed and purified protein was validated by Western blotting target specific antibodies (Fig 6). The overall purification yielded 20 mg per 1 L total culture volume of approximately 80 % purity (Fig 6).

Fig 6. Heterologous expression and purification of recombinant K0513 using nickel affinity chromatography.

SDS-PAGE (12 %) analysis of K0513 expression in *E. coli* XL1-Blue cells. Lanes 1 – Protein marker; 2 – total extract of cells transformed with pQE60 as control; 3 – total extract of cells transformed with pQE60-KIAA prior to IPTG induction; 4 - 8 – total cell lysates obtained up to 5 hours post IPTG induction; 9, 10 – pellet and soluble fractions obtained from total lysate of cells transformed with pQE60-K0513, respectively; 11 - flow through fraction; 12 – wash fraction, 13 - 15 – elution fractions. Western blotting was performed using anti-K0513 and Anti-Histidine antibodies.

3.5. Investigation of the effect of nucleotides on the tertiary structure of K0513

We investigated the effects of nucleotides on the tertiary structure of the recombinant K0513 protein using tryptophan-based fluorescence assay. The K0513 protein sequence possesses a total of six tryptophan residues at positions W65, W134, W218, W300, W306 and W339. The conformation of K0513 in the presence of nucleotides ATP/ADP and GTP/GDP was assessed (Fig 7). Recombinant K0513 had a maxima of 332 nm at 0 mM ATP, ADP, GTP and GDP. We observed a marginal red shift in the emission maxima of recombinant K0513 exposed to increasing concentrations of ATP, ADP, GTP and GDP of approximately 3 nm, 3

nm, 4 nm and 2 nm, respectively (Fig 7). In addition, K0513 was incubated with a fixed concentration of GTP and the fluorescence was monitored over time (Fig 7). There was no significant red or blue shift of the 332 nm peak. This data signifies that the conformation of K0513 remains unaltered in the presence of nucleotides, which would suggest that K0513 is not binding nor hydrolysing nucleotides.

Fig 7. Effect of nucleotides on the tertiary structure of K0513.

The conformational change of K0513 was assessed in the presence of various concentrations of ATP and the shift in wavelength was plotted against concentration (A), the same was followed for ADP (B), GTP (C) and GDP (D). The conformational change of K0513 was assessed in the presence of a fixed concentration of GTP at different time points, and the shift in wavelength was plotted against time point in minutes (E). Initial excitation was at 295 nm and the spectra were monitored between 280 to 500 nm, 1000 nm/min scan speed, 3 accumulations and 10 nm emission bandwidth.

To determine whether the observed fluorescence spectral shifts were due to conformational changes in the protein tertiary structure, chaotropic denaturants were used. The maximum fluorescence intensity for native K0513 was observed at 332 nm, which is expected [50]. This would suggest that K0513 assumes a folded state prior to introduction of the denaturants. The addition of urea and guanidine HCl to recombinant K0513 at increasing concentrations resulted in a significant red shift for both denaturants (Fig 8). The red shift signifies that recombinant K0513 is now in a more unfolded and extended state due to resulting surface exposed charged residues [51]. Taken together, this data would suggest that K0513 conformation is not perturbed by the presence of nucleotides.

Fig 8. Conformation of K0513 in the presence of guanidine-HCl and urea.

K0513 conformation was assessed using tryptophan fluorescence spectrophotometry in the presence of various guanidine-HCl concentrations, and the wavelength shift was plotted against its concentration (A); the same was followed in the presence of urea (B). Initial excitation was at 295 nm and the spectra were monitored between 280 to 500 nm, 1000 nm/min scan speed, 3 accumulations and 10 nm emission bandwidth.

To further validate the results that were obtained using the tryptophan-based fluorescence spectroscopy, we performed a trypsin based limited proteolysis assay in which we monitored the effect of nucleotides on the conformation of recombinant K0513 (Fig 9). The protein was digested to smaller fragments, as was represented by smaller bands on SDS-PAGE. This demonstrated that recombinant K0513 is a proteolytically susceptible molecule. The fragmentation profiles obtained by limited proteolysis of K0513 in the presence of various nucleotides ATP/ADP and GTP/GDP were not distinct (Fig 9). This suggests that the various nucleotides did not uniquely modulate the structure of the protein. This is in support of the data obtained by intrinsic fluorescence analysis (Fig 9).

Fig 9. Proteolytic digest of K0513 in the absence and presence of nucleotides.

SDS-PAGE analysis of recombinant K0513 that was subjected to limited proteolysis (LP) in the presence of ATP/ADP (A) and in the presence of GTP/GDP (B). Lanes M - Protein marker; U – undigested K0513; 1 - K0513 subjected to LP; 2 – K0513 in the presence of 5 mM ATP subjected to LP; 3 - K0513 in the presence of 5 mM ADP subjected to LP; 4 - K0513 in the presence of 1 mM GTP subjected to LP; 5 - K0513 in the presence of 1 mM GDP subjected to LP. The limited proteolysis by trypsin was conducted using an enzyme to substrate ratio 1 : 2000. The duration of exposure to trypsin is given in minutes.

We further investigated the presence of K0513 in various lysates prepared from both normal and cancerous human cell lines. Using Western blotting, K0513 was detected in all the cell lines tested including retinal pigment epithelial cells, keratinocytes, lung carcinoma cells, fibrosarcoma cells, and cervical adenocarcinoma cells, albeit at very low levels in the latter (Fig 10). Notably, K0513 was detected as a doublet band in both the epithelial and fibrosarcoma lysates at ~ 45 kDa and ~ 46 kDa, which could signify the presence of two K0513 isoforms (Fig 10). This finding opens the door for further establishment of the role of K0513 in cancer cells.

Fig 10. Detection of K0513 in human lysates prepared from different cell lines to assess the anti-K0513 antibody.

K0513 was detected in various human cell lines by Western blotting using the rabbit raised anti-K0513 antibody at a 1:100 dilution. Lanes M – Protein marker; 1 - Human retinal pigment epithelial (RPE-1) whole cell lysate; 2 - Human immortalized keratinocyte (HaCaTs) whole cell lysate; 3 – Human lung carcinoma (A549) whole cell lysate; 4 - Human cervical adenocarcinoma (HeLa) whole cell lysate; 5 – Human fibrosarcoma whole cell lysate.

4. DISCUSSION

Our findings which confirmed that human K0513 contains conserved SBF2 and dDENN domains, and a transmembrane region, suggest a possible role of this protein as nucleotide exchange factor for Rab GTPases. *In silico* analysis revealed the protein to be predominantly composed of alpha helices and coils, and homology modelling showed the fold of K0513 to be globular in nature. Investigating the predicted interactome of K0513 revealed its possible association with proteins related to GTPase pathways, in particular, Rab3a, a small GTPase

protein. Furthermore, we optimised a protocol for the expression and purification of K0513. Biophysical characterisation of the protein established that the tertiary fold of K0513 is not modulated by nucleotides, suggesting that it may not directly interact with nucleotides, as observed using both tryptophan-based fluorescence and limited proteolysis assays. In addition, data obtained from the tryptophan-based fluorescence assay revealed a pronounced red shift of the protein in the presence of denaturants, suggesting that, in its native state, K0513 exists as a globular protein. This is the first report of recombinant K0513 protein production from a bacterial system, and analysis of tertiary structure. Considering that we found the conformation of K0513 to be unperturbed by nucleotides, together with the *in silico* analysis, we propose that K0513 may function as a GEF of Rab3a. Our findings further suggest that K0513 does not necessarily bind to G-proteins (GAP function) to accelerate their GTPase activity through hydrolysis of GTP.

Conservation of protein structure often mirrors similarities in protein function, and so this study sought to identify the structure-function features of “uncharacterized protein” K0513. K0513 comprises a conserved SBF2 domain, which is present in MTMR13. The SBF2 domain is just downstream of the dDENN domain (Fig 2). The DENN domain containing proteins are reported to be involved in a variety of signalling pathways, however, there is a significant degree of variation within these domains from protein to protein [29]. The DENN domain found in MTMR13 has been suggested to facilitate nucleotide exchange function of specific small GTPases from the Rab family [30]. The DENN domain binds to switches I and II of Rab21 GTPase resulting in conformational change in the region to allow GDP to GTP exchange [52]. The tripartite DENN module is comprised of the upstream DENN (uDENN), followed by central DENN (cDENN), and the downstream DENN (dDENN). The uDENN may occur independently of the DENN module and interact with various GTPases [53]. The central cDENN and dDENN occur together, the d-DENN is reported to make fewer contact with the GTPase thus it always associates with cDENN to maintain the GEF function [52]. A multiple protein sequence alignment of MTMR13 and the K0513 isoforms interestingly exhibited some similarity of K0513

over the cDENN and dDENN domains of MTMR13 as well as the SBF2 domain (Fig 2). From the sequence conservation observed between K0513 and the DENN and SBF2 domains of MTMR13, we asked whether protein K0513 cross-talks to a small GTPase protein, possibly through a nucleotide exchange function. There is no evidence on how the SBF2 domain facilitates the nucleotide exchange activity of the DENN domain, however, their co-existence strongly suggests they have a functional relationship. Homology determination is important in ascertaining functional features in protein structures especially in the absence of a crystal structure, as well as structural prediction of proteins. The query protein sequence can be aligned with identified homologues for determination of domains and motifs based on the level of conservation between the protein. Although K0513 is well conserved amongst its homologues, K0513 exhibits low conservation to proteins that have had their crystal structures resolved, and as a result the three-dimensional model of K0513 was generated using segments of different templates (Fig 4). The three-dimensional homology model of K0513 shows a globular fold, and the SBF2 domain appears to be surface exposed and a coiled coil in structure (Fig 4). The Phyre² tool also showed segments of proteins that align to regions of K0513; it was notable that some of these proteins are associated with the GTPase pathway, in particular two GAP proteins (S2 Table).

Protein-protein interaction prediction is crucial in ascertaining pathways and cellular processes that the protein is involved in. This study used the Integrated Interaction Database (IID) to predict protein-protein interactions [19]. IID is the only database with context-specific networks for the common model organisms and domesticated species [19], it integrates experimentally detected protein interactions from nine curated databases (BioGRID, IntAct, I2D, MINT, InnateDB, DIP, HPRD, BIND, BCI). Predicting the K0513 interactome may reveal signalling pathways that K0513 could be involved in and so enable prediction of function. Physicochemical analysis showed that K0513 possess a transmembrane region. The 17 residues long alpha-helical region consists of hydrophobic residues positioned from amino acid residue 150 to 167. Transmembrane proteins can be integrally anchored to the lipid

bilayer for interaction with membrane proteins. For example, general receptor for phosphoinositides 1 (Grp1), a known GEF interact with PI3-kinase in the inner membrane through its pleckstrin homology (PH) domain [54]. When anchored to the membrane, Grp1 catalyses the nucleotide exchange and activation of ADP-ribosylation factor-6 (Arf6) [55]. The presence of the transmembrane suggest that K0513 performs its nucleotide exchange function on membrane localised Rab GTPases when it is anchored to the lipid bilayer. Furthermore, it may directly or indirectly facilitate cellular processes such as vesicle trafficking or serve in transducing signal in pathways such as neuroplasticity, cytoskeletal regulation and apoptosis [6]. The predicted interaction with the transmembrane proteins, most notably the Na⁺/Ca²⁺ Solute Carrier Family 8 Member A2 (SLC8A2) further substantiate this finding. Most importantly, Rab3a, a GTPase involved in vesicle trafficking has been predicted to interact with K0513 supporting the role of K0513 along the membrane. Another important finding is the identification of the possible phosphorylation at S279. Protein phosphorylation serves as an initial step crucial for facilitating cellular and organic functions such as regulation of metabolism, subcellular trafficking, and signalling transduction [56]. Phosphoresidues can be activated or deactivated by kinases or phosphatases, this modification induces conformational change when the protein interacts with other proteins. Some GEFs require phosphorylation to activate their GEF activity, p115, the smallest member of the RH-RhoGEF subfamily gets phosphorylated at Tyr-738 by Janus kinase 2 (JAK2) to positively regulate its GEF activity [57]. The presence of the phosphoserine residue suggests that K0513 may be required to undergo phosphorylation to perform its GEF function or other cellular processes.

Escherichia coli has been considered a pioneering host for recombinant protein production. Using *E. coli* remains a popular choice when attempting the production of a new protein because of the existing extensive knowledge on its genetics and physiology, and the number of genetic engineering tools adapted to *E. coli*. In this study, codon harmonization was applied to the human protein, K0513, through adjustment of the codon usage bias, adjustment of the GC content and removal of repeat sequences [58]. Codon harmonization by synonymous

substitution of the coding sequence is designed to increase the expression level of the protein of interest [58]. K0513 was produced as a soluble recombinant protein. Its solubility may suggest cytosolic localisation (Fig 5) [6].

The fluorescence data provided preliminary insights on the tertiary structural features of K0513. Denaturing the protein with urea and guanidine-HCl resulted in a significant red shift of the intensity maxima in a concentration dependent manner (Fig 7). This reveals that the fluorescing Trp residues are located in a nonpolar hydrophobic interior [59]. This is because Trp has two nearby isoenergetic excitation states, 1L_a and 1L_b [60] which differs with sensitivity to solvent. In hydrophobic environments, 1L_b dominates the Trp emission and display a blue shift of Trp fluorescence due to an increase in non-polarity [61]. In contrast, denatured proteins with solvent exposed Trp residues displays a red shift [62]. The increase in the concentration of denaturants resulted in a significant red shift in the wavelength maxima (Fig 8), this shift suggests the exposure of Trp residues that are positioned within the hydrophobic interior [59]. This further suggests that, in its native state, K0513 exists as a globular protein. K0513 depicted no significant conformational changes when exposed to ATP, ADP, GTP nor GDP which suggests that K0513 may not be binding these nucleotides directly (Fig 7). This finding was supported by investigating the effect of nucleotides on the conformation of K0513 using limited proteolysis. The limited proteolysis assay showed that K0513 is a proteolytically susceptible molecule, however, its conformation remains unperturbed in the presence of nucleotides (Fig 9). K0513 was detected in a series of human cell lines, including lung carcinoma, fibrosarcoma and cervical adenocarcinoma cells, which could suggest its involvement in various pathways related to tumorigenesis, and possibly the importance of characterizing the function of K0513.

The *in silico* and biophysical analyses reported on in this study suggest that K0513 is involved in the Rab3a GTPase signalling pathway. Furthermore, due to lack of of significant conformational response to nucleotides, we suggest that K0513 may facilitate nucleotide

exchange of Rab3a, by functioning as a GEF, inducing conformational changes in the Rab GTPase, and thereby activating the Rab3a. GEFs are potential targets for cancer therapy due to their role in many signalling pathways, especially in cell proliferation. For instance, several cancers result from mutations in the MAPK/ERK pathway that leads to proliferation. The GEF SOS1 activates Ras, whose target is the kinase Raf. Raf is a protooncogene due to the numerous mutations it exhibits in various cancers [63,64]. The Rho GTPase Vav1, which could be activated by the GEF receptor, has been shown to promote tumour proliferation in pancreatic cancer; GEFs represent possible therapeutic targets as they can potentially play a role in regulating these pathways through their activation of GTPases.

5. CONCLUSIONS

This is the first study to describe the expression, purification and characterisation of human K0513. Biophysical and *in silico* studies conducted on human K0513 here suggest that it is a soluble, globular protein that possesses SBF2 and dDENN domains. This suggests a role for this protein as nucleotide exchange factor for Rab3a. This implicates K0513 in neurotransmitter signalling pathways, possibly facilitating the role of Rab3a which is involved in the transport of synaptic vesicles in the brain [65,66], thus suggesting a crucial role for this protein in cancer development. Not only is the *K0513* upregulated in glioblastoma patients, but *K0513* has also recently been highlighted as one of four genes in a model to predict prognosis of pancreatic cancer [7]. The current study findings are important for further elucidation of the role of K0513, particularly in signalling pathways associated with tumorigenesis.

6. ACKNOWLEDGEMENTS

We thank the Hair and Skin Research Unit in the Department of Dermatology at the University of Cape Town Medical School for expertise and use of their cell culture facility.

7. FUNDING STATEMENT

This work was supported by the University of Venda Research Grant [grant number SMNS/17/BCM/17] awarded to AB. The authors are grateful to the Department of Science and Technology/National Research Foundation (NRF) of South Africa for providing an equipment grant (UID, 75464) awarded to AS. NN is a recipient of the NRF of South Africa Freestanding, Innovation and Scarce Skills Masters Scholarship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

8. DATA AVAILABILITY

All relevant data are within the manuscript and its Supporting Information files.

9. DECLARATION OF COMPETING INTEREST

The authors have declared that no competing interests exist.

10. REFERENCES

1. Zhu P, Du XL, Lu G, Zhu JJ. Survival benefit of glioblastoma patients after FDA approval of temozolomide concomitant with radiation and bevacizumab: a population-based study. *Oncotarget*. 2017;8: 44015–31. doi: 10.18632/oncotarget.17054.
2. Ostrom QT, Gittleman H, Fulop J, Liu M, Blanda R, Kromer C, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2008–2012. *Neuro Oncol*. 2015;17(Suppl. 4):iv1–62. doi: 10.1093/neuonc/nov189.
3. Kanu OO, Mehta A, Di C, Lin N, Bortoff K, Bigner DD. Glioblastoma multiforme: a review of therapeutic targets. *Expert Opin Ther Targets*. 2019;13(6): 701-718.
4. Rao SK, Edwards J, Joshi AD, Siu IM, Riggins GJ. (2010). A survey of glioblastoma genomic amplifications and deletions. *J. Neurooncol*. 2010;96(2): 169-179.

5. Hoelzinger DB, Mariani L, Weis J, Woyke T, Berens TJ, McDonough WS, et al. Gene expression profile of glioblastoma multiforme invasive phenotype points to new therapeutic targets. *Neoplasia*. 2005;7(1): 7-16. doi:10.1593/neo.04535.
6. Lauriat TL, Dracheva S, Kremerskothen J, Duning K, Haroutunian V, Buxbaum JD, et al. Characterization of K0513, a novel signaling molecule that interacts with modulators of neuroplasticity, apoptosis, and the cytoskeleton. *Brain Res*. 2006;1121: 1-11.
7. Chen K, He Y, Liu Y, Yang X. Gene signature associated with neuro-endocrine activity predicting prognosis of pancreatic carcinoma. *Mol. Gen. Genom. Med.* 2019;7:e729 doi:10.1002/mgg3.729.
8. Olsson M, Beck S, Kogner P, Martinsson T, Carén H. Genome-wide methylation profiling identifies novel methylated genes in neuroblastoma tumors. *Epigen*. 2016;11(1): 74-84.
9. Uhlén M, Zhang XC, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G. et al. A pathology atlas of the human cancer transcriptome. *Science*. 2017;357(6352). doi: 10.1126/science.aan2507. Database: Protein Atlas. Available from: <http://www.proteinatlas.org>.
10. Standfuß C, Parczyk J, Ruhnau J, Klein A. Genome organization in different cancer types:detection of cancer specific breakpoint regions. *Mol. Cytogen*. 2019;12: 28-35.
11. NCBI Resource Coordinators. Database resources of the National Centre for Biotechnology Information. *Nucleic Acids Res*. 2018; 46(D1), D8–D13. <https://doi.org/10.1093/nar/gkx1095>. Database: NCBI. Available from: <http://www.ncbi.nlm.nih.gov>.
12. Sievers F, Higgins DG (2002). Clustal Omega. *Current Protocols in Bioinformatics*. 48. John Wiley & Sons, Inc. pp. 3.13.1–16. doi:10.1002/0471250953.bi0313s48. ISBN 9780471250951. PMID 25501942. Available from: <https://www.ebi.ac.uk/Tools/msa/clustalo/>
13. Pei J, Kim B-H, Grishin NV. PROMALS3D: a tool for multiple sequence and structure alignment. *Nucleic Acids Res*. 2008;36(7): 2295-2300.

14. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modelling, prediction and analysis. *Nat Protoc.* 2015;10: 845-858. Database: Phyre2. Available from: <http://www.sbg.bio.ic.ac.uk/phyre2>.
15. Ison, J. Rapacki K, Ménager H, Kalaš M, Rydza E, Chmura P. et al. Tools and data services registry: a community effort to document bioinformatics resources. *Nucleic Acids Res.* 2015. doi: 10.1093/nar/gkv1116. Database: Biotoools. Available from: <http://biotoools.tcd.org/barwhat2.html>.
16. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 2019;47(W1): W636–W641. doi: <https://doi.org/10.1093/nar/gkz268>. Database: PEPSTAT. Available from: https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/.
17. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. 2005. pp. 571-607. Database: ProtParam. Available from: <https://web.expasy.org/protparam/>.
18. Dassault Systèmes BIOVIA, [Biovia Discovery Studio Visualizer], [Software version 20.1], San Diego: Dassault Systèmes, [2019].
19. Kotlyar M, Pastrello C, Malik Z, Jurisica I. IID 2018 update: context-specific physical protein-protein interactions in human, model organisms and domesticated species. *Nucleic Acids Res.* 2019;47(D1): D581-D589. doi:10.1093/nar/gky1037. Available from: <http://ophid.utoronto.ca/iid>.
20. Burger A, Ludewig MH, Boshoff A. Investigating the Chaperone Properties of a Novel Heat Shock Protein, Hsp70.c, from *Trypanosoma brucei*. *J. Parasitol. Res.* 2014;172582 doi: 10.1155/2014/172582.
21. Zininga T, Makumire S, Gitau GW, Njunge JM, Pooe OJ, Klimek H, *et al.* Plasmodium falciparum Hop (PfHop) interacts with the Hsp70 chaperone in a nucleotide-dependent fashion and exhibits ligand selectivity. *PloS One.* 2015;10(8).

791 22. Holden P, Horton WA. Crude subcellular fractionation of cultured mammalian cell lines.
792 BMC research notes 2009;2(1): 243.

793 23. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M,
794 Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D,
795 Tosatto SCE, Finn RD. The Pfam protein families database in 2019. Nucleic Acids
796 Res. 2019; doi: 10.1093/nar/gky995

797 24. Levivier E, Goud B, Souchet M, Calmels TP, Mornon JP, Callebaut I. uDENN, DENN, and
798 dDENN: indissociable domains in Rab and MAP kinases signaling pathways. Biochem.
799 Biophys. Res. Commun. 2001;287: 688–695 10.1006/bbrc.2001.5652.

800 25. Laporte J, Bedez F, Bolino A, Mandel JL. Myotubularins, a large disease-associated family
801 of cooperating catalytically active and inactive phosphoinositide phosphatases. Hum. Mol.
802 Genet. 2003;12: R285– R292.

803 26. Isakoff SJ, Cardozo T, Andreev J, Ferguson KM, Abagya R, Lemmon MA, et al.
804 Identification and analysis of PH domain-containing targets of phosphatidylinositol 3-kinase
805 using a novel in vivo assay in yeast. EMBO J. 1998;17(18): 5374-5387.
806 doi:10.1093/emboj/17.18.5374.

807 27. Robinson FL, Dixon JE. Myotubularin phosphatases: policing 3-phosphoinositides. Trends
808 Cell Biol. 2006;16(8): 403-412. doi:10.1016/j.tcb.2006.06.001.

809 28. Berger P, Niemann A, Suter U. Schwann cells and the pathogenesis of inherited motor
810 and sensory neuropathies (Charcot-Marie-Tooth disease). Glia. 2006;54: 243–257.

811 29. Marat AL, Dokainish H, McPherson PS. DENN domain proteins: regulators of Rab
812 GTPases. J Biol Chem. 2011;286(16): 13791-13800.

813 30. Yoshimure S, Gerondopoulos A, Linford A, Rigden DJ, Barr FA. Family-wide
814 characterization of the DENN domain Rab GDP-GTP exchange factors. J. Cell Biol.
815 2010;191(2): 367-381.

816 31. Barr F, Lambright DG. Rab GEFs and GAPs. Curr. Opin. Cell Biol. 2010;22: 461-470.
817 doi:10.1016/j.ceb.2010.04.007.

32. Diekmann Y, Seixas E, Gouw M, Tavares-Cadete F, Seabra MC, Pereira-Leal JB. Thousands of rab GTPases for the cell biologist. PLoS. Comput. Biol. 2011;7, e1002217. doi:10.1371/journal.pcbi.1002217.
33. Wandinger-Ness A, Zerial M. Rab proteins and the compartmentalization of the endosomal system. Cold Spring Harb. Perspect. Biol. 2014;6: a022616. doi:10.1101/cshperspect.a022616.
34. Stenmark, H. Rab GTPases as coordinators of vesicle traffic. Nat. Rev. Mol. Cell Biol. 2009;10: 513-525. doi:10.1038/nrm2728.
35. Hutagalung AH, Novick PJ. Role of Rab GTPases in membrane traffic and cell physiology. Physiol. Rev. 2011;91: 119-149. doi:10.1152/physrev.00059.2009.
36. Calero M, Chen CZ, Zhu W, Winand N, Havas KA, Gilbert PM, et al. Dual prenylation is required for Rab protein localization and function. Mol Biol Cell. 2003;14(5): 1852-1867. doi:10.1091/mbc.e02-11-0707.
37. Zhen Y, Stenmark H. Cellular functions of Rab GTPases at a glance. J Cell Sci. 2015;128(17): 3171-3176. doi:10.1242/jcs.166074.
38. Alexandrov K, Simon I, Yurchenko V, Iakovenko A, Rostkova E, Scheidig AJ, et al. Characterization of the ternary complex between Rab7, REP-1 and Rab geranylgeranyl transferase. Eur. J. Biochem. 1999; 265:160-170.
39. Oesterlin LK, Pylypenko O, Goud B. Effectors of Rab GTPases: Rab Binding Specificity and Their Role in Coordination of Rab Function and Localization. In: Wittinghofer A. (eds) Ras Superfamily Small G Proteins: Biology and Mechanisms 2. Springer, Cham. 2018.
40. Pfeffer S, Aivazian D. Targeting Rab GTPases to distinct membrane compartments. Nat Rev Mol Cell Biol. 2004;5(11): 886-896. doi:10.1038/nrm1500.
41. Allaire PD, Marat AL, Dall'Armi C, Di Paolo G, McPherson PS, Ritter B. The Connecdenn DENN domain: a GEF for Rab35 mediating cargo-specific exit from early endosomes. Mol Cell. 2010;37(3): 370-382. doi:10.1016/j.molcel.2009.12.037.

42. Marat AL, McPherson PS. The connecdenn family, Rab35 guanine nucleotide exchange factors interfacing with the clathrin machinery. *J Biol Chem.* 2010;285(14): 10627-10637. doi:10.1074/jbc.M109.050930.
43. Cabrera M, Ungermann C. Guanine nucleotide exchange factors (GEFs) have a critical but not exclusive role in organelle localization of Rab GTPases. *J Biol Chem.* 2013;288: 28704-28712. doi: 10.1074/jbc.M113.488213.
44. Luo HR, Saiardi A, Nagata E, Ye K, Yu H, Jung TS, et al. GRAB: a physiologic guanine nucleotide exchange factor for Rab3A, which interacts with inositol hexakisphosphate kinase. *Neuron.* 2001;31: 439–451.
45. Glatz JFC. Lipids and lipid binding proteins: A perfect match. *PLEFA.* 2015;93: 45-49.
46. Linder A, Thulin MH, Damber JE, Welén K. Analysis of regulator of G-protein signalling 2 (RGS2) expression and function during prostate cancer progression. *Sci Rep.* 2018;8: 17259. <https://doi.org/10.1038/s41598-018-35332-4>.
47. Huang C-C, Yang D-M, Lin C-C, Kao L-S. Involvement of Rab3A in vesicle priming during exocytosis:interaction with Mun13-1 and Munc18-1. *Traffic.* 2011;12(10): 1356-1370. doi:10.1111/j.1600-0854.2011.01237.x.
48. Van Acker T, Tavernier J, Peelman F. The Small GTPase Arf6: An Overview of Its Mechanisms of Action and of Its Role in Host-Pathogen Interactions and Innate Immunity. *Int J Mol Sci.* 2019;20(9): 2209. doi: <https://doi.org/10.3390/ijms20092209>.
49. Prendergast, G. Actin' up: RhoB in cancer and apoptosis. *Nat Rev Cancer.* 2001;1: 162168. doi: <https://doi.org/10.1038/35101096>.
50. Ramachandran S, Kota P, Ding F, Dokholyan NV. Automated minimization of steric clashes in protein structures. *Proteins.* 2011;79(1): 261-270. doi:10.1002/prot.22879.
51. Vivian JT, Callis PR. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys J.* 2001;80(5): 2093-2109. doi:10.1016/S0006-3495(01)76183-8.
52. Wu X, Bradley MJ, Cai Y, Kümmel D, De La Cruz EM, Barr FA, et al. Insights regarding guanine nucleotide exchange from the structure of a DENN-domain protein complexed with

871 its Rab GTPase substrate. *Proc Natl Acad Sci U S A*. 2011;108(46):18672-18677.
872 doi:10.1073/pnas.1110415108.

873 53. Schlenker O, Hendricks A, Sinning I, Wild K. The structure of the mammalian signal
874 recognition particle (SRP) receptor as prototype for the interaction of small GTPases with
875 Longin domains. *J Biol Chem*. 2006;281(13): 8898-8906. doi:10.1074/jbc.M512415200.

876 54. Albers J, Wayne RR. *Cell Membrane Structures and Functions*. Basic Neurochem. 2012.
877 pp. 26-39. doi: 10.1016/B978-0-12-374947-5.00002-X.

878 55. Venkataraman A, Nevrvy DJ, Filtz TM, Leid M. Grp1-associated scaffold protein (GRASP)
879 is a regulator of the ADP ribosylation factor 6 (Arf6)-dependent membrane trafficking pathway.
880 *Cell Biol Int*. 2012;36: 1115-1128. doi:10.1042/CBI20120221.

881 56. Ardito F, Giuliani M, Perrone D, Troiano G, Lo Muzio L. The crucial role of protein
882 phosphorylation in cell signaling and its use as targeted therapy (Review). *Int J Mol Med*.
883 2017;40(2): 271-280. doi:10.3892/ijmm.2017.3036.

884 57. Patel M, Karginov AV. Phosphorylation-mediated regulation of GEFs for RhoA. *Cell Adh*
885 *Migr*. 2014;8(1): 11-18. doi: 10.4161/cam.28058.

886 58. Brule CE, Grayhack EJ. Synonymous codons: choose wisely for expression. *Trends*
887 *Genet*. 2017;33: 283-297.

888 59. Keshav GC, Davidnhan T, Jayalath K, Abeysirigunawardena S. Discovery of a novel small
889 molecular peptide that disrupts helix 34 of bacterial ribosomal RNA. *RSC Adv*. 2019;9: 40268-
890 40276. doi: 10.1039/C9RA07812F.

891 60. Albani JR Origin of tryptophan fluorescence lifetimes part 1. Fluorescence lifetimes origin
892 of tryptophan free in solution. *J fluoresce*. 2014;24(1): 93-104.

893 61. Ghisaidoobe AB, Chung SJ. Intrinsic tryptophan fluorescence in the detection and
894 analysis of proteins: a focus on Förster resonance energy transfer techniques. *Int J Mol Sci*.
895 2014;15(12): 22518-22538. doi:10.3390/ijms151222518.

896 62. Eftink MR. Intrinsic fluorescence of proteins. In *Topics in Fluorescence Spectroscopy*. J.
897 R. Lakowicz, editor. Kluwer Academic/Plenum Publishers, New York. 2000. pp. 1–15.

63. Quilliam LA, Rebhun JF, Castro AF. A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases. *Prog Nucleic Acid Res Mol Biol.* 2002;71: 391–444. doi:10.1016/S0079-6603(02)71047-7.
64. Schmidt A, Hall A. Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. *Genes Dev.* 2002;16 (13): 1587–609. doi:10.1101/gad.1003302.
65. Geppert M, Bolshakov VY, Siegelbaum SA, Takei K, De Camilli P, Hammer RE, et al. The role of Rab3A in neurotransmitter release. *Nature.* 1994;369(6480): 493-497. doi:10.1038/369493a0.
66. Tang B. C9orf72's Interaction with Rab GTPases—Modulation of Membrane Traffic and Autophagy. *Front Cell Neuro.* 2016;10. doi: 10.3389/fncel.2016.00228.

12. SUPPORTING INFORMATION

S1 Table. Summary of the seven transcript variants of human *K0513*. The table shows features of the four transcripts of the human gene *K0513* (HGNC accession no. 29058) including a description, gene and protein sizes and accession numbers. *K0513* is located on chromosome 16: 85,027,751-85,094,230 forward strand. This gene has 238 orthologues (Ensembl; Cunningham *et al.*, 2019). The predicted *K0513* transcript variants X1 and X2 are not included in this table; they both code for isoform a.

S2 Table. Proteins of interest that aligned to regions of *K0513* as predicted by Phyre². The table shows selected proteins that align partially to regions of human *K0513* predicted using Phyre², their sequence identities, and a description of each protein.

S3 Table. The predicted interactors of human K0513.

S1 Fig. Pairwise sequence alignment of human K0513 and MTMR13. Pairwise sequence alignment of human K0513 (Accession: NP_055547.1) and human MTMR13 (Accession: Q86WG5.1). The uDENN (Olive green) is situated upstream of the DENN (green) followed by the dDENN (light green) domain. The hypothetical position of the SBF2 is represented by red, the GRAM (Glucosyltransferase/Rab-like GTPase activator/myotubularin domain) domain is represented by purple followed by the pseudo-phosphatase (Blue), and lastly the PH (Plenkin homolog) situated towards the C-terminal. The alignment was generated using EMBL-EBI Pairwise Alignment Tool (<https://www.ebi.ac.uk/Tools/>).

S2 Fig. Prediction of human K0513 Cysteines disulfide bonding state and connectivity.

No disulphide patterns were predicted for any of the eight cysteine residues present in human K0513 using the server DISULFIND. Symbol AA indicates amino acid sequence; DB_state indicates predicted disulfide bonding state where 1=disulfide bonded and 0=not disulfide bonded; and DB_conf indicates confidence of disulfide bonding state prediction where 0=low to 9=high.

953 Fig 1

955 Fig 2

Fig 3

Fig 4

Fig 5

997 Fig 6

998

999

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013 Fig 7

1014

1015

1016

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

1027

Fig 8

Fig 9

Fig 10

S1 Table

K0513 TRANSCRIPT VARIANT	DESCRIPTION	mRNA bp	PROTEIN ISOFORM	CCDS	PROTEIN ID (NCBI Ref Seq)	M. WT (KDA)
1	This variant (1) represents the longest transcript and encodes the longer isoform (a). Variants 1 and 2 both encode the same isoform (a)	7781	Isoform a (411 aa)	CCDS32499.1	NP_001273494.1	46,508
2	This variant (2) differs in the 5' UTR compared to variant 1. Variants 1 and 2 both encode the same isoform (a)	7423	Isoform a (411 aa)	CCDS32499.1	NP_055547.1	46,508
3	This variant (3) differs in the 5' UTR and lacks an alternate in-frame exon compared to variant 1. The resulting isoform (b) has the same N- and C-termini but is shorter compared to isoform a.	7400	Isoform b (401 aa)	CCDS67091.1	NP_001273495.1	45,407
4	This variant (4) differs in the 5' and 3' UTRs and has multiple coding region differences, compared to variant 1. The resulting protein (isoform c) has a distinct C-terminus and is shorter than isoform a.	1951	Isoform c (301 aa)	CCDS73919.1	NP_001284695.1	33,586

The table shows features of the four transcripts of the human gene *K0513* (HGNC accession no. 29058) including a description, gene and protein sizes and accession numbers. *K0513* is located on chromosome 16: 85,027,751-85,094,230 forward strand. This gene has 238 orthologues (Ensembl; Cunningham *et al.*, 2019). The predicted *K0513* transcript variants X1 and X2 are not included in this table; they both code for isoform a.

1097 S2 Table

PROTEIN	ALIGNMENT CONVERGE OF K0513	% IDENTITY	TEMPLATE INFORMATION
c2jobA	Residues 378 - 404	19	PDB header:lipid binding protein Chain: A: PDB Molecule:antilipopolsaccharide factor; PDTitle: solution structure of an antilipopolsaccharide factor from2 shrimp and its possible lipid a binding site
c2d9jA	Residues 262 - 310	16	PDB header:signaling protein Chain: A: PDB Molecule:regulator of g-protein signalling 7; PDTitle: solution structure of the rgs domain of regulator of g-2 protein signaling 7
c2af0A	Residues 103 - 145	20	PDB header:signaling protein Chain: A: PDB Molecule:regulator of g-protein signaling 2; PDTitle: structure of the regulator of g-protein signaling domain of2 rgs2
d2j9ua1		27	PDB header:transport protein Chain: A: PDB Molecule:vacuolar protein sorting- associated protein 54; PDTitle: c-terminal domain of vps54 subunit of the garp complex
d2j9ua1		15	Fold:Four-helical up-and-down bundle Superfamily:VPS28 C-terminal domain-like Family:VPS28 C-terminal domain-like

1098 The table shows selected proteins that align partially to regions of human K0513 predicted
1099 using Phyre², their sequence identities, and a description of each protein.

1100

1101

1102

1103

1104

1105

1106

1107

1108

1109

1110 S3 Table

UNIPROT ID	NAME	DESCRIPTION	PMIDS
Q9NTI2	ATP8A2	Phospholipid-transporting ATPase IB, Catalytic component of a P4-ATPase flippase complex	25402006
P60880	SNAP25	Synaptosomal-associated protein 25, t-SNARE involved in the molecular regulation of neurotransmitter release	25402006
Q2M2I8	AAK1	AP2-associated protein kinase 1, Regulates clathrin-mediated endocytosis	25402006
P42356	PI4KA	Phosphatidylinositol 4-kinase alpha, Acts on phosphatidylinositol (PtdIns) in the first committed step in the production of the second messenger inositol-1,4,5,-triphosphate	25402006
Q9NWB1	RBFOX1	RNA binding protein fox-1 homolog 1, RNA-binding protein that regulates alternative splicing events	25402006
P14867	GABRA1	Gamma-aminobutyric acid receptor subunit alpha-1, Ligand-gated chloride channel which is a component of the heteropentameric receptor for GABA	25402006
P09471	GNAO1	Guanine nucleotide-binding protein G(o) subunit alpha, involved as modulators or transducers in various transmembrane signaling systems	25402006
P61764	STXBP1	Syntaxin-binding protein 1, Participates in the regulation of synaptic vesicle docking and fusion through interaction with GTP-binding proteins	25402006
P37840	SNCA	Alpha-synuclein, Neuronal protein that plays several roles in synaptic activity	25402006
Q59EK9	RUNDC3A	RUN domain-containing protein 3A, May act as an effector of RAP2A in neuronal cells	25402006
Q70YC5	ZNF365	Protein ZNF365, Involved in the regulation of neurogenesis	25402006
Q9BWQ8	FAIM2	Protein lifeguard 2, Antiapoptotic protein which protects cells uniquely from Fas-induced apoptosis	25402006
Q9UI12	ATP6V1H	V-type proton ATPase subunit H, ubunit of the peripheral V1 complex of vacuolar ATPase	25402006
Q9UM19	HPCAL4	Hippocalcin-like protein 4, May be involved in the calcium-dependent regulation of rhodopsin phosphorylation	25402006
P78357	CNTNAP1	Contactin-associated protein 1, Required, with CNTNAP2, for radial and longitudinal organization of myelinated axons	23023127
Q9NZU7	CABP1	Calcium-binding protein 1, Modulates calcium-dependent activity of inositol 1,4,5-triphosphate receptors	25402006

Q9UPR5	SLC8A2	Sodium/calcium exchanger 2, Mediates the electrogenic exchange of Ca ²⁺ against Na ⁺ ions across the cell membrane	25402006
Q9UI15	TAGLN3	Transgelin-3, central nervous system development	25402006
Q9UBL0	ARPP21	cAMP-regulated phosphoprotein 21, may act as a competitive inhibitor of calmodulin-dependent enzymes	25402006
P20336	RAB3A	Ras-related protein Rab-3A, Small GTP-binding protein that plays a central role in regulated exocytosis and secretion	25402006
P17600	SYN1	Synapsin-1, Neuronal phosphoprotein that coats synaptic vesicles and binds to the cytoskeleton	25402006
O60359	CACNG3	Voltage-dependent calcium channel gamma-3 subunit, Regulates the trafficking to the somatodendritic compartment and gating properties of AMPA-selective glutamate receptors (AMPA receptors)	25402006
P08247	SYP	Synaptophysin, Possibly involved in structural functions as organizing other membrane components or in targeting the vesicles to the plasma membrane	25402006
Q3SXP7	SHISAL1	Protein shisa-like-1,	25402006
P62760	VSNL1	Visinin-like protein 1, Regulates (in vitro) the inhibition of rhodopsin phosphorylation in a calcium-dependent manner	25402006
Q99784	OLFM1	Noelin, Contributes to the regulation of axonal growth in the embryonic and adult central nervous system	25402006
Q96HW7	INTS4	Integrator complex subunit 4, Component of the Integrator (INT) complex, a complex involved in the small nuclear RNAs (snRNA) U1 and U2 transcription	17010949;21836163
O75493	CA11	Carbonic anhydrase-related protein 11	25402006
P85037	FOXK1	Forkhead box protein K1, Transcriptional regulator involved in different processes such as glucose metabolism, aerobic glycolysis, muscle cell differentiation and autophagy	25609649
O95741	CPNE6	Copine-6, Calcium-dependent phospholipid-binding protein that plays a role in calcium-mediated intracellular processes	25402006
Q9UPA5	BSN	Protein bassoon, Scaffold protein of the presynaptic cytomatrix at the active zone (CAZ)	25402006
Q92581	SLC9A6	Sodium/hydrogen exchanger 6, Electroneutral exchange of protons for Na ⁺ and K ⁺ across the early and recycling endosome membranes	25402006
P21579	SYT1	Synaptotagmin-1, Calcium sensor that participates in triggering neurotransmitter release at the synapse	25402006

Q9UPV7	PHF24	PHD finger protein 24	25402006
Q9P2U7	SLC17A7	Vesicular glutamate transporter 1, Mediates the uptake of glutamate into synaptic vesicles at presynaptic nerve terminals of excitatory neural cells	25402006
Q7L0J3	SV2A	Synaptic vesicle glycoprotein 2A, Plays a role in the control of regulated secretion in neural and endocrine cells	25402006
Q9Y4E6	WDR7	WD repeat-containing protein 7	25402006
Q05193	DNM1	Dynamin-1, Microtubule-associated force-producing protein involved in producing microtubule bundles and able to bind and hydrolyse GTP	25402006
Q16799	RTN1	Reticulon-1, May be involved in neuroendocrine secretion or in membrane trafficking in neuroendocrine cells	25402006
Q9UPP2	IQSEC3	IQ motif and SEC7 domain-containing protein 3, Acts as a guanine nucleotide exchange factor (GEF) for ARF1	25402006
Q01167	FOXK2	Forkhead box protein K2, Transcriptional regulator involved in different processes such as glucose metabolism, aerobic glycolysis and autophagy	25609649
Q15818	NPTX1	Neuronal pentraxin-1, May be involved in mediating uptake of synaptic material during synapse remodelling	25402006
Q92561	PHYHIP	Phytanoyl-CoA hydroxylase-interacting protein, Its interaction with PHYH suggests a role in the development of the central system	25402006
O60307	MAST3	Microtubule-associated serine/threonine-protein kinase 3	25402006
Q8TDI0	CHD5	Chromodomain-helicase-DNA-binding protein 5, Chromatin-remodeling protein that binds DNA through histones and regulates gene transcription	25402006
Q13367	AP3B2	AP-3 complex subunit beta-2, Subunit of non-clathrin- and clathrin-associated adaptor protein complex 3 (AP-3)	25402006
Q13387	MAPK8IP2	C-Jun-amino-terminal kinase-interacting protein 2, the JNK-interacting protein (JIP) group of scaffold proteins selectively mediates JNK signaling	25402006
Q9BR01	SULT4A1	Sulfotransferase 4A1, Atypical sulfotransferase family member with very low affinity for 3'-phospho-5'-adenylyl sulfate (PAPS)	25402006
O60641	SNAP91	Clathrin coat assembly protein AP180, Adaptins are components of the adapter complexes which link clathrin to receptors in coated vesicles	25402006

Q9UPP5	KIAA1107	AP2-interacting clathrin-endocytosis protein, Involved in clathrin-mediated endocytosis at the synapse	25402006
Q7L1I2	SV2B	Synaptic vesicle glycoprotein 2B, probably plays a role in the control of regulated secretion in neural and endocrine cells	25402006
Q9NYX4	CALY	Neuron-specific vesicular protein calcyon, nteracts with clathrin light chain A and stimulates clathrin self-assembly and clathrin-mediated endocytosis	25402006
P46459	NSF	Vesicle-fusing ATPase, Required for vesicle-mediated transport	25402006
P07196	NEFL	Neurofilament light polypeptide, contain three intermediate filament proteins: L, M, and H which are involved in the maintenance of neuronal caliber	25402006
P49418	AMPH	Amphiphysin, May participate in mechanisms of regulated exocytosis in synapses and certain endocrine cell types	25402006
Q16143	SNCB	Beta-synuclein, Non-amyloid component of senile plaques found in Alzheimer disease	25402006
Q08209	PPP3CA	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform, Calcium-dependent, calmodulin-stimulated protein phosphatase	25402006
Q8IX03	WWC1	Protein KIBRA, plays a pivotal role in tumor suppression by restricting proliferation and promoting apoptosis	17010949;21836163
P63027	VAMP2	Vesicle-associated membrane protein 2, Involved in the targeting and/or fusion of transport vesicles to their target membrane	25402006
Q9P2S2	NRXN2	Neurexin-2, Neuronal cell surface protein that may be involved in cell recognition and cell adhesion	23023127
Q96BY2	MOAP1	Modulator of apoptosis 1, Required for death receptor-dependent apoptosis	25402006
Q99819	ARHGDIG	Rho GDP-dissociation inhibitor 3, Inhibits GDP/GTP exchange reaction of RhoB	25402006
P05771	PRKCB	Protein kinase C beta type, Calcium-activated, phospholipid- and diacylglycerol (DAG)-dependent serine/threonine-protein kinase	25402006
Q9H2X9	SLC12A5	Solute carrier family 12 member 5, Mediates electroneutral potassium-chloride cotransport in mature neurons and is required for neuronal Cl ⁻ homeostasis	25402006
P05067	APP	Amyloid-beta precursor protein, Functions as a cell surface receptor	21832049
O60268	DYNC111	Cytoplasmic dynein 1 intermediate chain 1, Acts as one of several non-catalytic accessory components of the cytoplasmic dynein 1 complex	25402006

O60268	SNPH	Syntaphilin, Inhibits SNARE complex formation by absorbing free syntaxin-1	25402006
O60268	EPHB6	Ephrin type-B receptor 6, Kinase-defective receptor for members of the ephrin-B family	25402006
O60268	B4GAT1	Beta-1,4-glucuronyltransferase 1, Beta-1,4-glucuronyltransferase involved in O-mannosylation of alpha-dystroglycan (DAG1)	25402006
O60268	HAX1	HCLS1-associated protein X-1, Recruits the Arp2/3 complex to the cell cortex and regulates reorganization of the cortical actin cytoskeleton	17010949;21836163

1111

1112

1113

1114

1115

1116

1117

1118

1119

1120

1121

1122

1123

1124

1125

1126

1127

1128

1129

1130

1131

1132

1133 S1 Fig

1145 S2 Fig

```

      .....10.....20.....30.....40.....50.....60.....70.....
AA      METPEVPVGLIDFGPEAPTSSPLEAPPVLQDGDGSLGDGASESETTESADSENDMGESPSHPSWDQDRRSSNESFS
DB_state
DB_conf

      80.....90.....100.....110.....120.....130.....140.....150.....
AA      SNQSTESTQDEETLALRDFMRGYVEKIFSGGEDLDQEEKAKFGGEYCSENGKGREWFARYVSAQRCNSKCVSEATFYRL
DB_state                                0                                0    0
DB_conf                                2                                4    5

      .160.....170.....180.....190.....200.....210.....220.....230.....
AA      VQSFVAVLFECHQMDDFGPAKNLMTMCFTYYHIGKPQLLPESREKPAGSIDSYLKSANSWLAEEKKDIAERLLKNTSAR
DB_state                                0                                0
DB_conf                                6                                5

      ..240.....250.....260.....270.....280.....290.....300.....310....
AA      TENVKGFFGGLETCLKGPLARRNEEDENKPKQEKRPRAVTAYSPEDKKGEKIYLYTHLKQQPIWHTLRFWNAAFFDAVH
DB_state
DB_conf

      ...320.....330.....340.....350.....360.....370.....380.....390...
AA      CERTKRSPTRGDAGEEEEKREKWCHMTQEERDDSLRFNENITFGQLGTFTHNMLAFGLNKKLCNDFLKKQAVIGNLDE
DB_state 0                                0                                0
DB_conf 6                                6                                5

      ....400.....410
AA      EQYKLLSDHIEQMATE
DB_state
DB_conf

```

1146

1147

1148