MOONLIGHTING PROTEINS AND HUMAN DISEASES

Universitat Autònoma de Barcelona

MIREIA JIMÉNEZ ROSÉS

DEGREE IN BIOLOGY, UNIVERSITAT AUTÒNOMA DE BARCELONA , BELLATERRA, BARCELONA

1. INTRODUCTION

Multifunctional or multitasking proteins, also named moonlighting proteins, are proteins with more than one function. The first function ever disclosed of the protein is named canonical function and the other discovered lately are moonlighting functions.

These proteins complicate the interpretation of knock-outs/ knock-ins, DNA arrays, metabolomics and systems biology analyses, drug targeting, drug pharmacokinetics, pharmacodynamics and toxicity assays, etc [1,2].

Generally, they are discovered by serendipity while investigating other issues, but it would be useful to find other methods to to identify moonlighting proteins. An approach is bioinformatics analyses, although the final demonstration remains experimental. The bioinformatics approach has been the objective of the group of Bioinformatics for Molecular Biology, which supervises this End-of-Degree Project [3, 4] which has also created the first database of these proteins [5].

There are some examples of proteins in which the second function is not a "normal" function and are involved in different diseases (Neomorphic Moonlighting Function, gain-of-toxic-function) [6].

Thus, the complex phenotypes of several disorders may be related to the involvement of moonlighting proteins. A major challenge associated with therapeutic manipulation of moonlight protein is designing a drug which inhibits the function involved in the pathology while minimizing disruption of the other functions [2].

2. OBJECTIVES

The objectives of the Final Degree Project are:

A. to compile those moonlighting proteins associated with genetic diseases or disorders, which appear because of a mutation that occur in the human gene of each protein, and to join, as an update, into the database of multitasking protein (multitaskProtDB), where researchers will be able to find out quickly and easily the information they are looking for.

B. Identify and specify which proteins are known drug targets and, if possible, what different kind of side effect could be attributable to the different functions they have.

3. MATERIAL AND METHODS

PubMed (www.ncbi.nlm.nih.gov/pubmed)
Following: moonlighting, multitasking, bioinformatics, disease, disorders, drug targeting and drug side effects.

Databases:

- MultitaskProtDB (http://wallace.uab.es/multitask/)
- Online Mendelian Inheritance in Man (OMIM) (http://www.ncbi.nlm.nih.gov/omim/)
- Human Gene Mutation Database" (HGMD, http://www.hgmd.org/)
- The Drug Gene Interaction Database (http://dgidb.genome.wustl.edu/)
- Side Effect Resource (SIDER) http://sideeffects.embl.de/

We have used only the *Homo sapiens* moonlighting proteins and it has been built a table with the disorders related, the drugs and the side effects

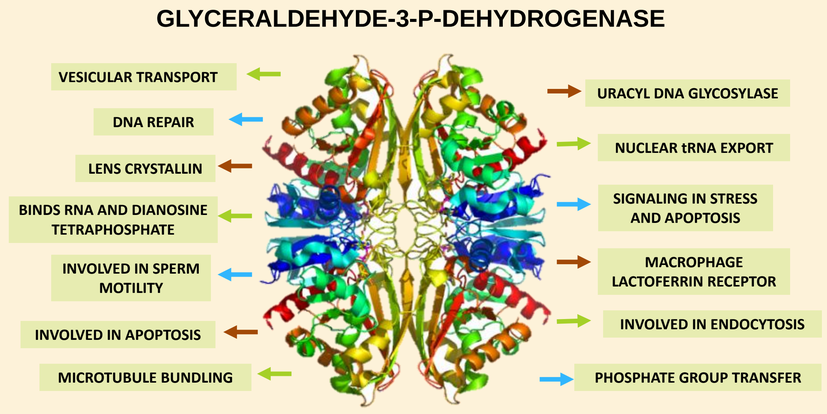
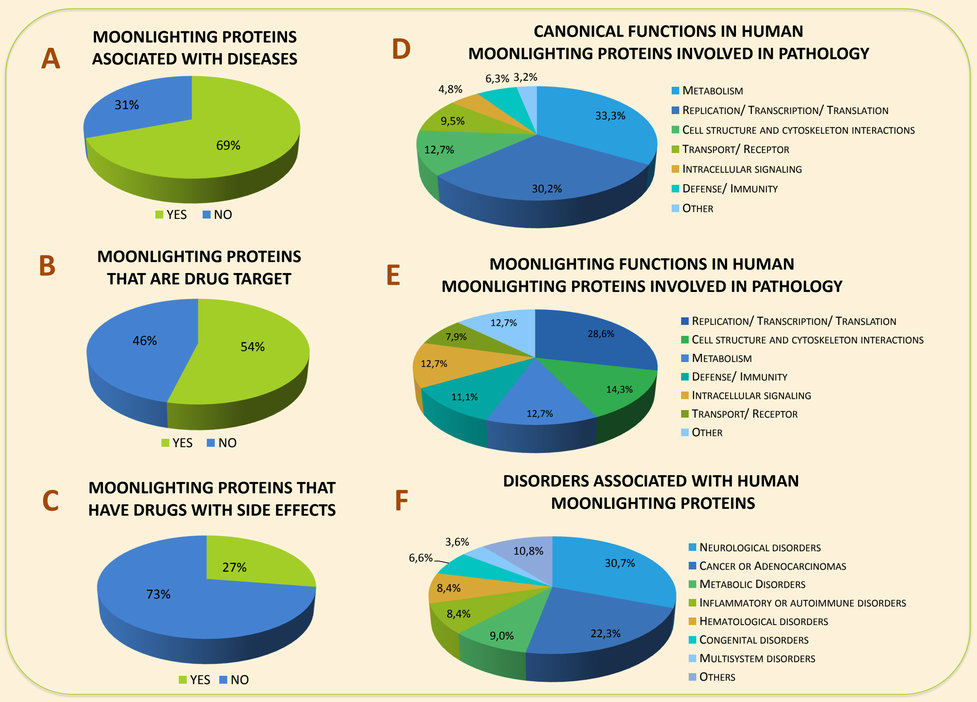


Figure 1. Glyceraldehyde-3P-dehydrogenase, a moonlighting protein with 13 functions.

4. RESULTS

Table 1. Some of the 91 human moonlighting proteins that are found in the database MultitaskProtDB.

ID	NCBI Code	UniProt Code	Protein Name	Canonical Function	Moonlighting Function	Organism	PDB	Oligomeric State	Reference
1	AAD19351	Q71UF1	Aconitase EC:4.2.1.3	Catalyses the stereo-specific isomerization of citrate to isocitrate via cis-aco More ...	Doom homeostasis / IREBP (cytosol), mtDNA maintenance (mitochondrion)	Homo sapiens			8041788
4	P21399	P21399	Cytoplasmatic aconitate hydratase/IRP1 EC:4.2.1.3	Catalyses the stereo-specific isomerization of citrate to isocitrate via cis-aco More ...	mRNA binding protein	Homo sapiens	2B3X		17698960
5	O00337	O00337	hCNT1 (Sodium/nucleoside cotransporter 1)	Nucleosides Transport (selective for pyrimidine nucleosides and adenosine)	Inhibition of tumor growth (likely to be relevant in tumor biology)	Homo sapiens			23722537
7	AAI07699	P15336	ATF2 protein (Cyclic AMP-dependent transcription factor) EC:2.3.1.48	Transcription factor (stimulates CRE (cAMP responsive element)-dependent transcr More...	DNA damage response	Homo sapiens	1BHI		15916964
8	NP061820	P99999	Cytochrome c	It transfers electrons between Complexes III (Coenzyme Q - Cyt C reductase) and More ...	Apoptosis	Various (Homo sapiens)	3NVV		15907471
9	AAB01381	P09622	DLD (Dihydrolipoyl dehydrogenase, mitochondrial) EC:1.8.1.4	Lipoamide dehydrogenase is a component of the glycine cleavage system as well as More ...	Protease	Homo sapiens	3DL2		17404228
10	AAA58459	P28482	ERK2 (signal-regulated kinases) EC:2.7.11.24	Mitogen-activated protein kinase 1 (MAP kinase)	Transcriptional repressor	Homo sapiens	3TEI		19879846
30	NP_647541	P07323	Enolase (Gamma enolase) EC:4.2.1.11	Conversion of 2-phosphoglycerate to phosphoenolpyruvate (Glycolysis)	It has neurotrophic and neuroprotective effects on rather a broad spectrum of ne More...	Rattus norvegicus (& Homo sapiens)		Mammalian enolase is composed of 3 isozyme subunits, alpha, beta and gamma, whic More ...	7753509
74	AAN04009	Q9BPW9	hRoDH-E2 (Dehydrogenase/reductase SDR family member 9)	Retinoid and hydroxy-steroid Dehydrogenase/Reductase	Transcriptional repressor	Homo sapiens		Homotetramer	16691198
75	AAA51885	P05455	La protein (Lupus La protein)	Protect RNA from 3'-end digestion	RNA folding chaperone	Homo sapiens	1S79		16799560
78	AAB29184	Q99217	Amelogenin	Regulates the size and shape of mineral crystallites	Mitochondrial DNA maintenance	Homo sapiens			16674683
79	AAH32695	P36969	Gpx4 - Phospholipid hydroperoxide glutathione peroxidase, mitochondrial (Glutath More ...	Peroxidase (Protects from radiation and oxidative damage)	Formation of the mitochondrial capsule during sperm maturation	Homo sapiens	2OBI	Monomer	17630701
84	NP_001171651	P06744	Phosphoglucose isomerase EC:5.3.1.9	Catalyzes the conversion of glucose-6-phosphate into fructose 6-phosphate (glyco More ...	Neuroleukin, autocrine motility factor, differentiation and maturation factor / More ...	Homo sapiens	1JLH	Homodimer in the catalytically active form, monomer in the secreted form	3352744
88	P19971	P19971	Thymidine phosphorylase EC:2.4.2.4	Phosphorylates thymidine to produce thymine and 2-deoxy-alpha-D-ribose 1-phospha More ...	Platelet-derived endothelial cell growth factor	Homo sapiens	2WKS	Homodimer	1570012
89	AAC12921	O14786	Neuropilin (VEGF receptor)	Vascular endothelial cell growth factor receptor	Receptor for semaphorin III (nerve axons) / VEFGR, regulation of angiogenesis	Homo sapiens	4DEQ	Homodimer, and heterodimer with NRP2.	9529250
91	CAA25833	P04406	Glyceraldehyde-3-phosphate dehydrogenase EC:1.2.1.12	Conversion of D-glyceraldehyde 3-phosphate (G3P) into 3-phospho-D-glyceroyl phos More ...	Uracil-DNA glycosylase, Transferrin receptor, involved in endocytosis, microtubu More ...	Homo sapiens	1U8F	Homotetramer	15877277
104	NP000526	P54278	PMS2 mismatch-repair enzyme	PMS2 mismatch-repair enzyme	Hypermutation of antibody variable chains	Homo sapiens	1H7S	Heterodimer of PMS2 and MLH1	9469811
105	P13569	P13569	CFTR chloride channel (Cystic fibrosis transmembrane conductance regulator) EC:3 More ...	Involved in the transport of chloride ions	Regulator of other epitelial anion channels (Regulator of Na+ channels)	Homo sapiens	3ISW		7543698
148	AAA52576	P32189	Glycerol Kinase EC:2.7.1.30	It converts glycerol to sn-glycerol 3-phosphate (glycolysis)	ATP-stimulated translocation protein (ASTP), porin binding, apoptosis, lysosome More ...	Homo sapiens			8884278
149	AAC50455	Q13584	Isocitrate dehydrogenase EC:1.1.1.42	TCA cycle enzyme	Mitochondrial mRNA binding	Homo sapiens			7505425
151	AAD17940	P53597	Succinyl-coA synthetase (Succinate thiokinase) EC:6.2.1.4	TCA cycle enzyme	Mitochondrial DNA maintenance	Homo sapiens		Heterodimer of an alpha and a beta subunit	15692048
152	P21912.3	P21912	Succinate dehydrogenase EC:1.3.99.1	TCA cycle enzyme	Tumor supression (it should play specific tumor-suppressing roles, directly or i More ...	Homo Sapiens		Component of complex II composed of four subunits	11404820
153	AAH17444	P07954	Fumarate hydratase EC:4.2.1.2	Catalyzes the reversible hydration/dehydration of fumarate to malate (TCA cycle)	Tumor suppressor	Homo sapiens	3E04	Homotetramer	11865300
154	AAA36154	P13010	Ku70/Ku80	Single stranded DNA-dependent ATP-dependent helicase. Has a role in chromosome t More ...	Extracellular protease (function at the cell surface is likely to be important i More ...	Homo sapiens	1Q2Z	Heterodimer	15738653
155	AAF15890	Q9UH91	Xanthine oxidoreductase	Purine catabolism enzyme	Structural role in the mammary gland	Homo sapiens			12502743
156	AAL48323	P18074	ERCC2 - TFIIH basal transcription factor complex helicase XPD subunit	DNA helicase involved in the repair of DNA damaged by exposure to ultraviolet li More ...	It is also a subunit of TFIIH, a basal transcription factor	Homo sapiens		One of the six subunits forming the core of TFIIH	14726016



Figures 2. Percentage of moonlighting proteins that have a disease or diseases-associated (**A**), that are drug target (**B**) and that have drugs with side effects (**C**). Distribution of the different types of canonical functions (**D**) or moonlighting functions (**E**) in the human moonlighting proteins associated with diseases. Distribution of the different types of disorders in the human moonlighting proteins (**F**).

5. CONCLUSIONS

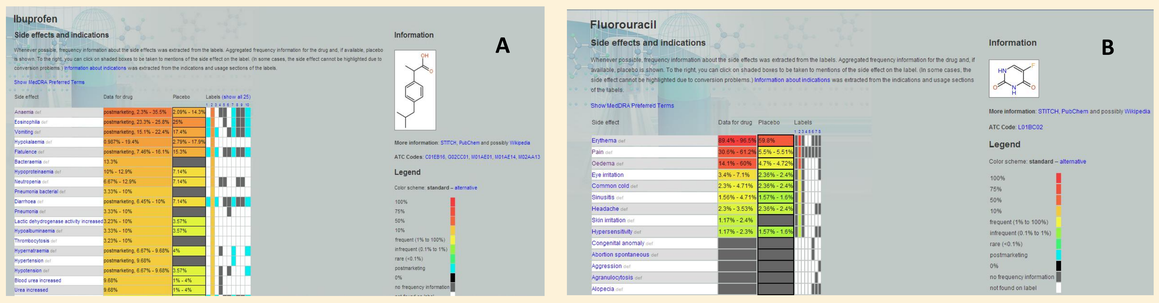
- From 288 moonlighting proteins in the database *MultitaskProtDB,* only 91 of them are human proteins. That means around 30% of all multitasking proteins which have been described until nowadays. There are 69% of them which have diseases associated or have susceptibility to a disorder. These suggest that moonlighting proteins are more prone to be involved in human disease than the average. Although unknown what percentage of the estimated 1-5 million proteins in the human proteome it is difficult to think in about 70% of them involved in pathology.
- Of these human moonlighting proteins involved in pathology, more than 75% are proteins whose canonical function is involved in the metabolism or in the replication, transcription and translation or in the cell structure or cytoskeleton interactions, whilst these functions as moonlighting ones are found in 55%. Concluding that it has been an increase of the other functions and also from the category "others" from 3,2% to 12,7%, where are included the proteins with more than one category of functions.
- According to the clinical disorder, 62% of the diseases can be classified into neurological disorders, different ks of cancer and metabolic disorders (30,7%, 22,3% and 9%, respectively).
- 54% of these 91 proteins are drug targets, what mean that there are 49 proteins on which may act different drugs in which there are also included the substrates.
- But, up to now, only 27% of the human moonlighting proteins have been associated side effects related to some of the drugs that act in the protein. Future work will be gaining insight into the side effects at molecular level.

In conclusion, with the information that has been compiled of moonlighting proteins, we can contribute to identify molecular mechanisms of diseases, and to understand why some side effects appear linked to some drugs.

The above shown results will be included in the next update of the mentioned MultitaskProtDB database.

Table 2. Disease, drugs and side effects links of the human moonlighting proteins that appear in Table 1.

PROTEIN NAME	UNIPROT	DISEASE	DRUG TARGET	SIDE EFFECTS
Aconitase	Q71UF1	Infantile cerebellar-retinal degeneration (MIM number 614559)	4-hydroxy-aconitate ion Aconitase ion Alpha-methylisocitric acid Tricarballylic acid	
Cytoplasmatic aconitate hydratase/IRP1	P21399	Association with myophathy (MIM number 255125)		
hCNT1 (Sodium/nucleoside cotransporter 1)	O00337	Concentrative nucleoside transporter deficiency	Anthracyclines and related substances	
ATF2 protein (Cyclic AMP-dependent transcription factor)	P15336			
Cytochrome c	P99999	Thrombocytopaenia 4 (MIM number 612004)	Heme Heme C Minocycline (negative modulador) N-trimethyllysine Protoporphyrin IX containing Co or Zn Zinc substituted heme C	sideeffects.embl.de/drugs/4200/
DLD (Dyhidrolipoyl dehydrogenase, mitochondrial)	P09622	Dihydrolipoamide dehydrogenase deficiency (MIM number 246900)	Flavin-adenine dinucleotide NADH Nicotinamide-adenine-dinucleotide	
ERK (signal-regulated kinases, MAPK1)	P28482	Truncus arteriosus (MIM number 217095)	Arsenic trioxide (inducer) Isoproterenol (inducer) Sorafenib ERK2.txt	sideeffects.embl.de/drugs/14888/ sideeffects.embl.de/drugs/3779/ sideeffects.embl.de/drugs/216239/
Enolase (Gamma enolase)	P09104		2-Phosphoglycolic acid	
hRoDH-E2 (Dehydrogenase/ reductase SDR family member 9)	Q9BPW9			
La protein (Lupus La protein)	P05455	Association with Sjogren syndrome (MIM number 270150) Association with systemic lupus erythematosus (SLE; MIM number 152700)		
Amelogenin	Q99217	Amelogenesis imperfecta, hypoplastic/ hypomaturation type 1E (MIM number 301200)		
Gpx4 (Glutathione peroxidase-4)	P36969	Increased 5-lipoxygenase metabolism	Glutathione (cofactor)	
Phosphoglucose isomerase	P06744	Hemolytic anemia, nonspherocytic, due to glucose phosphate isomerase deficiency (MIM number 613470) PGI deficiency, autosomal recessive (MIM number 172400)	5-Phospho-D-arabinohydroxamic acid 5-Phsphoarabinonic acid 6-Phosphogluconic acid Alpha-D-glucose-phosphate Erythose-4-phosphate Fructose-6-phosphate Glucose-6-phosphate Sorbitol 6-phosphate	
Thymidine phosphorylase	P19971	Mitochondrial DNA depletion syndrome 1 (MNGIE type; MIM number 603041) Peripheral neuropathy with intestinal pseudo-occlusion	5-Chloro-6-[(2-iminopyrrolidin-1-yl)methyl]pyrimidine-2,4(1H,3H)-dione Fluorouracil	sideeffects.embl.de/drugs/3385/
Neuropilin (VEGF receptor)	O14786		Anti-NRP1 (inhibitor) Palifermin Pegaptanib	sideeffects.embl.de/drugs/158786/
Glyceraldehyde-3-phosphate dehydrogenase	P04406	Huntington disease (MIM number 143100) Alzheimer disease (MIM number 104300 and 104310) Parkinson disease (MIM number168601) Brain ischemia	4-(2-Aminoethyl)benzenesulfonyl fluoride Adenosine-5-diphosphoribose NADH Nicotinamide-adenine-dinucleotide Thiocotinamide-adenine-dinucleotide	
PMS2 mismatch-repair enzyme	P54278	Colorectal cancer, hereditary nonpolyposis, type 4 (MIM number 614337) Mismatch repair cancer Syndrome (MIM number 276300) Colorectal adenoma Defective protein-protein interaction with MLH1 Neuroectodermal tumors and cafe-au-lait spots	Adenosine-5'-diphosphate Phosphothiophosphoric acid-adenylate ester	
CFTR chloride channel (Cystic fibrosis transmembrane conductance regulator)	P13569	Congenital bilateral absence of vas deferens (MIM number 277180) Cystic fibrosis (MIM number 219700) Sweat chloride elevation without CF Modifier of Bronchiectasis with or without elevated sweat chloride 1 (MIM number 211400) Neonatal hypertrypsinemia; Idiopathic Pancreatitis (MIM number 167800) Fetal hyperechogenic bowel Primary sclerosing cholangitis Association with respiratory/pancreatic disease Association with Typhoid fever (protection against)	Adenosine-5'-diphosphate Bumetanide (antagonist) Glyburide (antagonist) Ivacaftor (potentiator) Phosphoaminophosphonic acid-adenylate ester Phosphonoserine	sideeffects.embl.de/drugs/2471/ sideeffects.embl.de/drugs/3488/
Glycerol Kinase	P32189	Glycerol kinase deficiency (or hyperglycerolemia; MIM number 307030)		
Isocitrate dehydrogenase	Q13584	D-2-hydroxyglutaric aciduria 2 (MIM number 613657)	Isocitric acid	
Succinyl-coA synthetase (Succinate thiokinase)	P53597	Mitochondrial DNA depletion syndrome 9 (encephalomyopathic type with methylmalon aciduria; MIM number 245400) Succinyl-coenzyme A synthetase deficiency	Succinic acid	
Succinate dehydrogenase	O14521	Cowden syndrome 3 (MIM number 615106) Carcinoid tumors, intestinal (MIM number 114900) Merkel cell carcinoma, somatic Paraganglioma and gastric stromal sarcoma (MIM number 606864) Paragangliomas 1, with or without deafness (MIM number 168000) Pheochromocytoma (MIM number 171300)	2-Hexyloxy-6-hydroxymethyl-tetrahydro-pyran-3,4,5-triol Hexachlorophene (inhibitor) Succinic acid Ubiquinone-1	sideeffects.embl.de/drugs/3598/
Fumarate hydratase	P07954	Fumarase deficiency (MIM number 606812) Leiomyomatosis and renal cell cancer (HLRCC; MIM number 150800) Leydig cell tumour of the testis Ovarian mucinous cystadenoma		
Ku70/Ku80	P13010	Association with Bladder cancer (increased risk)	Thalidomide	sideeffects.embl.de/drugs/5426/
Xanthine oxidoreductase	Q9UH91	Xanthinuria type I (or XDH Deficiency; MIM number 278300)	4-[5-PYRIDIN-4-YL-1H-[1,2,4]TRIAZOL-3-YL]-PYRIDINE-2-CARBONITRILE Allopurinol Dioxothuomolybdenum (VI) ion Doxorubicin Febuxostat Flavin-adenine dinucleotide L-carnitine Leptin Y-700	sideeffects.embl.de/drugs/2094/ - sideeffects.embl.de/drugs/1690/ sideeffects.embl.de/drugs/134018/ sideeffects.embl.de/drugs/85/
ERCC2 - TFIIH basal transcription factor complex helicase XPD subunit	P18074	Cerebrooculafacioskeletal syndrome 2 (MIM number 610756) Trichothiodystrophy (MIM number 601675) Xeroderma pigmentosum group D (MIM number 278730) Association with oligodendroglioma development (reduce risk) Cockayne syndrome (MIM number 216400)	Doxorubicin Ifosfamide Leucovorin Methotrexate Oxaliplatin Platinum compound Vitamin C	sideeffects.embl.de/drugs/1690/ sideeffects.embl.de/drugs/3690/ sideeffects.embl.de/drugs/143/ sideeffects.embl.de/drugs/4112/ sideeffects.embl.de/drugs/4609/



Figures 3. Side effects which appears in *Side Effect Resource Webside* (SIDER). (**A**) Ibuprofen, a drug that acts in cyclooxygenase-1 gene (COX-1). (**B**) Fluorouracil, a drug that acts in thymidine phosphorylase (marked with a brown arrow), thymidylate synthase and endothelial nitric oxide synthase.

REFERENCES:

[1] Copley SD (2012) Moonlighting is mainstream: paradigm adjustment required. Bioessays, 34:578-588.
[2] Gómez A, Domedel N, Cedano J, Piñol J, Querol E. (2003) Do current sequence analysis algorithms disclose multifunctional (moonlighting) proteins?. Bioinformatics, 19:895-896.
[3] Gómez A, Hernández S, Amela I, Piñol J, Cedano J, Querol E. (2011) Do protein-protein interaction databases identify moonlighting proteins? Mol Biosyst, 7:2379-2382.
[4] Jeffery CJ (2011) Proteins with neomorphic moonlighting functions in disease. IUBMB Life, 63:489-494.
[5] Hernández S, Ferragut G, Amela I, Perez-Pons J, Piñol J, Mozo-Villarias A, Cedano J, Querol E. (2014) MultitaskProtDB: a database of multitasking proteins. Nucleic Acids Res., 42: D517-520.
[6] Butler GS, Overall CM. (2009) Proteomic identification of multitasking proteins in unexpected locations complicates drug targeting. Nat Rev Drug Discov, 8:935-948

TUTORS: Enric Querol and Isaac Amela
Institute of Biothecnology and Biomedicine (IBB), Departament of Biochemistry and Molecular Biology, Universitat Autònoma de Barcelona