

MoonProt: a database for proteins that are known to moonlight

Mathew Mani¹, Chang Chen¹, Vaishak Ambler², Haipeng Liu³, Tanu Mathur², Grant Zwicke², Shadi Zabad⁴, Bansi Patel², Jagravi Thakkar² and Constance J. Jeffery^{1,2,*}

¹Department of Bioengineering, University of Illinois at Chicago, Chicago, IL 60607, USA, ²Department of Biological Sciences, MC567, University of Illinois at Chicago, 900 S. Ashland Ave., Chicago, IL 60607, USA, ³Center for Pharmaceutical Biotechnology, College of Pharmacy, University of Illinois at Chicago, 900 S. Ashland Ave., Chicago, IL 60607, USA and ⁴Illinois Institute of Technology, 3300 S Federal St, Chicago, IL 60616, USA

Received June 18, 2014; Revised September 18, 2014; Accepted September 30, 2014

ABSTRACT

Moonlighting proteins comprise a class of multifunctional proteins in which a single polypeptide chain performs multiple biochemical functions that are not due to gene fusions, multiple RNA splice variants or pleiotropic effects. The known moonlighting proteins perform a variety of diverse functions in many different cell types and species, and information about their structures and functions is scattered in many publications. We have constructed the manually curated, searchable, internet-based MoonProt Database (<http://www.moonlightingproteins.org>) with information about the over 200 proteins that have been experimentally verified to be moonlighting proteins. The availability of this organized information provides a more complete picture of what is currently known about moonlighting proteins. The database will also aid researchers in other fields, including determining the functions of genes identified in genome sequencing projects, interpreting data from proteomics projects and annotating protein sequence and structural databases. In addition, information about the structures and functions of moonlighting proteins can be helpful in understanding how novel protein functional sites evolved on an ancient protein scaffold, which can also help in the design of proteins with novel functions.

INTRODUCTION

It has been known for many years that some soluble enzymes have a second function as structural proteins in the lens of the eye (crystallins) (1,2) or in binding to DNA or RNA to regulate translation or transcription (3). These proteins belong to a subclass of multifunctional proteins called

‘moonlighting proteins’ (4), in which more than one biochemical or biophysical function is contained within one polypeptide chain. In this subclass of multifunctional proteins, the multiple functions are not due to gene fusions, multiple RNA splice variants or multiple proteolytic fragments. Moonlighting proteins do not include pleiotropic proteins, where a protein has multiple downstream cellular roles in different pathways or physiological processes that result from a single biochemical function of a protein. Moonlighting proteins also do not include families of homologous proteins, if the different functions are performed by different members of the protein family.

Over 200 proteins have been found to moonlight, and it is possible that many other proteins also have additional functions that have not yet been found. The known examples of moonlighting proteins include diverse types of proteins, including receptors, enzymes, transcription factors, adhesins and scaffolds [for reviews see (3–16)]. Moonlighting proteins are expressed in many cell types and function in many different biochemical pathways. They are found in mammals, yeast, bacteria, plants and other organisms. Different combinations of functions are observed, for example, a cytoplasmic enzymatic function and an extracellular cytokine function are found in phosphoglucose isomerase/autocrine motility factor (17–20).

Some moonlighting proteins can perform both functions simultaneously, but for others, the functions of a moonlighting protein can vary due to changes occurring within the cell. A moonlighting protein can adjust to these changes by having different functions in different cellular locations, in different cell types, by interacting with other proteins in various oligomeric states or when it senses a change in the cellular concentration of a ligand, substrate, cofactor or a product of an enzymatic reaction (4). Two proposed mechanisms/models (1,4) for the evolution of a moonlighting function make use of general features of protein structure and could apply to many protein types, potentially resulting in a large number of proteins that moonlight. The

*To whom correspondence should be addressed. Tel: +312 996 3168; Fax: +312 413 2691; Email: cjeffery@uic.edu

diverse examples of moonlighting proteins already identified, the methods by which one protein can moonlight, the two proposed models for the evolution of moonlighting functions and the potential benefits moonlighting proteins might provide through coordinating cellular activities, suggest that moonlighting proteins are likely to be common.

The ability of a single protein to have multiple, sometimes unrelated, functions and to be involved in different multiprotein complexes or biochemical pathways can make more complicated predicting functions of the many genes identified through genome sequencing projects, interpreting results from proteomics experiments and annotating protein structure and sequence databases. In proteomics studies, the additional functions can result in complicated protein expression and interaction patterns and make it more difficult to determine if 'false' positives are truly false or perhaps instead due to a second function of a protein. Moonlighting proteins complicate the interpretation and annotation of genome sequences because sequence homologs can share all, none, one or some functions. For example, the delta 1 and delta 2 crystallins in the duck lens share 94% amino acid sequence activity, but only the delta 2 isoform retains arginosuccinate lyase catalytic activity (21–24). In systems biology, moonlighting adds another level to our understanding of the complex but organized cellular protein network. For example, a moonlighting protein can provide a means of communication and coordination between the many pathways within a cell or a mechanism to switch between pathways, helping the cell to respond to changes in its environment. A moonlighting protein can also be involved in communication and coordination between different cell types within an organism.

Information about the structures and functions of moonlighting proteins can be helpful in studies of protein evolution, for example, in understanding how novel protein functional sites evolved on an ancient protein scaffold. This information can also be useful for the design of proteins with novel functions because the structures and functions of moonlighting proteins can serve as a guide of how to add new functional site to the fold of a stable protein. In moonlighting proteins nature has created new functions on old protein scaffolds many times, and for some proteins it is clear which function was the original functional site on a protein structure and which functional site evolved later.

For the reasons listed above, it would be useful to have a way to identify moonlighting proteins, but there is currently no straightforward method to identify which proteins have moonlighting functions or for determining if a protein of interest moonlights. The lack of an identified common physical feature has prevented a means of quickly identifying other moonlighting proteins. In addition, characterization of a novel protein generally involves finding a function for a protein, but it does not necessarily include a search for all possible additional functions of a protein. For most moonlighting proteins a second function has been found by serendipity.

Information about moonlighting proteins is scattered in many publications. There is currently some confusion in the literature because some authors have used the term 'moonlighting' when the publication is actually describing a protein family in which different members have different func-

tions, or a protein that has the same function but in two different locations (two biological roles but not two biochemical functions), or a protein that is another type of multifunctional protein. The presence of a protein in an unexpected location can suggest that the protein has a second function, but additional evidence that the protein truly performs two different biochemical functions in the two locations is needed before concluding that a protein moonlights.

The MoonProt Database provides a centralized, searchable, web-based database for knowledge about the known moonlighting proteins. It contains information about the sequences, structures and functions for over 200 moonlighting proteins for which multiple biochemical functions have been experimentally verified within one protein. The MoonProt database is also the first database that is strictly devoted to moonlighting proteins and not other types of multifunctional proteins. It differs from another recent resource that includes multifunctional or 'multitasking' proteins that have multiple functions due to gene fusions, proteins that have a pleiotropic effect on multiple pathways and/or 'multifunctional proteins' for which the two functions are not exhibited by a single protein but instead by different protein products within a multiprotein complex (25).

In contrast, for the MoonProt Database, we have carefully selected only proteins that have two different biochemical or biophysical functions, such as being a crystallin in the lens of the eye in addition to being an enzyme with catalytic activity. This specific requirement adds to the value of the MoonProt database and is particularly important for some studies of protein structure and function, protein evolution, protein function prediction and the design of proteins with novel proteins.

The MoonProt Database was created in order to be a resource for researchers and for educational purposes. The availability of this organized, curated information provides a more complete picture of what is currently known about moonlighting proteins. This repository of information integrates information scattered across the literature and provides links to other databases. It helps avoid the need of searching through hundreds of journal articles. It will also serve as a resource for the future development of methods to identify other moonlighting proteins. The database will be important for many labs that are working on determining the functions of genes identified in the genome sequencing projects, interpreting data from proteomics projects, studying the evolution of new functions on proteins, designing proteins with novel functions and annotating protein sequence and structural databases. Also, researchers in other fields, such as cell biology or genetics, will be able to look up the proteins they are studying and see if they are known to moonlight.

MATERIALS AND METHODS

Selection of moonlighting proteins included in the database

The goal of creating the MoonProt Database is to bring together information about the experimentally validated moonlighting proteins. We manually collected information from peer-reviewed journals with an emphasis on finding references describing experimental evidence of multiple biochemical and/or biophysical functions within one polypep-

tide chain. Published biochemical, mutagenic or other evidence to support the presence of multiple functions was required and was critically reviewed by the Principle Investigator (PI) before the protein was accepted for inclusion in the MoonProt Database. Ideally, independence of both functions has been shown by demonstrating that inactivation of one function (i.e. through mutagenesis) does not inactivate the second function, and vice versa.

Proteins were not included if they belong to a different class of multifunctional proteins, for example, if the 'multiple functions' are due to gene fusions, different RNA splice variants, the same function in two different locations, pleiotropic affects on multiple pathways or multiple physiological processes or a family of proteins in which the different functions are performed by different proteins. We also have not included proteins for which the 'multiple functions' are simply different aspects of the same function (i.e. 'membrane protein' and 'transmembrane channel'). In this version 1.0 of the database, we also include a list of 'potentially moonlighting proteins', for which a second function is only suggested by the results of gene knockout or other genetics experiments, in the FAQs page to illustrate examples of additional proteins that might also moonlight.

Information included about individual proteins

Information about each protein was manually curated from published journal articles and online resources (Protein Data Bank (PDB), Uniprot, etc.). In some cases, the curators communicated with the labs that are experts in studying the individual proteins for assistance in finding some of the information (for example, finding the correct amino acid sequence, identifying which species was used in the experiments or checking the accuracy in our descriptions of the biochemical functions). The protein-specific entries include a description of each biochemical or biophysical function and a list of citations for evidence of that function. Alternative names are included because many moonlighting proteins have multiple names relating to their multiple functions. Including the alternative names aids in finding a protein of interest in a text search of the database. The specific species in which each protein has two or more functions was identified; it is especially important that the correct species and amino acid sequence is identified for the moonlighting protein and not for a homolog that might or might not have both functions. UniProtKB (26) or Pubmed [<http://www.ncbi.nlm.nih.gov/pubmed/>] resources were used to identify the correct amino acid sequences. The sequences are included in FASTA format. Basic Local Alignment Search Tool [<http://blast.ncbi.nlm.nih.gov/Blast.cgi>] was used to identify structures in the PDB (27) that correspond to the amino acid sequence, if available. Protein structures are not listed if they belong to a homolog from a different species. Gene Ontology (GO) terms (28) were identified from the UniProtKB Database (26) in order to illustrate the different types of proteins annotated in the MoonProt Database. UniProtKB (26) entry IDs and Enzyme Commission (EC) numbers are included for easy connection to external resources. When available, information is included about the cellular location in which the protein exhibits each function.

Members of the Jeffery laboratory will be in charge of updating the database and verifying the multiple functions of additional moonlighting proteins that are continuing to be identified. It is also anticipated that more information about the proteins on the list will become available, including new protein structures, additional functions, new information about functional sites on the protein structures, additional evidence of each function (i.e. additional biochemical or biophysical experiments), as well as the corresponding references.

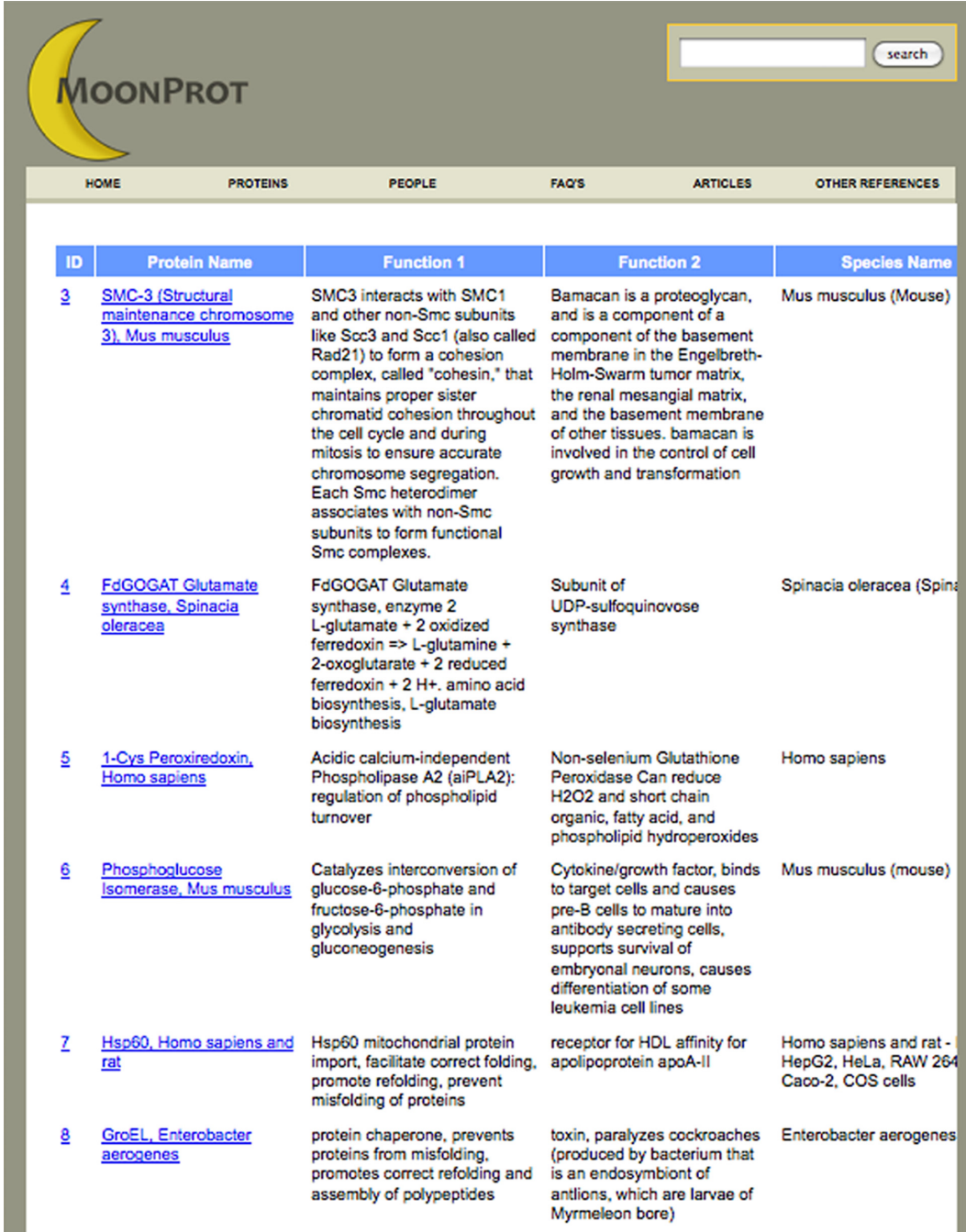
DATABASE ARCHITECTURE AND WEB INTERFACE

The MoonProt Database is based on a MySQL database for storing data (<http://www.mysql.com/>) and uses PHP 5.2.11 (<http://www.php.net/>). The web pages are written in Hypertext Markup Language and provide a user-friendly interface for locating and retrieving information from the database.

The moonlighting protein database includes a total of seven types of pages: Home, Proteins, People, FAQ's, Articles, Other References and a Protein Details page that can be used to bring up the annotated information about each moonlighting protein in the database. The website features a prominent header that makes it easy to explore and discover relevant information. In the header, there is a navigation bar that provides an easy way to navigate between the main sections of the website, as well as a text search box that makes it easy for the user to query and retrieve specific information from the database. The Home page provides a summary about moonlighting proteins and a dynamic illustration of several moonlighting proteins and their functions. The Proteins page (Figure 1) provides a list of all the proteins included in the database. From the Proteins page, the user can select a protein and bring up the annotated information for that protein in the Protein Details page (Figure 2). The annotated information for each protein contains the names of the protein, a UniProt accession number, GO terms from the UniProt database, the species of organism for which both functions have been demonstrated to be in one protein, the length of the amino acid sequence, the amino acid sequence in FASTA format, PDB IDs for any available protein structures in the PDB, descriptions of at least two functions, references for experiments demonstrating the functions and EC numbers for enzymes. Additional information is included for some proteins, including cellular location where that function is observed, and information about quaternary structure. The FAQ's page includes frequently asked questions about moonlighting proteins and information about how to contact the database curators. The Articles page includes a list of review articles about moonlighting proteins. The Other References page includes references to the UniProtKB, PubMed and other resources used in the creation of the database.

RESULTS AND CONCLUSIONS

The MoonProt Database brings together information about over 200 moonlighting proteins in a centralized, web-based resource. It includes proteins from species throughout the evolutionary tree, archaea, bacteria and eukaryotes (mammals, plants, reptiles, insects, worms, protozoa, etc.). Re-



ID	Protein Name	Function 1	Function 2	Species Name
3	SMC-3 (Structural maintenance chromosome 3), <i>Mus musculus</i>	SMC3 interacts with SMC1 and other non-Smc subunits like Scc3 and Scc1 (also called Rad21) to form a cohesion complex, called "cohesin," that maintains proper sister chromatid cohesion throughout the cell cycle and during mitosis to ensure accurate chromosome segregation. Each Smc heterodimer associates with non-Smc subunits to form functional Smc complexes.	Bamacan is a proteoglycan, and is a component of a component of the basement membrane in the Engelbreth-Holm-Swarm tumor matrix, the renal mesangial matrix, and the basement membrane of other tissues. bamacan is involved in the control of cell growth and transformation	<i>Mus musculus</i> (Mouse)
4	FdGOGAT Glutamate synthase, <i>Spinacia oleracea</i>	FdGOGAT Glutamate synthase, enzyme 2 L-glutamate + 2 oxidized ferredoxin => L-glutamine + 2-oxoglutarate + 2 reduced ferredoxin + 2 H+. amino acid biosynthesis, L-glutamate biosynthesis	Subunit of UDP-sulfoquinovose synthase	<i>Spinacia oleracea</i> (Spinach)
5	1-Cys Peroxiredoxin, <i>Homo sapiens</i>	Acidic calcium-independent Phospholipase A2 (aiPLA2): regulation of phospholipid turnover	Non-selenium Glutathione Peroxidase Can reduce H2O2 and short chain organic, fatty acid, and phospholipid hydroperoxides	<i>Homo sapiens</i>
6	Phosphoglucose isomerase, <i>Mus musculus</i>	Catalyzes interconversion of glucose-6-phosphate and fructose-6-phosphate in glycolysis and gluconeogenesis	Cytokine/growth factor, binds to target cells and causes pre-B cells to mature into antibody secreting cells, supports survival of embryonal neurons, causes differentiation of some leukemia cell lines	<i>Mus musculus</i> (mouse)
7	Hsp60, <i>Homo sapiens</i> and <i>rat</i>	Hsp60 mitochondrial protein import, facilitate correct folding, promote refolding, prevent misfolding of proteins	receptor for HDL affinity for apolipoprotein apoA-II	<i>Homo sapiens</i> and <i>rat</i> - HepG2, HeLa, RAW 264 Caco-2, COS cells
8	GroEL, <i>Enterobacter aerogenes</i>	protein chaperone, prevents proteins from misfolding, promotes correct refolding and assembly of polypeptides	toxin, paralyzes cockroaches (produced by bacterium that is an endosymbiont of antlions, which are larvae of Myrmeleon bore)	<i>Enterobacter aerogenes</i>

Figure 1. The Proteins Page enables users to browse a list of all the proteins annotated in the MoonProt Database. Version 1.0 of the database contains over 200 moonlighting proteins.

ceptors, enzymes, transcription factors, adhesins, chaperones and scaffolds, and many other types of proteins are included.

Having information about confirmed moonlighting proteins together in one place and separate from other types of multifunctional proteins (i.e. gene fusions) will facilitate studies of their sequences, structures and functions. The proteins in the database can also serve as a test set for testing algorithms and programs to predict protein functions in general and for predicting which other proteins might

moonlight. For example, because of the different physical arrangement of the functional sites in moonlighting proteins versus gene fusion proteins, a program may score well in predicting from sequence or structure the two functions of a gene fusion protein but not identify the multiple functions of a moonlighting protein. The information in the database is already being used in tests of protein function prediction methods.

The database can also serve as a set of examples of proteins and data (protein sequences, functions and in

General Information	
MoonProt ID:	6
First appeared in release:	1.0
Name(s):	Phosphoglucose isomerase GPI GPI AMF NLK Glucose-6-phosphate isomerase Autocrine motility factor Neuroleukin Differentiation and maturation mediator Phosphohexose isomerase PHI
UniProt:	P06745 (GPI_MOUSE) Reviewed
GO terms:	GO:0001525 angiogenesis GO:0001701 in utero embryonic development GO:0001707 mesoderm formation GO:0006994 gluconeogenesis GO:0006096 glycolytic process GO:0034101 erythrocyte homeostasis GO:0042593 glucose homeostasis GO:0004347 glucose-6-phosphate isomerase activity GO:0005125 cytokine activity GO:0005515 protein binding GO:0008083 growth factor activity GO:0016853 isomerase activity GO:0005576 extracellular region GO:0005615 extracellular space GO:0005737 cytoplasm GO:0005886 plasma membrane GO:0006170 ciliary membrane
Organism(s) for which both functions have been demonstrated:	Mus musculus (mouse)
Sequence length:	558
Quaternary structure:	Dimer of identical subunits
FASTA sequence:	>sp P06745 GPI_MOUSE Glucose-6-phosphate isomerase OS=Mus musculus GN=Gpi PE=1 SV=4 MALTRNPPQKLELWHRKNSANLKLRELFEADPERFN NLSLNLNTHLHLYDYKSNLV NKEVMQMLVELAKSRGVEAARDNMFSGKSNYENRAV LHWLRNRSNTPKVGGKDVMP EVNVRVLDKMKSFQVRVSGDWGKGTGKISDIIINIGGG SDLGPLMTEALKPKYSKGGPR YWFVSNIDGTHAKTLASLSPELFIASKFTTQETITNA ETAKEVFLKAKDPSAIA KHFVLSLNTAKVKEGIDPQNMFEFWDWVGGRYSLVWS AIGLSIALHVGDFHFEQLLSGA HWMDDQHLKTPLEKNAPVLLALLGWYNYCYGETHALL PYDQYMRFAAYFQGGDMESN GKYITKSGARVDHQTGPVWGEPTNGQAFYQLIHOG TKMIPCDFLIPVQGHPRKGL HKVLLANFLADTEALMKGKLPPEARKEKLAQAGSPEDL EKLLPHKVFEGGRPTNSIVET KLPFILGALIAMYEHKIFVQGMWDINSFDQWVGLGKQ LAKKIEPELGGSAVTSHDS STNGLISFKQQRDTKLE
Structure Information	
PDB ID:	1U0E, 1U0F, 1U0G
Information about functions	
One Function:	Catalyzes interconversion of glucose-6-phosphate and fructose-6-phosphate in glycolysis and gluconeogenesis
References for function:	Lohmann, K. (1933) <i>Biochem. Z.</i> 262, 137. Rose, I. A., and O'Connell, E. L. (1961) <i>J. Biol. Chem.</i> 236, 3368-3392. Rose, I. A. (1975) <i>Adv. Enzymol. Relat. Areas Mol. Biol.</i> 43, 491-517. Rose, I. A. (1981) <i>Philos. Trans. R. Soc. London, Ser. B</i> 293, 131-143. Willem, R., Malaisse-Lagae, F., Ottinger, R., and Malaisse, W. J. (1990) <i>Biochem. J.</i> 265, 519-524.
E.C. number:	5.3.1.9
Location of functional site(s) and reference(s) for that site:	Active site pocket containing His388, Glu357, Arg272, Lys518, Thr214, Ser209, Thr211, Ser159
Cellular Location of Function:	Cytoplasm
Comments:	Multistep catalytic reaction involves ring opening followed by isomerization step, strict specificity for glucose-6-phosphate and fructose-6-phosphate
Another Function:	Cytokine/growth factor: binds to target cells and causes pre-B cells to mature into antibody secreting cells, supports survival of embryonal neurons, causes differentiation of some leukemia cell lines
References for function:	Chaput, M., Cless, V., Portetelle, D., Cludts, I., Cravador, A., Burny, A., Gras, H., and Tartar, A. (1988) <i>Nature</i> 332, 454-455. Faik, P., Walker, J. I. H., Redmil, A. A. M., and Morgan, M. J. (1988) <i>Nature</i> 332, 455-457. Watanabe, H., Takehana, K., Date, M., Shinozaki, T., and Raz, A. (1996) <i>Cancer Res.</i> 56, 2960-2963. Xu, W., Selter, K., Feldman, E., Ahmed, T., and Chiao, J. W. (1996) <i>Blood</i> 87, 4502-4506. Gurney, M. E., Heinrich, S. P., Lee, M. R., and Yin, H. S. (1986) <i>Science</i> 234, 566-574. Gurney, M. E., Apatoff, B. R., Spear, G. T., Baumel, M. J., Antel, J. P., Bania, M. B., and Roder, A. T. (1986) <i>Science</i> 234, 574-581. Watanabe, H., Carmi, P., Hogan, V., Raz, T., Silletti, S., Nabi, I. R., and Raz, A. (1991) <i>J. Biol. Chem.</i> 266, 13442-13448.
E.C. number:	N/A
Location of functional site(s) and reference(s) for that site:	
Cellular Location of Function:	Extracellular
Comments:	

Figure 2. The Protein Details page brings up annotation for a selected protein that was manually curated to include alternative names of the protein, information about its biochemical and/or biophysical functions, citations describing experimental evidence of the functions, the specific species in which the protein has two or more functions, the amino acid sequence in FASTA format, PDB IDs of any available structures, GO terms from the UniProtKB resource (23), EC numbers for enzyme functions and other information.

many cases structures) to aid in developing hypotheses about which characteristics, motifs or other features might be shared by moonlighting proteins, or some subtypes of moonlighting proteins, that could potentially be used in developing new methods to identify moonlighting proteins. In addition, by comparing the list of proteins with the results of proteomics studies that compare protein–protein interactions, protein expression levels or protein cellular locations, an interaction, expression or location pattern might be identified for proteins that moonlight that can be used to predict which other proteins in those studies might also moonlight.

In proteomics studies, the MoonProt Database can also be useful in another way. Because moonlighting proteins are often found in more than one cellular location or expressed in more than one cell type or interact with multiple binding partners, they might help explain unexpected protein localization, protein expression or protein–protein interaction results. Proteins identified in the proteomics experiments can be compared to the list of known moonlighting proteins to see if the presence of a second function can help explain the unexpected results.

The MoonProt Database may also be valuable to researchers who identify a protein of interest in cell biology, genetics or other areas of biology or medicine. Information about all the known functions of a protein might help in identifying the specific function, and the associated biochemical pathway or multifunctional complex, that might be involved in a physiological process or disease being studied.

Finally, the MoonProt Database includes links to the 3D structures of 115 moonlighting proteins in the PDB. These include proteins that have been adopted for a second function without significant changes to sequence or structure, proteins that have gained a second binding site during evolution, proteins that undergo conformational changes to switch between function and examples of other ways in which a protein can have two different functions. These structures, as well as the amino acid sequences for these and many other moonlighting proteins, can provide interesting examples for studying the evolution of new protein functions, and they can provide examples and potential guides for designing and regulating the functions of proteins with novel functions,

The Jeffery lab is continuing to review the literature to add additional moonlighting proteins for future updates of the database as well as to add additional information and references for the proteins included in version 1.0. We welcome suggestions of additional moonlighting proteins and information about those proteins from colleagues.

AVAILABILITY

The MoonProt Database is freely available via a user-friendly graphical user interface at the web address www.moonlightingproteins.org. The interface enables text search for a protein name, species, amino acid sequence or a UniProtKB or PDB identifier. The user can also browse a list of all the proteins in the database. The database is ‘read and search only’ by the public, but additional information about the known moonlighting proteins and suggestions of other proteins that might also be moonlighting are welcome

and can be sent to the curators for possible inclusion in the database.

FUNDING

UIC Faculty Scholarship Support Program Award [to C.J.J.]. Funding for open access charge: University of Illinois at Chicago Faculty Scholarship Support Program Award [to C.J.J.]; Research Open Access [ROAAP] Publishing Fund of the University of Illinois at Chicago.

Conflict of interest statement. None declared.

REFERENCES

- Piatigorsky, J. and Wistow, G.J. (1989) Enzyme/crystallins: gene sharing as an evolutionary strategy. *Cell*, **57**, 197–199.
- Wistow, G. and Piatigorsky, J. (1987) Recruitment of enzymes as lens structural proteins. *Science*, **236**, 1554–1556.
- Commichau, F.M. and Stülke, J. (2008) Trigger enzymes: bifunctional proteins active in metabolism and in controlling gene expression. *Mol. Microbiol.*, **67**, 692–702.
- Jeffery, C.J. (1999) Moonlighting proteins. *Trends Biochem. Sci.*, **24**, 8–11.
- Piatigorsky, J. (2007) *Gene Sharing and Evolution*. Harvard University Press, Cambridge, MA.
- Henderson, B. and Martin, A. (2011) Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infect. Immun.*, **79**, 3476–3491.
- Henderson, B. and Martin, A. (2013) Bacterial moonlighting proteins and bacterial virulence. *Curr. Top Microbiol. Immunol.*, **358**, 155–213.
- Gancedo, C. and Flores, C.L. (2008) Moonlighting proteins in yeasts. *Microbiol. Mol. Biol. Rev.*, **72**, 197–210.
- Huberts, D.H. and van der Klei, I.J. (2010) Moonlighting proteins: an intriguing mode of multitasking. *Biochim. Biophys. Acta*, **1803**, 520–525.
- Jeffery, C.J. (2003) Moonlighting proteins: old proteins learning new tricks. *Trends Genet.*, **19**, 415–417.
- Jeffery, C.J. (2003) Multifunctional proteins: examples of gene sharing. *Ann. Med.*, **35**, 28–35.
- Jeffery, C.J. (2004) Molecular mechanisms for multitasking: recent crystal structures of moonlighting proteins. *Curr. Opin. Struct. Biol.*, **14**, 663–668.
- Jeffery, C.J. (2004) Moonlighting proteins: complications and implications for proteomics research. *Drug Discov. Today Targets*, **3**, 71–78.
- Jeffery, C.J. (2009) Moonlighting proteins—an update. *Mol. Biosyst.*, **5**, 345–350.
- Nobeli, I., Favia, A.D. and Wool, I.G. (1996) Extraribosomal functions of ribosomal proteins. *Trends Biochem. Sci.*, **21**, 164–165.
- Guo, M. and Schimmel, P. (2013) Essential nontranslational functions of tRNA synthetases. *Nat. Chem. Biol.*, **9**, 145–153.
- Gurney, M.E., Apatoff, B.R., Spear, G.T., Baumel, M.J., Antel, J.P., Bania, M.B. and Reder, A.T. (1986) Neuroleukin: a lymphokine product of lectin-stimulated T cells. *Science*, **234**, 574–581.
- Gurney, M.E., Heinrich, S.P., Lee, M.R. and Yin, H.S. (1986) Molecular cloning and expression of neuroleukin, a neurotrophic factor for spinal and sensory neurons. *Science*, **234**, 566–574.
- Watanabe, H., Takehana, K., Date, M., Shinozaki, T. and Raz, A. (1996) Tumor cell autocrine motility factor is the neuroleukin/phosphohexose isomerase polypeptide. *Cancer Res.*, **56**, 2960–2963.
- Xu, W., Seiter, K., Feldman, E., Ahmed, T. and Chiao, J.W. (1996) The differentiation and maturation mediator for human myeloid leukemia cells shares homology with neuroleukin or phosphoglucose isomerase. *Blood*, **87**, 4502–4506.
- Barbosa, P., Wistow, G.J., Cialkowski, M., Piatigorsky, J. and O'Brien, W.E. (1991) Expression of duck lens delta-crystallin cDNAs in yeast and bacterial hosts. Delta 2-crystallin is an active argininosuccinate lyase. *J. Biol. Chem.*, **266**, 22319–22322.
- Chiou, S.H., Lo, C.H., Chang, C.Y., Itoh, T., Kaji, H. and Samejima, T. (1991) Ostrich crystallins. Structural characterization of delta-crystallin with enzymic activity. *Biochem. J.*, **273**, 295–300.
- Piatigorsky, J., O'Brien, W.E., Norman, B.L., Kalumuck, K., Wistow, G.J., Borras, T., Nickerson, J.M. and Wawrousek, E.F. (1988) Gene sharing by delta-crystallin and argininosuccinate lyase. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 3479–3483.
- Piatigorsky, J. and Horwitz, J. (1996) Characterization and enzyme activity of argininosuccinate lyase/delta-crystallin of the embryonic duck lens. *Biochim. Biophys. Acta.*, **1295**, 158–164.
- Hernández, S., Ferragut, G., Amela, I., Perez-Pons, J., Piñol, J., Mozo-Villarias, A., Cedano, J. and Querol, E. (2014) MultitaskProtDB: a database of multitasking proteins. *Nucleic Acids Res.*, **42**, D517–D520.
- The UniProt Consortium. (2013) Update on activities at the universal protein resource (UniProt) in 2013. *Nucleic Acids Res.* **41**, D43–D47.
- Rose, P.W., Bi, C., Bluhm, W.F., Christie, C.H., Dimitropoulos, D., Dutta, S., Green, R.K., Goodsell, D.S., Prlic, A., Quesada, M. *et al.* (2013) The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res.* **41**, D475–D482.
- Gene Ontology Consortium. (2013) Gene ontology annotations and resources. *Nucleic Acids Res.*, **41**, D530–D535.