BIOINFORMATICS APPLICATIONS NOTE Vol. 20 no. 18 2004, pages 3652–3653
doi:10.1093/bioinformatics/bth399

PLPMDB: Pyridoxal-5′-phosphate dependent enzymes mutants database

Paolo Di Giovine

Dipartimento di Scienze Biochimiche A. Rossi Fanelli, Universita degli Studi di Roma La Sapienza, Piazzale Aldo Moro, 5, 00185 Rome, Italy

Received on March 14, 2004; revised on May 28, 2004; accepted on July 1, 2004
Advance Access publication July 9, 2004

ABSTRACT

Summary: The searchable mutant database PLPMDB has been developed to provide rapid and simple access to relevant mutant information on pyridoxal-5′-phosphate dependent enzymes. All data have been extracted from publications and publicly available databases, then organized in a relational database to enable searching via a web-based search form. The current version of PLPMDB contains 688 mutants described in 220 research papers. The database is a useful tool for planning mutant experiments and for interpretation of information from such experiments.

Availability: PLPMDB is freely accessible from http://www.studiofmp.com/plpmdb/index.htm

Contact: paolo.digiovine@studiofmp.com

Pyridoxal-5′-phosphate-dependent enzymes constitute a ubiquitous class of biocatalysts. They catalyze a variety of reactions involving amino acids, including isomerizations, transaminations, racemizations, aldol cleavage, decarboxylations, and side chain eliminations or replacements (Christen and Mehta, 2001). Mutation studies are an important source of data on the role of individual amino acids and together with structural data, they play a central role in the rational drug design process. A large amount of data on structural and functional features of pyridoxal-5′-phosphate-dependent enzymes have been developed by the creation, expression and subsequent characterization of mutants (Hayashi, 1995). Several mutation information resources are available. At this point, mutant data on pyridoxal-5′-phosphate-dependent enzymes are distributed in several databases and hundreds of research papers. In order to facilitate access to these data, they have been extracted from publications and publicly available database, organized in a relational database and made accessible to the research community via a web-based search form. SWISS-PROT (Bairoch and Apweiler, 2000) contains sequences of about 1400 pyridoxal-5′-phosphate-dependent enzymes. From their annotation, information about 194 variants and 42 mutagens have been extracted using a Perl script. Other mutation information have been manually extracted from Protein Mutant Database (Kawabata *et al.*, 1999), on-line Mendelian inheritance in man (OMIM) (McKusick, 1998) and papers identified by manual literature searches in PubMed (http://www.ncbi.nlm.nih.gov/). Table 1 summarizes the resources on the basis of the PLPMDB database. Sequences and three-dimensional (3D) structures of pyridoxal-5′-phosphate-dependent enzymes have been extracted from Swiss-Prot and the Protein Data Bank (Berman *et al.*, 2000) via SRS system (Zdobnov *et al.*, 2002). All data collected have been stored in a Microsoft Access relational database. A schematic diagram of the entity-relationship (ER) model of PLPMDB can be found at website http://www.studiofmp.com/plpmdb/ermodel.gif. To guarantee high-quality data, the pyridoxal-5′-phosphate-dependent enzymes that have been mutated have to be present in the annotated database Swiss-Prot. The PLPMDB contains 688 mutants described in 220 research papers. Each mutant entry contains references and hyperlinks to other databases. Table 2 summarizes the information stored per mutant with hyperlinks to the other Web resources. The website was made using HTML, Active Server Pages and SQL programming languages. The database covers all known types of pyridoxal-5′-phosphate-dependent enzymes. The papers dealing exclusively with descriptions of clinical findings related to splice variants and polymorphisms are included. A WWW-based form gives access to a system that allows the user to query the database for specific substitutions, reported biological effects, specific protein or species. Furthermore, bibliographic-oriented queries may be performed.

Table 1. Mutation information resources

Name	Mutants found	URL address
Swiss-Prot	236	http://www.expasy.org
OMIM	34	http://www.ncbi.nlm.nih.gov/omim/
PubMed	241	http://www.ncbi.nlm.nih.gov/
Protein Mutant Database	177	http://spock.genes.nig.ac.jp/~pmd/

Mutants collected in the PLPMDB database. The table holds the name of the resource, the number of mutants we found in the resource and its location in the WWW.

Table 2. Information per mutant

Field	Example	Description
Mutant Accession Number	16	Unique identifier
Protein Name	Serinehydroxymethyltransferase (EC 2.1.2.1) (Serine methylase) (SHMT)	Common name of the protein and its E.C. number
Swiss Prot Entry	GLYA_ECOLI	Link to Swiss-Prot entry
Organism	*Escherichia coli*	Species name
Gene Name	*GLYA*	Enzyme gene name
PMD entry	A880266	Link to Mutant entry in the Protein Mutation Database
Mutations	POINT His 228 Asn	Mutation notation
Variant/Mutagen	Mutagen	Indicates whether a mutation is a natural occurring variant or an experimental mutagen
Effects:	Catalytic activity [−]: affinity for both D-Ala and L-Ala [+]: form quinonoid complexes with both isomers at similar rates:	Description of the observed effect of the mutant
PubmedID:	3069126	Link to PubMed abstract
Title:	Serine hydroxymethyltransferase: mechanism of the racemization and transamination of D- and L-alanine	Title of the reference
Authors:	Shostak,K. and Schirch,V.	Authors
Reference:	Biochemistry (1988) **27**(21), 8007–8014	Reference
3D Structures	1EQB 1KL2	Link to protein structure in Protein Structure Databank

Left column, the name of the data field; middle column, example of actual data stored in PLPMDB; and right column, description of the field.

I have also prepared some lists that may be useful entry points into the database (list of papers referenced, list of all sequences and 3D structures, list of all mutations and mutants in the PLPMDB database). In summary, PLPMDB is a web-based bioinformatics tool for pyridoxal-5′-phosphate-dependent enzymes research that combines mutation information from different sources. The database is a useful tool for planning mutant experiments and for interpretation of information from such experiments. The approaches that I have used to create this database are generally applicable to building databases for mutagenesis studies of other protein families. Moreover, mutation information from the literature continues to be added manually on a monthly basis. Additional information and statistical data about PLPMDB are accessible at http://www.studiofmp.com/plpmdb/index.htm.

ACKNOWLEDGEMENTS

Author would like to thank Maria Nasso for her technical support.

REFERENCES

Bairoch,A. and Apweiler,R. (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.*, **28**, 45–48.

Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.

Christen,P. and Mehta,P.K. (2001) From cofactor to enzymes. The molecular evolution of pyridoxal-5′-phosphate-dependent enzymes. *Chem. Rec.*, **6**, 436–437.

Hayashi,H. (1995) Pyridoxal enzymes: mechanistic diversity and uniformity. *J. Biochem. (Tokyo)*, **3**, 463–473.

Kawabata,T., Ota,M. and Nishikawa,K. (1999) The Protein Mutant Database. *Nucleic Acids Res.*, **27**, 355–357.

McKusick,V.A. (1998) *Mendelian Inheritance in Man. Catalogs of Human Genes and Genetic Disorders*, 12th edn. The Johns Hopkins University Press, Baltimore, MD.

Zdobnov,E.M., Lopez,R., Apweiler,R. and Etzold,T. (2002) The EBI SRS server-new features. *Bioinformatics*, **8**, 1149–1150.