View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by RERO DOC Digital Library

310–311 *Nucleic Acids Research, 1999, Vol. 27, No. 1* *© 1999 Oxford University Press*

The ENZYME data bank in 1999

Amos Bairoch*

Swiss Institute of Bioinformatics, Centre Medical Universitaire, 1 rue Michel Servet, 1211 Geneva 4, Switzerland

Received October 7, 1998; Revised October 9, 1998; Accepted October 28, 1998

ABSTRACT

The ENZYME data bank is a repository of information related to the nomenclature of enzymes. In recent years it has become an indispensable resource for the development of metabolic databases. The current version contains information on 3704 enzymes. It is available through the ExPASy WWW server (http://www.expasy.ch/).

INTRODUCTION

The ENZYME data bank is a repository of information related to the nomenclature of enzymes. It is primarily based on the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB) (1) and it contains the following data for each type of characterized enzyme for which an EC (Enzyme Commission) number has been provided: (i) EC number; (ii) recommended name; (iii) alternative names (if any); (iv) catalytic activity; (v) cofactors (if any); (vi) pointers to the SWISS-PROT (2) protein sequence entrie(s) that correspond to the enzyme (if any); (vii) pointers to the PROSITE (3) entrie(s) describing the protein familie(s) of which the enzyme is a member (if any); (viii) pointers to human disease(s) (4) associated with a deficiency of the enzyme (if any).

We believe that the ENZYME data bank would be useful to anybody working with enzymes and that it can be of help in the development of computer programs involved in the manipulation of metabolic pathways. In recent years it has became an indispensable resource for the development of metabolic databases (5). Such databases typically describe collections of enzymes, reactions and biochemical pathways and are used in conjunction with software that allows the user to query and visualize metabolic information. They are used in various contexts and have gained recognition in the context of the reconstruction of metabolic pathways from the sequence of complete bacterial or archebacterial genomes (6,7).

The main source for the data in the ENZYME data bank comes from the recommendations of the IUBMB, but additional information has been extracted from the literature. Finally, it is important to note that the tight coupling that exists between the ENZYME database and the SWISS-PROT protein sequence database is of benefit to both resources as it allows updates and corrections to be propagated efficiently between them.

```
ID   1.14.17.3
DE   Peptidylglycine monooxygenase.
AN   Peptidyl alpha-amidating enzyme.
AN   Peptidylglycine 2-hydroxylase.
CA   PEPTIDYLGLYCINE + ASCORBATE + O(2) = PEPTIDYL(2-HYDROXYGLYCINE) +
CA   DEHYDROASCORBATE + H(2)O.
CF   Copper.
CC   -!- Peptidylglycines with a neutral amino acid residue in the penultimate
CC       position are the best substrates for the enzyme.
CC   -!- The enzyme also catalyses the dismutation of the product to
CC       glyoxylate and the corresponding desglycine peptide amide.
CC   -!- Involved in the final step of biosynthesis of alpha-melanotropin
CC       and related biologically active peptides.
PR   PROSITE; PDOC00080;
DR   P08478, AMD1_XENLA;  P12890, AMD2_XENLA;  P10731, AMD_BOVIN ;
DR   P19021, AMD_HUMAN ;  P97467, AMD_MOUSE ;  P14925, AMD_RAT   ;
//

ID   2.3.1.43
DE   Phosphatidylcholine--sterol O-acyltransferase.
AN   Lecithin--cholesterol acyltransferase.
AN   LCAT.
AN   Phospholipid--cholesterol acyltransferase.
CA   PHOSPHATIDYLCHOLINE + STEROL = STEROL ESTER +
CA   1-ACYLGLYCEROPHOSPHOCHOLINE.
CC   -!- Palmitoyl, oleoyl, and linoleoyl can be transferred; a number of
CC       sterols, including cholesterol, can act as acceptor.
CC   -!- The bacterial enzyme also catalyses the reactions of EC 3.1.1.4 and
CC       EC 3.1.1.5.
DI   NORUM DISEASE; MIM:245900.
DI   FISH-EYE DISEASE; MIM:136120.
PR   PROSITE; PDOC00110;
DR   P10480, GCAT_AERHY;  P53760, LCAT_CHICK;  P04180, LCAT_HUMAN;
DR   P16301, LCAT_MOUSE;  Q08758, LCAT_PAPAN;  P30930, LCAT_PIG  ;
DR   P53761, LCAT_RABIT;  P18424, LCAT_RAT  ;
//
```

Figure 1. Two sample ENZYME entries.

FORMAT

The entries in the database are structured so as to be usable by human readers as well as by computer programs. An entry in the database is composed of defined line types, each with its own format; they are used to record the various types of data which make up the entry. For standardization purposes the format of ENZYME follows as closely as possible that of the SWISS-PROT (2) protein and EMBL (8) nucleotide sequence databases. Two sample ENZYME entries are shown in Figure 1.

PRACTICAL INFORMATION

Content of the current release

Release 23.0 of ENZYME (July 1998) contains information on 3704 enzymes. The data file (ENZYME.DAT) requires ~1.5 Mb of disk storage space. The database is distributed with a user manual (ENZUSER.TXT); a file describing the various classes, subclasses and sub-subclasses of enzymes (ENZCLASS.TXT); and a file that describes how the database can be obtained (ENZYME.GET).

*Tel: +41 22 702 5477; Fax: +41 22 702 5502; Email: amos.bairoch@medecine.unige.ch

The present distribution frequency is four releases per year. No restrictions are placed on the use or redistribution of the data.

Interactive access to SWISS-PROT and TrEMBL

The most efficient and user-friendly way to browse interactively in ENZYME is to use the World-Wide Web (WWW) molecular biology server ExPASy (9). The ExPASy Web server was made available to the public in September 1993. In October 1998 a cumulative total of 33 million connections was attained. It may be accessed through its URL, which is: http://www.expasy.ch/

You can directly access the section of ExPASy that allows you to browse through the ENZYME database by opening the URL: http://www.expasy.ch/sprot/enzyme.html

The electronic version of the Boehringer Mannheim Biochemical Pathways Wallchart

The Biochemical Pathways Wallchart edited by retired Boehringer Mannheim researcher Dr Gerhard Michal (see http://biochem.boehringer-mannheim.com/techserv/metmap.htm), has a long tradition of prominence on the walls of life sciences laboratories. It consists of a graphical representation of the main metabolic pathways. We provide, on the ExPASy server, an electronic version of the chart as a series of linked images. Each enzyme mentioned in the chart is linked to its corresponding entry in ENZYME. The converse is also true. The Biochemical Pathways can be accessed from the URL: http://www.expasy.ch/cgi-bin/search-biochem-index

How to obtain ENZYME

If you have access to a computer system linked to the Internet you can obtain ENZYME using anonymous FTP (File Transfer Protocol), from the following servers: ExPASy (Expert Protein Analysis System), ftp.expasy.ch; EBI, ftp.ebi.ac.uk.

A version of the database in the ASN.1 data exchange format compatible with the databases and software developed by the National Center for Biotechnology Information (NCBI) (10) is also available on the above servers.

How to submit new data or updates/corrections to ENZYME

We do not assign EC numbers for newly characterized enzymes, this is the responsibility of the Nomenclature Committee of IUBMB (NC-IUBMB) (see http://www.chem.qmw.ac.uk/iupac/jcbn/). To contact the person responsible for the assignment of EC numbers in that committee one should write to: Prof. K. Tipton, Department of Biochemistry, Trinity College, Dublin 2, Republic of Ireland. Tel: +353 1 608 1608; Fax: +353 1 677 2400; Email: ktipton@tcd.ie.

The ENZYME data bank is distributed with a form that can be used to fill in the information necessary for the NC-IUBMB to assign an EC number. Such a form is also available from the Web at the URL: http://www.expasy.ch/sprot/enz_new_form.html

A separate form is available to send updates or corrections: http://www.expasy.ch/sprot/enz_update_form.html

The commission regularly sends us updates and additions to the nomenclature so that they can be integrated into the data bank in a timely manner.

REFERENCES

1 Enzyme Nomenclature, Recommendations of the Nomenclature Committe of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes, NC-IUBMB (1992) Academic Press, New York.
2 Bairoch,A. and Apweiler,R. (1998) *Nucleic Acids Res.*, **26**, 38–42.
3 Bairoch,A., Bucher,P. and Hofmann,K. (1997) *Nucleic Acids Res.*, **25**, 217–221.
4 Pearson,P.L., Francomano,C., Foster,P., Bocchini,C., Li,P. and McKusick,V.A. (1994) *Nucleic Acids Res.*, **22**, 3470–3473.
5 Karp,P. (1998) *Trends Biochem. Sci.*, **23**, 114–116.
6 Galperin,M.Y. and Brenner,S.E. (1998) *Trends Genet.*, **14**, 332–333.
7 Bono,H., Ogata,H., Goto,S. and Kanehisa,M. (1998) *Genome Res.*, **8**, 203–210.
8 Stoesser,G., Moseley,M.A., Sleep,J., McGowran,M., Garcia-Pastor,M. and Sterk,P. (1998) *Nucleic Acids Res.*, **26**, 8–15.
9 Appel,R.D., Bairoch,A. and Hochstrasser,D.F. (1994) *Trends Biochem. Sci.*, **19**, 258–260.
10 Benson,D.A., Boguski,M.S., Lipman,D.J., Ostell,J. and Ouellette,B.F.F. (1998) *Nucleic Acids Res.*, **26**, 1–7.