Nucleic Acids Research, 1996, Vol. 24, No. 1 221–222

The ENZYME data bank in 1995

Amos Bairoch

Department of Medical Biochemistry, University of Geneva, 1 rue Michel Servet, 1211 Geneva 4, Switzerland

Received October 3, 1995; Revised and Accepted October 13, 1995

ABSTRACT

The ENZYME data bank is a repository of information relative to the nomenclature of enzymes. The current version (October 1995) contains information relevant to 3594 enzymes. It is available from a variety of file and ftp servers as well as through the ExPASy World Wide Web server (http://expasy.hcuge.ch/).

INTRODUCTION

The ENZYME data bank is a repository of information relative 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 Enzyme Commission (EC) 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 human disease(s) associated with a deficiency of the enzyme (if any).

We believe that the ENZYME data bank can be useful to anybody working with enzymes and that it can be of help in the development of computer programs involved with the manipulation of metabolic pathways.

As stated above, 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. The information concerning human diseases originates from the MIM data base of Victor McKusick (3).

We do not assign EC numbers for newly characterized enzymes, this being the responsibility of the Nomenclature Committee of the IUBMB (NC-IUBMB). To contact the committee one should write to Prof. K.Tipton, Department of Biochemistry, Trinity College, Dublin 2, Republic of Ireland (tel. +353-1-702 1608; fax +353-1-677 2400; email ktipton@vax1.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. The commission regularly sends updates and additions to the nomenclature so that they can be integrated into the data bank in a timely manner.

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. These 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 (4) nucleotide sequence databases. Two sample ENZYME entries are shown below.

```
ID 1.14.17.3
DE PEPTIDYLGLYCINE MONOOXYGENASE.
AN PEPTIDYL ALPHA-AMIDATING ENZYME.
AN PEPTIDYLGLYCINE 2-HYDROXYLASE.
CA PEPTIDYLGLYCINE + ASCORBATE + O(2) = PEPTIDYL(2-
CA HYDROXYGLYCINE) + DEHYDROASCORBATE + H(2)O.
CF COPPER.
CC -!- PEPTIDYLGLYCINES WITH A NEUTRAL AMINO ACID RESIDUE IN
CC THE PENULTIMATE POSITION ARE THE BEST SUBSTRATES FOR
CC THE ENZYME.
CC -!- THE ENZYME ALSO CATALYZES THE DISMUTATION OF THE
CC PRODUCT TO GLYOXYLATE AND THE CORRESPONDING DESGLYCINE
CC PEPTIDE AMIDE.
DR P10731, AMD_BOVIN ; P19021, AMD_HUMAN ; P14925, AMD_RAT ;
DR P08478, AMD1_XENLA; P12890, AMD2_XENLA;
//
ID 2.3.1.43
DE PHOSPHATIDYLCHOLINE--STEROL 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
CC NUMBER OF STEROLS, INCLUDING CHOLESTEROL, CAN ACT AS
CC ACCEPTOR.
DI NORUM DISEASE; MIM:245900.
DI FISH-EYE DISEASE; MIM:136120.
DR P10480, GCAT_AERHY; P04180, LCAT_HUMAN; P16301, LCAT_MOUSE;
DR Q08758, LCAT_PAPAN; P30930, LCAT_PIG ; P18424, LCAT_RAT ;
//
```

PRACTICAL INFORMATION

Content of the current release

Release 19.0 of ENZYME (October 1995) contains information relevant to 3594 enzymes. The data file (enzyme.dat) requires ~1.1 Mb disk storage space. The database is distributed with a user's manual (enzuser.txt), a file describing the various classes, subclasses and subsubclasses of enzymes (enzclass.txt) and a file that describes how the database can be obtained (enzyme.get).

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

How to obtain ENZYME

ENZYME is distributed on CD-ROM by the EMBL Outstation–the European Bioinformatics Institute (EBI) (4). For all enquiries regarding subscription to and distribution of ENZYME one should contact The EMBL Outstation–The European Bioinformatics Institute, Hinxton Hall, Hinxton, Cambridge CB10 1RQ, UK (tel. +44 1223 494 400; fax +44 1223 494 468; email datalib@ebi.ac.uk).

ENZYME can be obtained from the EBI file server (5). Detailed instructions on how to make the best use of this service and, in particular, on how to obtain ENZYME can be obtained by query to the network address netserv@ebi.ac.uk

HELP

HELP ENZYME.

If you have access to a computer system linked to the Internet you can obtain ENZYME using ftp (File Transfer Protocol) from the following file servers:

EBI anonymous ftp server (ftp.ebi.ac.uk or 192.54.41.33);

NCBI Repository, National Library of Medicine, NIH, Washington, DC (ncbi.nlm.nih.gov or 130.14.20.1);

ExPASy (Expert Protein Analysis System) server, University of Geneva, Switzerland (expasy.hcuge.ch or 129.195.254.61);

National Institute of Genetics (Japan) ftp server (ftp.nig.ac.jp or 133.39.16.66).

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) (6) is also available on some of the above servers.

Browsing in ENZYME using the ExPASy WWW server

The most efficient and user friendly way to browse interactively in ENZYME is to use the World Wide Web (WWW) molecular biology server ExPASy (7). WWW is a global information retrieval system merging the power of worldwide networks, hypertext and multimedia. Through hypertext links it gives access to documents and information available on thousands of servers around the world. To access a WWW server one needs a WWW browser. Popular browsers available for most computer platforms include Mosaic™, developed at the National Center for Supercomputing Applications (NCSA) of the University of Illinois at Champaign (obtainable by anonymous ftp from ftp.ncsa.uiuc.edu), and Netscape Navigator™, from Netscape Communications Corp. (available from ftp.netscape.com). Using a WWW browser one has access to all the hypertext documents stored on the ExPASy server (as well as many other WWW servers) and also can make use of many sequence analysis software tools.

The ExPASy server may be accessed through its Uniform Resource Locator (URL, the addressing system defined in WWW), which is http://expasy.hcuge.ch/. You can directly access the 'top page' of the section of ExPASy that allows you to browse through the ENZYME database by opening the URL http://expasy.hcuge.ch/sprot/enzyme.html.

REFERENCES

1 NC-IUBMB (1992) *Enzyme Nomenclature, Recommendations of the Nomenclature Committe of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes*. Academic Press, New York, NY.
2 Bairoch,A. and Boeckmann,B. (1994) *Nucleic Acids Res*., **22**, 3578–3580.
3 McKusick,V.A. (1994) *Mendelian Inheritance in Man. Catalogs of Autosomal Dominant, Autosomal Recessive, and X-linked Phenotypes*, 11th Edn. Johns Hopkins University Press, Baltimore, MD.
4 Emmert,D.B., Stoehr,P.J., Stoesser,G. and Cameron,G.N. (1994) *Nucleic Acids Res*., **22**, 3445–3449.
5 Stoehr,P.J. and Omond,R.A. (1989) *Nucleic Acids Res*., **17**, 6763–6764.
6 Benson,D.A., Boguski,M.S., Lipman,D.J. and Ostell,J. (1990) *Genomics*, **6**, 389–391.
7 Appel,R.D., Bairoch,A. and Hochstrasser,D.F. (1994) *Trends Biochem. Sci*., **19**, 258–260.