304–305 *Nucleic Acids Research, 2000, Vol. 28, No. 1*

The ENZYME database in 2000

Amos Bairoch*

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

Received October 12, 1999; Accepted October 13, 1999

ABSTRACT

The ENZYME database is a repository of information related to the nomenclature of enzymes. In recent years it has became an indispensable resource for the development of metabolic databases. The current version contains information on 3705 enzymes. It is available through the ExPASy WWW server (http://www.expasy.ch/enzyme/).

INTRODUCTION

ENZYME 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:

- EC number;
- Recommended name;
- Alternative names (if any);
- Catalytic activity;
- Cofactors (if any);
- Pointers to the SWISS-PROT (2) protein sequence entrie(s) that correspond to the enzyme (if any);
- Pointers to the PROSITE (3) entrie(s) describing the protein familie(s) of which the enzyme is a member (if any);
- Pointers to human disease(s) (4) associated with a deficiency of the enzyme (if any).

We believe that the ENZYME database 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 the 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 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 archaebacterial genomes (6,7).

The main source for the data in ENZYME 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 ENZYME and SWISS-PROT is of benefit to both resources as it allows updates and corrections to be propagated efficiently between them.

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) and EMBL (8) 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-
     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 + a sterol = a 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 reac-
     tions 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 ;
//
```

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

PRACTICAL INFORMATION

Content of the current release

Release 25.0 of ENZYME (July 1999) contains information on 3705 enzymes. The data file (ENZYME.DAT) requires ~1.5 Mb of disk storage space. The database is distributed with a user's 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). 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 SWISS-PROT or TrEMBL 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. On October 1999 a cumulative total of 60 million connections was attained. Its address is http://www.expasy.ch/

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

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 (http://www.expasy.ch/cgi-bin/search-biochem-index), 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.

How to obtain ENZYME

You can obtain ENZYME by FTP from ftp.expasy.ch or 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

ENZYME is distributed with a form that can be used to fill in the information necessary for the NC-IUBMB to assign an EC number (see http://www.expasy.ch/sprot/enz_new_form.html). A separate form is available to send updates or corrections (see 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 database in a timely manner.

REFERENCES

1. Enzyme Nomenclature (1992) Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes, NC-IUBMB. Academic Press, New York, NY.
2. Bairoch,A. and Apweiler,R. (1999) *Nucleic Acids Res.*, **27**, 49–54. Updated article in this issue: *Nucleic Acids Res*. (2000), **28**, 45–48.
3. Hofmann,K., Bucher,P., Falquet,L. and Bairoch,A. (1999) *Nucleic Acids Res.*, **27**, 215–219.
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., Tuli,M.A., Lopez,R. and Sterk,P. (1999) *Nucleic Acids Res.*, **27**, 18–24. Updated article in this issue: *Nucleic Acids Res*. (2000), **28**, 19–23.
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., Ouellette,B.F.F., Rapp,B.A. and Wheeler,D.L. (1999) *Nucleic Acids Res.*, **27**, 12–17. Updated article in this issue: *Nucleic Acids Res*. (2000), **28**, 15–18.