Metabolic Engineering 4, 111–113 (2002)
doi:10.1006/mben.2001.0222, available online at http://www.idealibrary.com on IDEAL®

WEB SITE REVIEW

Review of the Biocatalysis/Biodegradation Database (UM-BBD)

Anthony P. Burgard and Costas D. Maranas[1]

Department of Chemical Engineering, Pennsylvania State University, University Park, Pennsylvania 16802

Received January 22, 2002; accepted January 22, 2002

INTRODUCTION

The University of Minnesota Biocatalysis/Biodegradation Database (Ellis *et al.*, 2001) is an online compilation of microbial catabolic enzymes, reactions, and pathways for primarily synthetic organic chemical compounds. This database can be accessed free of charge at http://umbbd.ahc.umn.edu/. As of January 2002, the database encompassed over 113 pathways containing 774 reactions, 696 compounds, and 476 enzymes present in over 306 microorganisms. This information is directly applicable toward enhancing the understanding of biocatalysis leading to specialty chemical manufacture and the biodegradation of environmental pollutants (Ellis *et al.*, 1999). Catabolic routes for compounds having any of the 50 functional groups shown in Table 1 are provided. Unlike pathway databases such as EcoCyc (Karp *et al.*, 2002a), MetaCyc (Karp *et al.*, 2002b), and Kegg (Kanehisa *et al.*, 2002), which focus primarily on intermediary metabolism, the UM-BBD is a key resource for biodegradation pathway information and is recently evolving to include the prediction of specialized catabolic routes for new compounds (Ellis *et al.*, 2000).

TABLE 1

List of 50 Functional Groups with Biodegradation Routes in the UM-BBD

Alcohol	Nitrate ester
Aldehyde	Nitrile
Alkane, primary	Nitro
Alkane, secondary	O-heterocyclic ring
Alkane, tertiary	Organoarsenical
Alkene	Organohalide
Alkyne	Organomercurial
Amide	Organophosphate ester
Amine, primary	Organosilicon
Amine, secondary	Organotin
Amine, tertiary	Oxime
Bicycloaliphatic ring	Oxygen ether
Biphenyl-type benzenoid ring	Peroxide
Carboxylic acid	Phosphinic acid
Carboxylic acid ester	Phosphonic acid
Carboxylic thioester	Polycyclic aromatic hydrocarbon
Cyanamide	S-heterocyclic ring
Cycloaliphatic ring	Sulfate ester
Diazo	Sulfonic acid
Epoxide	Thiocyanate
Ketone	Thioether
Methane	Thioketone
Monocyclic aromatic hydrocarbon	Thiol
N-heterocyclic ring, saturated	Thiophosphate ester
N-heterocyclic ring, unsaturated	Tricycloaliphatic ring

DESCRIPTION OF UM-BBD

The UM-BBD Web page contains links to 113 pathway maps by means of a scrollable list and includes a link to the search page. The database can be queried (i) for specific compounds via name, CAS number, or chemical formula; (ii) for enzymes by means of name or EC code; or (iii) for specific microorganisms. In addition, the complete collection of entries in the database can be browsed by compounds, enzymes, reactions, pathways, or microorganisms from the search page. Information in the database is accessed via four different formats: a compound page, an enzyme page, a reaction page, or a pathway map page. The pathway map pages illustrate the catabolic route(s) of 113 biochemical compounds and contain a short description explaining the significance of the pathway. All pathways are provided in a text format while a graphic representation is also available for most instances. An example of the text and graphic formats of the propylene degradation pathway from the UM-BBD database is shown in Fig. 1. The text-based pathways provide links to each individual compound, enzyme, or

[1] To whom correspondence and reprint requests should be addressed. Fax: (814)-865-7846. E-mail: costas@psu.edu.

1096-7176/02 $35.00

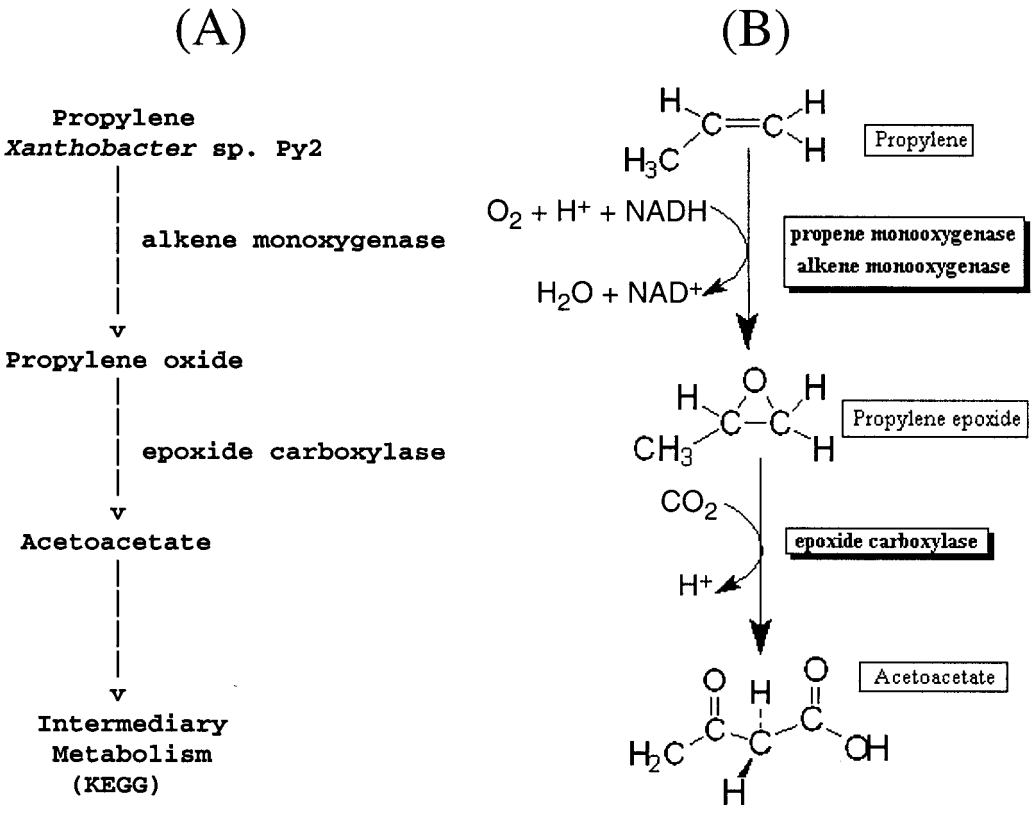

FIG. 1. (A) Text-based and (B) graphics-based representations of the biodegradation pathway for propylene obtained from the UM-BBD. Acetoacetate is the entry compound into intermediary metabolism.

reaction in the pathway as well as a list of organisms known to initiate the pathway. Also provided is a link to a KEGG (Kanehisa *et al.*, 2002) pathway graphic to show the user where the "end" compound enters intermediary metabolism.

Reaction pages include a text depiction of the reactions and links to graphics of the reactions, mechanisms of the reactions (if available), and the associated enzyme or compound pages. External links to the enzyme pages in the ligand (Goto *et al.*, 2002) chemical database and in the ENZYME (Bairoch, 2000) database on the ExPASy server are also available. Queries of the Medline (Wheeler *et al.*, 2001) and GenBank (Benson *et al.*, 2002) databases can also be carried out. As with the reaction pages, the enzyme pages include links to the Ligand database and the ENZYME database along with a search of the GenBank database. In addition, the enzyme pages include links to the corresponding UM-BBD reaction pages, links to enzyme pages in the BRENDA (Schomburg *et al.*, 2002) enzyme database and the IUBMB enzyme nomenclature database (http://www.chem.qmw.ac.uk/iubmb/enzyme/) for cases when a four-digit EC code exists, and searches of the GenPept and the Protein Data Bank (Westbrook *et al.*, 2002) wherever possible. Compound pages include general information such as chemical structure and molecular weight but also include links to other resources such as the National Toxicology Program (http://ntp-server.niehs.nih.gov/). Links to UM-BBD reaction pages for all reactions where the compound is a substrate or reactant are also provided. A more thorough overview of UM-BBD can be found in Ellis *et al.* (2002).

STRENGTHS AND WEAKNESSES

The key strength of the UM-BBD is its focus on specialized enzymes and genes not found in many databases of intermediary metabolism. This fact suggests utility not only in functional genomics, especially when dealing with less common genes, but also in predicting novel biosynthetic routes for specialty chemical manufacturing. The database's focus on functional group transformations enables the user to make educated hypotheses regarding the fate or environmental acceptability of new chemicals. In addition, numerous static and dynamic links to other

Metabolic Engineering **4**, 111–113 (2002)
doi:10.1006/mben.2001.0222

databases afford the user multiple opportunities to capture information outside of its scope. We found particularly useful the links to the KEGG pathway graphics that illustrate where the UM-BBD compounds fit into intermediary metabolism. Other strong points include the exceptional readability of the handcrafted mechanism, reaction, and pathway graphics and the user-friendliness of the entire website accompanied by excellent online documentation. One potential weakness, or perhaps strength depending on the user's perspective, is that unlike pathway representations in databases such as EcoCyc (Karp *et al.*, 2002a) or MetaCyc (Karp *et al.*, 2002b), complete UM-BBD pathways may not necessarily be present in their entirety within a single organism. Nevertheless, they do depict plausible metabolism and the author(s) names are listed for each contribution providing the user with an initial contact for questions regarding specifics of a particular pathway. In summary, the UM-BBD is an excellent online reference source for compounds, enzymes, reactions, and pathways for specialized metabolism in biocatalysis and biodegradation.

REFERENCES

Bairoch, A. (2000). The ENZYME database in 2000. *Nucleic Acids Res.* **28**, 304–305.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A., and Wheeler, D. L. (2002). GenBank. *Nucleic Acids Res.* **30**, 17–20.

Ellis, L. B. M., Hershberger, C. D., and Wackett, L. P. (1999). The University of Minnesota Biocatalysis/Biodegradation Database: Specialized metabolism for functional genomics. *Nucleic Acids Res.* **27**, 373–376.

Ellis, L. B. M., Hershberger, C. D., and Wackett, L. P. (2000). The University of Minnesota Biocatalysis/Biodegradation Database: Microorganisms, genomics, and prediction. *Nucleic Acids Res.* **28**, 377–379.

Ellis, L. B. M., Hershberger, C. D., Bryan, E. M., and Wackett, L. P. (2001). The University of Minnesota Biocatalysis/Biodegradation Database: Emphasizing enzymes. *Nucleic Acids Res.* **29**, 340–343.

Goto, S., Okuno, Y., Hattori, M., Nishioka, T., and Kanehisa, M. (2002). LIGAND: Database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res.* **30**, 402–404.

Kanehisa, M., Goto, S., Kawashima, S., and Nakaya, A. (2002). The KEGG databases at GenomeNet. *Nucleic Acids Res.* **30**, 42–46.

Karp, P. D., Riley, M., Saier, M., Paulsen, I. T., Collado-Vides, J., Paley, S. M., Pellegrini-Toole, A., Bonavides, C., and Gama-Castro, S. (2002). The EcoCyc Database. *Nucleic Acids Res.* **30**, 56–58.

Karp, P. D., Riley, M., Paley, S. M., and Pellegrini-Toole, A. (2002). The MetaCyc Database. *Nucleic Acids Res.* **30**, 59–61.

Schomburg, I., Chang, A., and Schomburg, D. (2002). BRENDA, enzyme data and metabolic information. *Nucleic Acids Res.* **30**, 47–49.

Westbrook, J., Feng, Z., Jain, S., Bhat, T. N., Thanki, N., Ravichandran, V., Gilliland, G. L., Bluhm, W. F., Weissig, H., Greer, D. S., Bourne, P. E., and Berman, H. M. (2002). The Protein Data Bank: Unifying the archive. *Nucleic Acids Res.* **30**, 245–248.

Wheeler, D. L., Church, D. M., Lash, A. E., Leipe, D. D., Madden, T. L., Pontius, J. U., Schuler, G. D., Schriml, L. M., Tatusova, T. A., Wagner, L., and Rapp, B. A. (2001). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **29**, 11–16.